LETTER FROM THE EDITOR

As William Dunham writes in the first article, the Möbius function "seems to conceal numerical information rather than reveal it." Dunham reviews Möbius' use in analysis of his namesake function and Euler's anticipation of the function. In the next article, Tom Edgar considers visually appealing dissection proofs of staircase series, proving known results in an intuitive and easy-to-reconstruct way. Francisco Sánchez and José María Sanchis follow by applying Darboux sums to sum the alternating harmonic series.

Two articles on group theory are next. Chase Saucier uses a combinatorial approach to prove Wilson's theorem for finite Abelian groups and Robert Heffernan, Des MacHale, and Brendan McCann are motivated by Cayley's theorem to ask five elementary questions about embeddings of finite groups, which they answer for groups of order 15 or less.

In a 1-page note, Mark Dalthorp uses a simple observation about integrals to derive the Taylor series of e^x, $\sin x$, and $\cos x$. Using multivariable calculus, Joseph Previte and Michelle Previte show that most triangles in n-dimensional space are acute for $n \geq 3$, a result not true for $n = 2$, which was considered by Lewis Carroll in 1893 and further develop by Richard Guy in THIS MAGAZINE in 1993. For triangles in the plane, Stefan Catiou and Allan Berrelle introduce a new curve which is the envelope of all lines that bisect the triangle's perimeter. They go on to describe the curve both analytically and geometrically.

Michael Lord constructs different representations of the power series for which the nth coefficient is the reciprocal of the nth Fibonacci number. He then compares the convergence of the resulting power series. This demonstrates that "the goal of the numerical analyst is to do the analysis first and then do the computation."

In a pictorally motivated proof, Roger Nelsen shows that even perfect numbers end in 6 or 28.

For a simple family of trinomials, the roots form intriguing patterns in the complex plane that lead to a conjecture about the number of roots that lie inside the unit circle. Using basic geometry, number-theoretic calculations, and an application of Rouché's theorem, Michael Brilleslyper and Lisbeth Schaubroeck prove the conjecture for a special case.

As with every issue in 2018, there is a Partiti puzzle (see the February issue for an introduction to the problem). And, as is often the case, the popular Problems and Solutions and the Reviews complete the issue.

Finally, I offer my thanks to two departing editorial board members (David Scott and Julie Beier) for their service. David steps down after 2^4 years, while Julie exits after 2^2 years. Best of luck to Julie in her new non-academic position.

Michael A. Jones, Editor

The Early (and Peculiar) History of the Möbius Function

WILLIAM DUNHAM
Bryn Mawr College
Bryn Mawr, PA 19010
bdunham@brynmawr.edu

We begin with this challenge from analysis: for $-1 < x < 1$, determine the exact value of the infinite series

$$\frac{x}{1-x} - \frac{x^2}{1-x^2} - \frac{x^3}{1-x^3} - \frac{x^5}{1-x^5} + \frac{x^6}{1-x^6} - \frac{x^7}{1-x^7} + \frac{x^{10}}{1-x^{10}} - \frac{x^{11}}{1-x^{11}} - \frac{x^{13}}{1-x^{13}} + \frac{x^{14}}{1-x^{14}} + \frac{x^{15}}{1-x^{15}} - \frac{x^{17}}{1-x^{17}} - \cdots.$$

The pattern here is anything but clear, as signs flip-flop in strange fashion and certain terms are unexpectedly missing. The sum turns out to be a simple one, but finding it, as we shall see, requires familiarity with something called the "Möbius function."

This function shows up in any comprehensive text on the theory of numbers. After chapters on primes and congruences, on the Euclidean algorithm and Diophantine equations, such a book will sooner or later introduce the Möbius function. It tends to appear alongside its number-theoretic cousins: the sigma, tau, and phi functions. The sigma function, $\sigma(k)$, sums all the divisors of k; the tau function, $\tau(k)$, counts all the divisors of k; and the phi function, $\phi(k)$, counts the numbers less than k and relatively prime to it. These three have an obvious utility in number theory.

The Möbius function, by contrast, seems neither useful nor obvious. It is denoted by $\mu(k)$ and defined, for a whole number k, as follows:

(a) $\mu(1) = 1$.
(b) $\mu(k) = 0$ if k is divisible by the square of some prime.
(c) $\mu(k) = (-1)^r$ if k is the product of r different primes.

Thus, for the first ten numbers, $\mu(1) = 1$; $\mu(2) = \mu(3) = \mu(5) = \mu(7) = -1$ because each is a prime; $\mu(4) = \mu(8) = \mu(9) = 0$ because these are divisible, respectively, by 2^2, 2^2, and 3^2; and $\mu(6) = \mu(10) = (-1)^2 = 1$ because these are the products of two different primes.

At first glance, these values seem uninformative. Unlike the sigma, Möbius's function doesn't *sum* anything. Unlike the tau and phi, Möbius's function doesn't *count* anything. Because every fourth number is divisible by $2^2 = 4$, the Möbius function takes the value zero more than a quarter of the time. And the Möbius function exhibits strange runs, like $\mu(33) = \mu(34) = \mu(35) = 1$ or $\mu(242) = \mu(243) = \mu(244) = \mu(245) = 0$. In such cases, the function seems to conceal numerical information rather

Math. Mag. **91** (2018) 83–91. doi:10.1080/0025570X.2017.1413921 © Mathematical Association of America
MSC: Primary 11-03, Secondary 01A50; 01A55

than to reveal it. Whereas the sigma, tau, and phi functions are "natural" concepts, the Möbius function comes across as something of an oddity. What possible interest could it hold?

Anyone who reads a bit further in that number theory text will discover that the Möbius function is not only important but highly so. The deeper one digs, the more significant the concept becomes. For instance, the prime number theorem can be recast in terms of the Möbius function. The (famously unresolved) Riemann hypothesis can be recast in a similar fashion. And there is even an application to quantum physics called the "free Riemann gas model of supersymmetry" that involves – yes! – the Möbius function. Indeed, this is a concept to be reckoned with.

Where did the idea come from?

August Ferdinand Möbius
(1790–1868)

Here the story holds its surprises, for the function's namesake, August Ferdinand Möbius (1790–1868), did not introduce it via the conditions (a)–(c) above. Rather, the concept appeared in his 1832 paper "*Ueber eine besondere Art von Umkehrung der Riehen*" (On a special kind of inversion of series) about a subject that, at first glance, had nothing to do with number theory [**4**, pp. 105–123].

Möbius began with a real function defined by the power series

$$f(x) = a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots,$$

which he sought to invert into an expression for x of the form

$$x = b_1f(x) + b_2f(x^2) + b_3f(x^3) + b_4f(x^4) + \cdots.$$

The challenge was to determine the values of the b_k based on the known values of the a_k.

To simplify things, Möbius set all the a_k equal to 1, thereby turning $f(x)$ into the geometric series

$$f(x) = x + x^2 + x^3 + x^4 + \cdots.$$

This, of course, sums to $x/(1 - x)$ for $-1 < x < 1$. His inversion therefore amounted to writing

$$\begin{aligned} x &= b_1f(x) + b_2f(x^2) + b_3f(x^3) + b_4f(x^4) + \cdots \\ &= b_1[x + x^2 + x^3 + x^4 + \cdots] + b_2[x^2 + x^4 + x^6 + x^8 + \cdots] \end{aligned}$$

$$+ b_3[x^3 + x^6 + x^9 + x^{12} + \cdots] + b_4[x^4 + x^8 + x^{12} + x^{16} + \cdots]$$
$$+ b_5[x^5 + x^{10} + x^{15} + x^{20} + \cdots] + \cdots.$$

Collecting powers of x, Möbius arrived at

$$x = b_1 x + [b_1 + b_2]\, x^2 + [b_1 + b_3]\, x^3 + [b_1 + b_2 + b_4]\, x^4 + [b_1 + b_5]\, x^5$$
$$+ [b_1 + b_2 + b_3 + b_6]\, x^6 + [b_1 + b_7]\, x^7 + [b_1 + b_2 + b_4 + b_8]\, x^8 + \cdots.$$

Matching coefficients of x, he saw that $b_1 = 1$. Using this fact and equating coefficients from consecutively higher powers, he got:

$$0 = 1 + b_2, \quad 0 = 1 + b_3, \quad 0 = 1 + b_2 + b_4, \quad 0 = 1 + b_5,$$
$$0 = 1 + b_2 + b_3 + b_6, \text{ and so on.} \tag{1}$$

Thus,

$$b_2 = -1,\ b_3 = -1,\ b_4 = -1 - b_2 = 0,\ b_5 = -1,\ b_6 = -1 - b_2 - b_3 = 1,$$
$$b_7 = -1,\ b_8 = -1 - b_2 - b_4 = 0, \text{ etc.}$$

Möbius wished to calculate the b_k by identifying a pattern in these coefficients rather than by solving a string of ever-longer equations. This he did, although here too his approach might strike the modern reader as strangely complicated.

First, Möbius observed, perhaps with some surprise, that “one finds these coefficients … are either −1, 0, or 1.” [4, p. 110] He then showed that they were generated by the following rules.

Rule 1: If p is prime, the pertinent equation to emerge from (1) is obviously $0 = 1 + b_p$. Hence $b_p = -1$ for all primes p.

Rule 2: If $k = p \cdot q$, where p and q are different primes, then the equation from (1) will be

$$0 = 1 + b_p + b_q + b_{pq}.$$

Rearranging terms and adding $b_p \cdot b_q$ to both sides of the equation, Möbius deduced that

$$-b_{pq} + b_p \cdot b_q = 1 + b_p + b_q + b_p \cdot b_q = \left(1 + b_p\right)\left(1 + b_q\right) = 0$$

by Rule 1. Thus $b_{pq} = b_p \cdot b_q = (-1)^2 = 1$. In like fashion, he saw that if k is the product of r different primes, then $b_k = (-1)^r$.

Rule 3: If $k = p^2$ for some prime p, then the corresponding equation from (1) will be

$$0 = 1 + b_p + b_{p^2},$$

and so $b_{p^2} = -(1 + b_p) = 0$ by Rule 1. Similarly, if $n = p^3$, he got $0 = 1 + b_p + b_{p^2} + b_{p^3}$ and so $b_{p^3} = -(1 + b_p) - b_{p^2} = 0$. The same outcome holds for any higher power of p.

Rule 4: Finally, Möbius considered the case where k is divisible by the square of a prime. For a simple example, we look at $k = p^2 q$ where p and q are different primes. From (1) we have

$$0 = 1 + b_p + b_{p^2} + b_q + b_{pq} + b_{p^2 q}.$$

Knowing that $b_{pq} = b_p \cdot b_q$ and adding $b_{p^2} \cdot b_q = 0$ to both sides, he got

$$-b_{p^2 q} + b_{p^2} \cdot b_q = 1 + b_p + b_{p^2} + b_q + b_p \cdot b_q + b_{p^2} \cdot b_q$$
$$= \left(1 + b_p + b_{p^2}\right)\,(1 + b_q) = 0$$

by Rule 1. Thus $b_{p^2 q} = b_{p^2} \cdot b_q = 0 \cdot b_q = 0$ by Rule 3.

In this fashion, Möbius determined rules for calculating his coefficients b_k. Of course, these yield precisely what is called $\mu(k)$ in today's textbooks. But it is interesting that, rather than simply *defining* a number theoretic concept, Möbius had *derived* it from the inversion of a geometric series.

We should say a word about the modern notation. This is not due to Möbius, who, as we just saw, used "b_k". Rather, it was the mathematician Franz Mertens (1840–1927) who, in an 1874 paper, introduced "μ", the Greek counterpart of "m," to denote this function [3]. We can read this choice as homage to "Möbius" … with the fortunate coincidence that it also celebrated "Mertens"!

Returning to the original inversion problem, we restate Möbius's conclusion as

$$x = b_1 f(x) + b_2 f(x^2) + b_3 f(x^3) + b_4 f(x^4) + \cdots = \sum_{k=1}^{\infty} b_k f(x^k) = \sum_{k=1}^{\infty} \mu(k) f(x^k),$$

or simply

$$x = \sum_{k=1}^{\infty} \frac{\mu(k)\, x^k}{1 - x^k}$$

because, as noted,

$$f(x) = \frac{x}{1-x}\ .$$

This, by the way, answers our opening question, for the infinite series

$$\frac{x}{1-x} - \frac{x^2}{1-x^2} - \frac{x^3}{1-x^3} - \frac{x^5}{1-x^5} + \frac{x^6}{1-x^6} - \frac{x^7}{1-x^7} + \frac{x^{10}}{1-x^{10}} - \frac{x^{11}}{1-x^{11}} - \cdots$$
$$= \sum_{k=1}^{\infty} \frac{\mu(k) x^k}{1-x^k} = x.$$

In short, our infinite series sums to x. What could be simpler?

So, it is tempting to conclude that the Möbius function first appeared in this paper from 1832 and was subsequently given the name of the author. Alas, that conclusion needs revision, for in 1748 Leonhard Euler (1707–1783) had stumbled upon the same idea, although in a very different fashion. This, by the way, was half a century before August Ferdinand Möbius was even born.

Leonhard Euler
(1707–1783)

Euler's discovery appeared in Chapter XV of his classic text, *Introductio in analysin infinitorum*. The title of this chapter translates as "On Series Which Arise from Products," and, as we shall see, it was aptly chosen [**1**, pp. 228–255].

First, for an arbitrary whole number n, we introduce the infinite quotient

$$M_n = \frac{1}{\left(1 - \frac{1}{2^n}\right)\left(1 - \frac{1}{3^n}\right)\left(1 - \frac{1}{5^n}\right)\left(1 - \frac{1}{7^n}\right)\cdots},$$

where the terms in the denominators run through the primes. (Note: Euler, who never employed subscripts, simply called this "M," but the subscript will prove useful below.) Because $1 + a + a^2 + a^3 + \cdots = 1/(1 - a)$, he expressed this as an infinite product of infinite series:

$$M_n = \left(1 + \frac{1}{2^n} + \frac{1}{4^n} + \frac{1}{8^n} + \cdots\right) \cdot \left(1 + \frac{1}{3^n} + \frac{1}{9^n} + \frac{1}{27^n} + \cdots\right) \cdot \left(1 + \frac{1}{5^n} + \frac{1}{25^n} + \cdots\right) \cdot \left(1 + \frac{1}{7^n} + \frac{1}{49^n} + \cdots\right)\cdots .$$

Upon multiplying these series, Euler concluded that

$$M_n = 1 + \frac{1}{2^n} + \frac{1}{3^n} + \frac{1}{4^n} + \frac{1}{5^n} + \frac{1}{6^n} + \frac{1}{7^n} + \frac{1}{8^n} + \frac{1}{9^n} + \frac{1}{10^n} + \cdots = \sum_{k=1}^{\infty} \frac{1}{k^n},$$

where every whole number k appears in one and only one denominator. This follows from the unique factorization of whole numbers into primes, the so-called fundamental theorem of arithmetic.

Perhaps more relevant to our story is the behavior of the reciprocal $1/M_n$, an infinite product that we shall denote by Q_n. Clearly,

$$Q_n = \left(1 - \frac{1}{2^n}\right)\left(1 - \frac{1}{3^n}\right)\left(1 - \frac{1}{5^n}\right)\left(1 - \frac{1}{7^n}\right)\cdots,$$

an expression whose product Euler sought to determine. We first observe that the primes occurring in these various denominators are different, so there is no way that something like

$$\frac{1}{9^n} = \frac{1}{3^n} \cdot \frac{1}{3^n} \quad \text{or} \quad \frac{1}{12^n} = \frac{1}{2^n} \cdot \frac{1}{2^n} \cdot \frac{1}{3^n}$$

could appear in the product. In other words, when the binomials of Q_n are multiplied, the resulting terms must look like $1/a^n$ where a is *not* divisible by the square of any prime. Such a number is called "square-free."

Thus, upon multiplying these binomials, Euler found that

$$Q_n = 1 - \frac{1}{2^n} - \frac{1}{3^n} - \frac{1}{5^n} + \frac{1}{6^n} - \frac{1}{7^n} + \frac{1}{10^n} - \frac{1}{11^n} - \frac{1}{13^n} + \frac{1}{14^n} + \frac{1}{15^n} - \cdots$$

The pattern here should look familiar. The signs obey Rules 1 to 4 of the Möbius function. Euler put it this way:

> "We note that the terms with primes, or products of three different primes, or any product of an odd number of different primes, appear with a negative sign. Those terms which are the product of two, four, six, or any even number of different primes, appear with a positive sign." [**1**, p. 230]

(To this we might add that terms divisible by the square of a prime do not appear at all.) As an example of the pattern, Euler observed that " ... the term $1/30^n$ appears with a negative sign, because 30 is the product of three different primes." [5]

This description exactly matches the rules Möbius would later generate. In modern notation, what Euler had found was

$$Q_n = \frac{1}{M_n} = \sum_{k=1}^{\infty} \frac{\mu(k)}{k^n}.$$

Next, Euler introduced an infinite quotient differing from M_n above only in the signs of the terms. We shall write it as

$$N_n = \frac{1}{\left(1+\frac{1}{2^n}\right)\left(1+\frac{1}{3^n}\right)\left(1+\frac{1}{5^n}\right)\left(1+\frac{1}{7^n}\right)\cdots}.$$

Letting R_n be the reciprocal of N_n, we see that

$$\begin{aligned} R_n &= \left(1+\frac{1}{2^n}\right)\left(1+\frac{1}{3^n}\right)\left(1+\frac{1}{5^n}\right)\left(1+\frac{1}{7^n}\right) \\ &= 1+\frac{1}{2^n}+\frac{1}{3^n}+\frac{1}{5^n}+\frac{1}{6^n}+\frac{1}{7^n}+\frac{1}{10^n}+\frac{1}{11^n}+\frac{1}{13^n}+\frac{1}{14^n}+\frac{1}{15^n}+\cdots \end{aligned}$$

This is the same series as that of Q_n above but with plus signs throughout. As before, only square-free terms appear in the denominators, so we can express this result as

$$R_n = \sum_{k=1}^{\infty} \frac{[\mu(k)]^2}{k^n}.$$

At this point, it is worth observing that Euler's derivations conformed to the fashion of the eighteenth century, when modern analytic rigor still lay far over the mathematical horizon. Later mathematicians would tidy up the logic of his results, but in this, as in so many other cases throughout his career, Euler's analytic intuition did not fail him.

Having generated these formulas, he was ready to specify values of n. To follow his line of attack, we recall three particular series whose sums were familiar to Euler.

- Since (at least) the previous century, the harmonic series was known to diverge. In Euler's day, this divergence was expressed as

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots = \infty.$$

- In 1734, Euler had evaluated the sum of reciprocals of the squares as

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

- Likewise, Euler had summed the reciprocals of the 4^{th} powers to get

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}.$$

Armed with these results, he returned to M_n, Q_n, and R_n above. For $n = 1$,

$$M_1 = 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\cdots = \infty$$

and so its reciprocal, Q_1, is given by

$$Q_1 = 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} + \frac{1}{6} - \frac{1}{7} + \frac{1}{10} - \frac{1}{11} - \frac{1}{13} + \frac{1}{14} + \frac{1}{15} - \cdots = \frac{1}{\infty} = 0.$$

In modern notation, this becomes the critical formula

$$\sum_{k=1}^{\infty} \frac{\mu(k)}{k} = 0.$$

Next, Euler let $n = 2$ to get

$$M_2 = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \frac{1}{7^2} + \cdots = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6},$$

and so

$$Q_2 = 1 - \frac{1}{2^2} - \frac{1}{3^2} - \frac{1}{5^2} + \frac{1}{6^2} - \frac{1}{7^2} + \frac{1}{10^2} - \frac{1}{11^2} - \frac{1}{13^2} + \frac{1}{14^2} + \frac{1}{15^2} - \cdots$$
$$= \frac{1}{M_2} = \frac{6}{\pi^2},$$

which we would write as

$$\sum_{k=1}^{\infty} \frac{\mu(k)}{k^2} = \frac{6}{\pi^2}.$$

This is a remarkable result, but Euler had one more trick up his sleeve. For $n = 4$, he knew that

$$M_4 = 1 + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \frac{1}{5^4} + \cdots = \frac{\pi^4}{90},$$

and so

$$\frac{M_2}{M_4} = \frac{\pi^2/6}{\pi^4/90} = \frac{15}{\pi^2}.$$

But the Ms and the Qs are reciprocals of one another, and so

$$\frac{15}{\pi^2} = \frac{M_2}{M_4} = \frac{Q_4}{Q_2} = \frac{\left(1 - \frac{1}{2^4}\right)\left(1 - \frac{1}{3^4}\right)\left(1 - \frac{1}{5^4}\right)\left(1 - \frac{1}{7^4}\right)\cdots}{\left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{3^2}\right)\left(1 - \frac{1}{5^2}\right)\left(1 - \frac{1}{7^2}\right)\cdots}$$
$$= \left(1 + \frac{1}{2^2}\right)\left(1 + \frac{1}{3^2}\right)\left(1 + \frac{1}{5^2}\right)\left(1 + \frac{1}{7^2}\right)\cdots = R_2,$$

because

$$\frac{\left(1 - \frac{1}{p^4}\right)}{\left(1 - \frac{1}{p^2}\right)} = \frac{\left(1 + \frac{1}{p^2}\right) \cdot \left(1 - \frac{1}{p^2}\right)}{\left(1 - \frac{1}{p^2}\right)} = \left(1 + \frac{1}{p^2}\right).$$

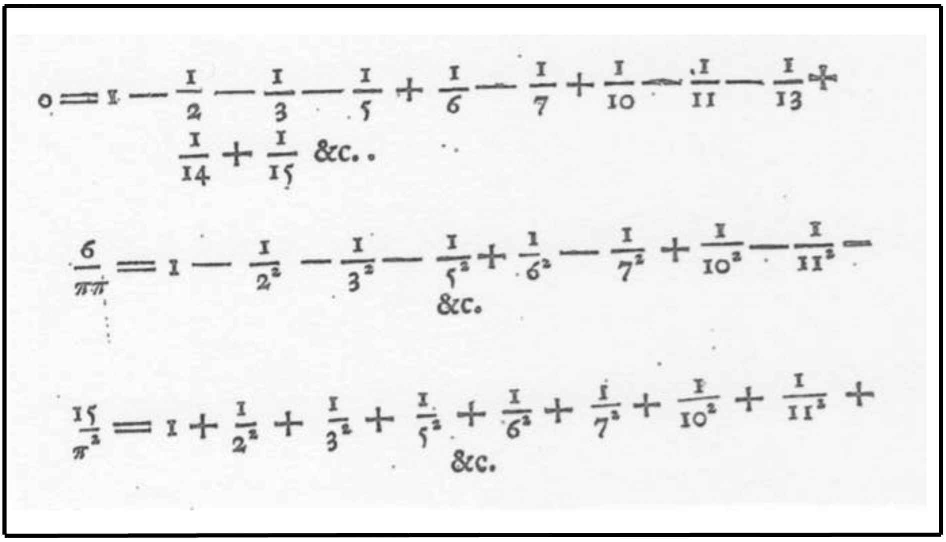
Consequently,

$$\frac{15}{\pi^2} = R_2 = \sum_{k=1}^{\infty} \frac{[\mu(k)]^2}{k^2}.$$

In words, this means that if we sum the reciprocals of the squares of the *square-free* whole numbers, we get

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{5^2}+\frac{1}{6^2}+\frac{1}{7^2}+\frac{1}{10^2}+\frac{1}{11^2}+\frac{1}{13^2}+\frac{1}{14^2}+\frac{1}{15^2}+\cdots$$
$$=1+\frac{1}{4}+\frac{1}{9}+\frac{1}{25}+\frac{1}{36}+\frac{1}{49}+\frac{1}{100}+\frac{1}{121}+\frac{1}{169}+\frac{1}{196}+\frac{1}{225}+\cdots=\frac{15}{\pi^2}.$$

These formulas involving the Möbius function can be spotted–albeit without the "μ" notation–in Euler's *Introductio* [2].



$$0=1-\frac{1}{2}-\frac{1}{3}-\frac{1}{5}+\frac{1}{6}-\frac{1}{7}+\frac{1}{10}-\frac{1}{11}-\frac{1}{13}+\frac{1}{14}+\frac{1}{15}\ \&c.$$

$$\frac{6}{\pi\pi}=1-\frac{1}{2^2}-\frac{1}{3^2}-\frac{1}{5^2}+\frac{1}{6^2}-\frac{1}{7^2}+\frac{1}{10^2}-\frac{1}{11^2}-\ \&c.$$

$$\frac{15}{\pi^2}=1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{5^2}+\frac{1}{6^2}+\frac{1}{7^2}+\frac{1}{10^2}+\frac{1}{11^2}+\ \&c.$$

From this last expression, it follows that the sum the reciprocals of the squares of those integers that are *not* square-free will be

$$\sum_{k=1}^{\infty}\frac{1}{k^2}-\sum_{k=1}^{\infty}\frac{[\mu\,(k)]^2}{k^2}=\frac{\pi^2}{6}-\frac{15}{\pi^2}.$$

That is,

$$\frac{1}{16}+\frac{1}{64}+\frac{1}{81}+\frac{1}{144}+\frac{1}{256}+\frac{1}{324}+\frac{1}{400}+\frac{1}{576}+\frac{1}{625}+\cdots=\frac{\pi^4-90}{6\pi^2}.$$

Look at this sum for a moment. It is exact. It is strange. It is astonishing. We have surely found our way into analytic territory where intuition is of no use whatever.

These wonderful results are examples of Euler being Euler, manipulating symbols with a gusto that can take one's breath away. In so doing, he not only anticipated the Möbius function but generated formulas more sophisticated than anything its namesake would discover eight decades later. Euler was, yet again, far ahead of his time.

With this, we conclude our story of a familiar number theoretic concept and its most peculiar ancestry. This tale reminds us–if we need reminding–that the history of mathematics can provide a host of unexpected rewards.

REFERENCES

[1] Euler, L. (1988). *Introduction to Analysis of the Infinite*. Vol. 1 (Blanton J., trans.) New York: Springer-Verlag.

[2] Euler, L. (1967). *Introductio in analysin infinitorum*. Vol. 1. Bruxelles: Impression Anastaltique Culture et Civilisation, p. 245.

[3] Mertens, F. (1874). Über einige asymptotische Gesatze der Zahlentheorie. *J. Für Die Reine Und Angewandte Mathematik*. 77:46–62.

[4] Möbius, A. F. (1832). Ueber eine besondere Art von Unkehrung der Riehen. *J. Für Die Reine Und Angewandte Mathematik*. 9. doi.org/10.1515/crll.1832.9.105

Summary. The Möbius function is a fixture of modern courses in number theory. It is usually traced back to an 1832 paper by August Ferdinand Möbius where the function unexpectedly arose in answer to an analytic, rather than a number theoretic, question. But perhaps more unexpected is that the function can be found in Leonhard Euler's classic text, *Introductio in analysis infinitorum*, from 1748. Besides presenting the origins of what might be called the "Euler/ Möbius" function, this article is a reminder that the history of mathematics holds its share of surprises.

WILLIAM DUNHAM (MR Author ID: 238997) is a research associate in mathematics at Bryn Mawr College. He has authored four books on the history of mathematics, among which are *Euler: The Master of Us All* (MAA, 1999) and *The Calculus Gallery* (Princeton, 2005), and he co-edited, with Don Albers and Jerry Alexanderson, *The G. H. Hardy Reader* (MAA, 2015). From 2014–2016, Dunham was the MAA's George Pólya Lecturer.

Partiti Puzzle

13	9	16	11	14	11
15	8	3	3	9	11
10	8	6	11	8	13
12	9	5	6	9	15
7	7	8	4	5	16
20	11	13	18	9	15

How to play. In each cell, place one or more distinct integers from 1 to 9 so that they sum to the value in the top left corner. No integer can be used more than once in horizontally, vertically, or diagonally adjacent cells.

For an introduction to the Partiti puzzle, see Caicedo, A. E., Shelton, B. (2018). Of puzzles and partitions: Introducing Partiti. *Math. Mag.* 91(1):20–23. The solution is on page 141.

—contributed by Lai Van Duc Thinh,
Vietnam; fibona2cis@gmail.com

Math. Mag. **91** (2018) 91. doi:10.1080/0025570X.2017.1421371

Staircase Series

TOM EDGAR
Pacific Lutheran University
Tacoma, WA 98447
edgartj@plu.edu

Swain [5] uses what he refers to as "Gabriel's Staircase" to prove visually that $\sum_{k=1}^{\infty} kr^k = \frac{r}{(1-r)^2}$ for $0 < r < 1$. In this note, we demonstrate how to modify his diagram to provide recursive wordless computations of series of the form $\sum_{k=0}^{\infty} k^n r^k$ where $n \geq 1$, which we will thus refer to as *staircase series*. For instance, in Figure 1 we vary the "stair heights" according to successive odd integers, resulting in a visual proof that

$$\sum_{k=1}^{\infty} k^2 r^k = \frac{1}{1-r} \sum_{k=1}^{\infty} (2k-1) r^k.$$

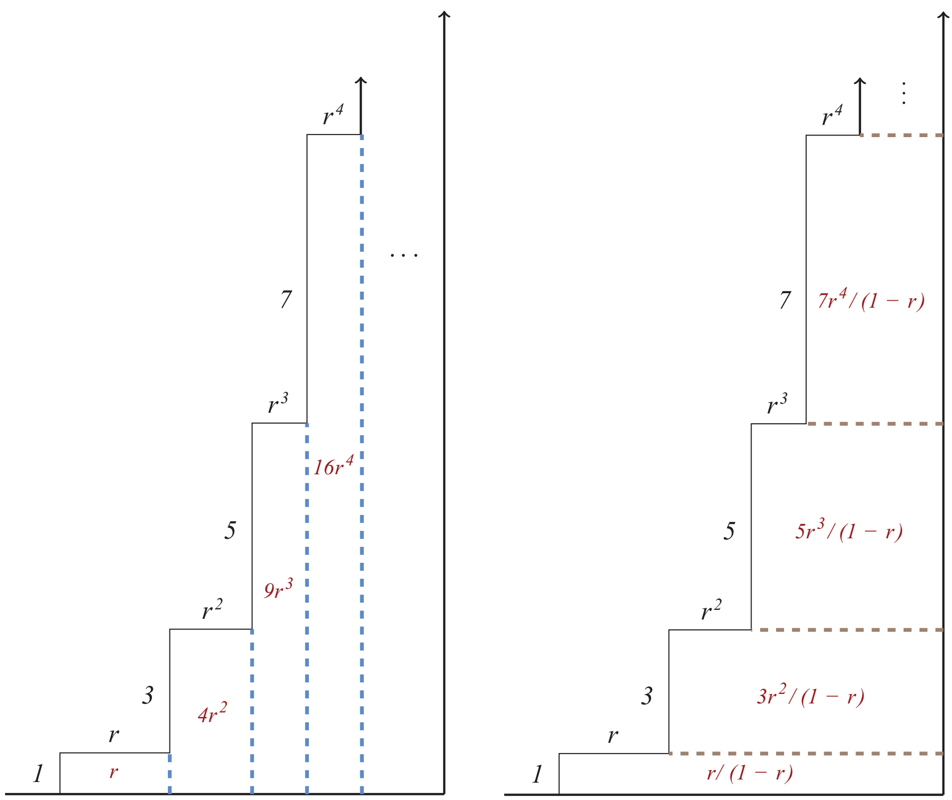


Figure 1 Staircase series with odd integer stair heights.

Math. Mag. **91** (2018) 92–95. doi:10.1080/0025570X.2017.1415584
MSC: Primary 40A05
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

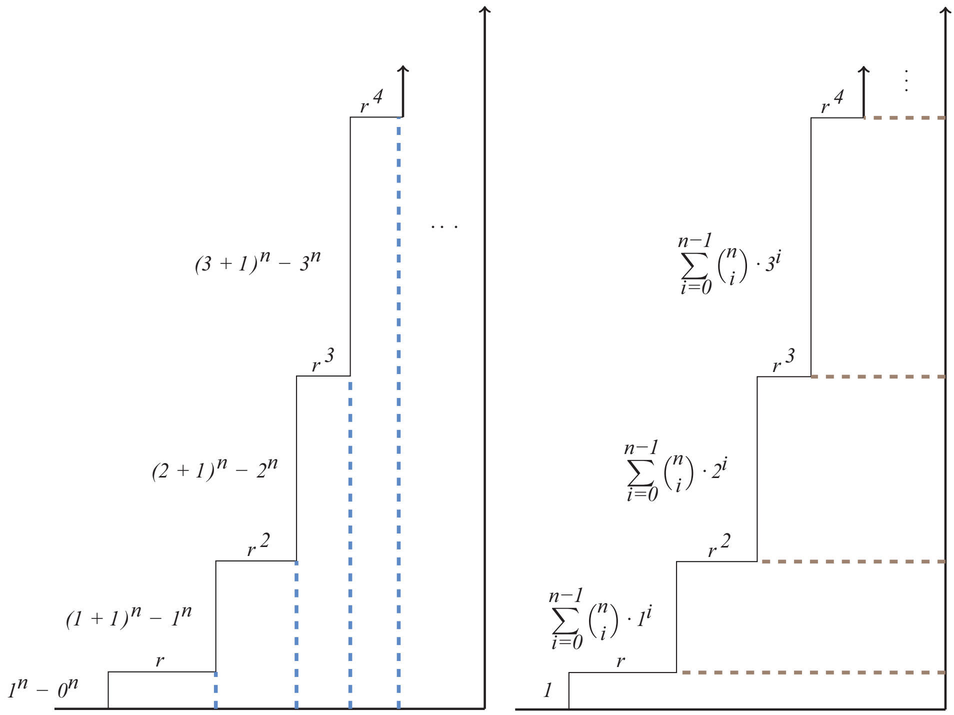


Figure 2 A general staircase series diagram.

Of course, if we couple this with Swain's result (and some elementary algebra) we obtain a visual proof that

$$\sum_{k=1}^{\infty} k^2 r^k = \frac{1}{1-r}\cdot\left(2\cdot\frac{r}{(1-r)^2} - \frac{r}{1-r}\right) = \frac{r^2+r}{(1-r)^3}.$$

This extension of Swain's diagram works because every square, n^2, is the sum of the first consecutive n odd integers so that the difference of consecutive squares is an odd number (and every odd number is the difference of two consecutive squares). This realization about Figure 1 helps us see that we can further generalize Swain's diagram to provide a reduction formula for the series $\sum_{k=0}^{\infty} k^n r^k$, which can be found in [2] or [3]. To build this diagram, we simply label "stair heights" with the differences of successive powers of n, which we have done in Figure 2. According to the binomial theorem, differences of successive powers of n can be computed in terms of binomial coefficients. Consequently, investigation of Figure 2 leads to the following reduction formula:

$$\sum_{k=0}^{\infty} k^n r^k = \sum_{k=0}^{\infty}\left(\sum_{i=0}^{n-1}\binom{n}{i}k^i\cdot\sum_{j=k+1}^{\infty} r^j\right) = \sum_{k=0}^{\infty}\left(\sum_{i=0}^{n-1}\binom{n}{i}k^i\cdot\frac{r^{k+1}}{1-r}\right).$$

Elementary algebraic manipulation (no derivatives needed) shows that for $n \geq 1$

$$\sum_{k=0}^{\infty} k^n r^k = \frac{r}{1-r}\cdot\sum_{i=0}^{n-1}\left(\binom{n}{i}\sum_{k=0}^{\infty} k^i r^k\right). \tag{1}$$

Exercise 1. Find a closed form for $\sum_{k=0}^{\infty} k^3 r^k$ in terms of r when $0 < r < 1$.

					1					
				1		1				
			1		4		1			
		1		11		11		1		
	1		26		66		26		1	
1		57		302		302		57		1

Figure 3 The triangle of Eulerian numbers up to row 5.

According to a variant of the Carlitz identity (see [1] or [4]), for $n \geq 1$, each staircase series has a closed form

$$\sum_{k=0}^{\infty} k^n r^k = \frac{r \cdot p_n(r)}{(1-r)^{n+1}},$$

where $p_n(r)$ is a polynomial in r of degree $n-1$. This identity was known to Euler; in fact the polynomials, $p_n(r)$, are known as the Eulerian polynomials, and their coefficients are given by the Eulerian numbers, which form a triangular array of numbers similar to the binomial coefficients. More precisely,

$$p_n(r) = \sum_{i=0}^{n-1} \left\langle {n \atop i} \right\rangle r^i,$$

where $\left\langle {n \atop i} \right\rangle$ is entry i in row n of the triangle of Eulerian numbers.

The Eulerian number $\left\langle {n \atop i} \right\rangle$ counts the number of permutations of $\{1, \ldots, n\}$ with i descents, where a descent of a permutation, π, is a position j such that $\pi(j) > \pi(j+1)$. The triangle of Eulerian numbers can be obtained recursively as follows:

$$\left\langle {n \atop i} \right\rangle = (n-i)\left\langle {n-1 \atop i-1} \right\rangle + (i+1)\left\langle {n-1 \atop i} \right\rangle,$$

where $\left\langle {1 \atop 0} \right\rangle = 1$ and $\left\langle {n \atop i} \right\rangle = 0$ if $i \geq n$ or if $i < 0$. Rows 1 through 5 of the triangle of Eulerian numbers are shown in Figure 3. For instance, we see that $p_3(r) = r^2 + 4r + 1$, and thus the answer to Exercise 1 is

$$\sum_{k=0}^{\infty} k^3 r^k = \frac{r^3 + 4r^2 + r}{(1-r)^4}.$$

So, our visual computation of staircase series, and its production of equation (1), coupled with the Carlitz identity, allows us to see that

$$\begin{aligned}\frac{r \cdot p_n(r)}{(1-r)^{n+1}} = \sum_{k=0}^{\infty} k^n r^k &= \frac{r}{1-r} \cdot \sum_{i=0}^{n-1} \left(\binom{n}{i} \sum_{k=0}^{\infty} k^i r^k \right) \\ &= \frac{r}{1-r} \cdot \sum_{i=0}^{n-1} \left(\binom{n}{i} \frac{r \cdot p_i(r)}{(1-r)^{i+1}} \right),\end{aligned}$$

which leads to the following recursive construction of the Eulerian polynomials:

$$p_n(r) = \sum_{i=0}^{n-1} \left(\binom{n}{i} r(1-r)^{n-i-1} p_i(r) \right), \tag{2}$$

where here we must set $p_0(r) = \frac{1}{r}$ since the variation on the Carlitz identity does not hold for $n = 0$.

Exercise 2. Use Figure 3 and equation (2) to find $p_7(r)$, then find the sum of the series $\sum_{k=0}^{\infty} k^7 (\frac{1}{2})^k$.

While "proofs without words" are not formal proofs, the point of such diagrams is to provide a visual that helps one recreate the theorem and/or proof. The reduction formula given in equation (1) can quickly be reconstructed by remembering the diagram in Figure 2. This reduction formula provides a method of computing the sum of a variety of series without having to apply calculus techniques making it suitable for introducing students to series. Surprisingly at first, the same reduction formula also has many interesting connections to combinatorics, as can be seen in [**2**] and [**3**] as well as our included recursion for Eulerian polynomials.

REFERENCES

[1] Carlitz, L. (1975). A combinatorial property of q-Eulerian numbers. *Amer. Math. Monthly* 82:51–54.

[2] Dasef, M. E., Kautz, S. M. (1997). Some sums of some significance. *College Math. J.* 28(1):52–55.

[3] Gorfin, A. (1989). Evaluating the sum of the series $\sum_{k=1}^{\infty} \frac{k^j}{M^k}$. *College Math. J.* 20(4):329–331.

[4] Petersen, T. K. (2015). *Eulerian Numbers*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. New York: Birkhäuser/Springer. With a foreword by Richard Stanley.

[5] Swain, S. G. (1994). Proof without words: Gabriel's staircase. *Math. Mag.* 67(3):209.

Summary. We provide a wordless calculation of the series $\sum_{i=0}^{\infty} k^2 r^k$ for any $0 < r < 1$. We extend this visual proof to give an algebraic reduction formula for $\sum_{i=0}^{\infty} k^n r^k$ where $n \geq 1$ is an integer, and we discuss a combinatorial consequence of this formula.

TOM EDGAR (MR Author ID: 821633) received his Ph.D. from the University of Notre Dame and is an associate professor at Pacific Lutheran University in Tacoma, WA. He enjoys helping his 3/2-year old son learn to explore the world and hopes that one day his son will learn how to climb staircase series.

Darboux Sums and the Sum of the Alternating Harmonic Series

FRANCISCO SÁNCHEZ
Departamento de Matemática Aplicada
Universidad Politécnica de Valencia, Spain
fsanchez@mat.upv.es

JOSÉ MARÍA SANCHIS
Departamento de Matemática Aplicada & IUMPA
Universidad Politécnica de Valencia, Spain
jmsanchis@mat.upv.es

We express a partial sum as a Darboux sum and compute a limit as an integral to prove that the alternating harmonic series sums to ln(2).

Theorem. $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + - \cdots = \ln(2)$

Proof. It suffices to prove that the sequence $\{S_{2n}\}$ of partial sums converges to $\ln(2)$.

$$\begin{aligned}
S_{2n} &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + - \cdots - \frac{1}{2n}\\
&= \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots + \frac{1}{2n}\right) - 2\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \cdots + \frac{1}{2n}\right)\\
&= \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots + \frac{1}{2n}\right) - \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n}\right)\\
&= \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{n+n} = \frac{1}{n}\left(\frac{n}{n+1} + \frac{n}{n+2} + \cdots + \frac{n}{n+n}\right)\\
&= \frac{1}{n}\left(\frac{1}{1+\frac{1}{n}} + \frac{1}{1+\frac{2}{n}} + \cdots + \frac{1}{1+\frac{n}{n}}\right) = \frac{1}{n}\left(f\left(\frac{1}{n}\right) + f\left(\frac{2}{n}\right) + \cdots + f\left(\frac{n}{n}\right)\right),
\end{aligned}$$

where $f(x) = \frac{1}{1+x}$. The last expression is a Darboux sum of $f(x)$ on the interval $[0, 1]$ for the partition $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\}$. Since $f(x)$ is continuous, and therefore Riemann integrable, on $[0, 1]$, we obtain

$$\lim_{n\to\infty} S_{2n} = \lim_{n\to\infty}\frac{1}{n}\left(f\left(\frac{1}{n}\right) + f\left(\frac{2}{n}\right) + \cdots + f\left(\frac{n}{n}\right)\right) = \int_0^1 \frac{1}{1+x}\,dx = \ln 2. \qquad \blacksquare$$

Summary. We provide a simple proof that the alternating harmonic series sums to the natural logarithm of 2 by expressing partial sums as a Darboux sum of a function on [0, 1] and on the computation of a limit as an integral.

FRANCISCO SÁNCHEZ GUIRADO (MR Author ID: 1233331) graduated in mathematics in 1980 at the University of Valencia. He has taught mathematics at the ETSII of the Polytechnic University of Valencia since 1981. He is an International Chess Master (IM) and has been one since 1987.

JOSÉ MARÍA SANCHIS LLOPIS (MR Author ID: 671500, ORCID 0000-0002-0039-8122) graduated in mathematics in 1985 and obtained his doctoral thesis in 1990, both at the University of Valencia. Since 1986 he has taught mathematics at the ETSII of the Polytechnic University of Valencia. He is an active researcher in the area of operations research, mainly in vehicle routing problems.

Math. Mag. **91** (2018) 96. doi:10.1080/0025570X.2017.1408380 © Mathematical Association of America
MSC: Primary 40A05, Secondary 97I30

A Combinatorial Approach to Wilson's Theorem for Finite Abelian Groups

CHASE SAUCIER
University of Florida
Gainesville, FL 32611
fugue@ufl.edu

Here, we will obtain a few results in group theory without using the standard theorems of Lagrange and Cauchy, or even the basic algebraic concepts of subgroups, quotient groups, or homomorphisms. Instead, we use simple counting and involutive arguments. Recall that Wilson's theorem says that $(p-1)! \equiv -1 \pmod p$ if p is prime (see [1]). Note that $(p-1)!$ is the product of all the units modulo p. We study this problem in a more general setting, and ask: Given an arbitrary finite Abelian group G, what can be said of the product over all the elements of G? The answer is well-known, but difficult to find in a standard textbook. We present our own approaches and give references to others' approaches below.

Preliminaries

In the following, G will denote a finite Abelian group unless otherwise stated. The identity of G will be denoted by 1, and $p(G)$ will denote the product over all the elements of G, that is,

$$p(G) = \prod_{g \in G} g.$$

We make a preliminary observation about $p(G)$: Since every element of G appears in $p(G)$, we may cancel each element g with its inverse g^{-1}, so long as $g \neq g^{-1}$. If an element is its own inverse, then it cannot be cancelled with itself in $p(G)$, since each element appears only once in the product. Now, an element g is its own inverse if and only if the order of g is less than or equal to 2. Thus, the product $p(G)$ can be immediately reduced to the product over all the elements of order 2. We summarize this as follows.

Reduction Lemma. *For an Abelian group G, then* $p(G) = \prod_{|g|=2} g$.

The existence of elements of order 2 is thus the only obstruction to achieving full cancellation.

Corollary. *If G has no elements of order 2, then* $p(G) = 1$.

We now note that Euler's theorem on congruences (see [1]) can be extended immediately to arbitrary finite Abelian groups.

Euler's Theorem. *Let G be a finite Abelian group. Then for any* $g \in G$*, we have* $g^{|G|} = 1$.

Math. Mag. **91** (2018) 97–102. doi:10.1080/0025570X.2017.1419748
MSC: Primary 05E15, Secondary 20K01

Proof. Recall that for any $g \in G$, the map $f(x) = gx$ is a permutation of G. Thus we have

$$p(G) = \prod_{x \in G} x = \prod_{x \in G} gx = g^{|G|} p(G)$$

and canceling $p(G)$ from both sides gives the result. ■

When the order of G is odd

Suppose the order of G is odd, so that $|G| = 2n + 1$. We can now show that $p(G)$ is trivial.

Theorem 1. *If $|G|$ is odd, then $p(G) = 1$.*

Proof. If $g \in G$ satisfies $g^2 = 1$, then by Euler's theorem

$$1 = g^{|G|} = g^{2n+1} = g^{2n} g = g.$$

By the reduction lemma, G has no elements of order 2 and the result follows. ■

We now give an alternate proof of Theorem 1 via the generalized pigeonhole principle, which does not require the reduction lemma or any consideration of elements of order 2. We also derive Euler's theorem (for groups of odd order) in a new way. Before we do so, we will need the following basic counting principle (see the first chapter of [**2**]).

Theorem (Generalized Pigeonhole Principle). *If n objects are chosen (with replacement) from a set of m objects, and $\frac{n}{m} > r$, then some object was chosen at least $r + 1$ times. In other words, if n pigeons fly into m holes, and $\frac{n}{m} > r$, then some hole contains at least $r + 1$ pigeons.*

Alternate Proof of Theorem 1. Let $z \in G$ be arbitrary. Let

$$G_z = \{\{x, y\} : x, y \in G \text{ and } x, y, z \text{ all distinct}\},$$

and let $\phi : G_z \to G$ be defined by $\phi(\{x, y\}) = xy$.

Note that $|G_z| = \binom{2n}{2} = 2n^2 - n$. It follows that ϕ makes $2n^2 - n$ choices from a set with $2n + 1$ elements as it maps G_z into G. Notice that $2n^2 - n$ divided by $2n + 1$ is $n - 1$ with remainder 1. By the generalized pigeonhole principle, there are at least n elements in G_z that get mapped to the same element $g \in G$. That is, we have

$$g_1 g_2 = g_3 g_4 = \cdots = g_{2n-1} g_{2n} = g, \tag{1}$$

where $g_i \neq z$ for $1 \leq i \leq 2n$ and the two elements appearing in each product are distinct from each other.

We claim that $\{g_1, \ldots, g_{2n}\} = G \setminus \{z\}$. Notice that all the g_i are distinct. Otherwise, if ab and ac were two of the products in equation (1), then $ab = ac$ implies $b = c$, contrary to the hypothesis that $\{a, b\}$ and $\{a, c\}$ were distinct elements of G_z. In other words, the preimage $\phi^{-1}(g)$ is a partition of $G \setminus \{z\}$. We could also say that equation (1) gives us all the ways to factor g into two elements not equal to z.

Further, it follows that $g = z^2$. To see this, let $x = z^{-1} g$. Then we must have $x = z$; otherwise from above, x appears in one of the above products, say $x g_i$, and we would have $x g_i = g = xz$, which implies $g_i = z$ contrary to our choice of g_i.

By combining equation (1) with the above two paragraphs and taking $z = 1$, then we can always pair each nonidentity element of G with a distinct inverse; thus, by our preliminary observation, G can have no elements of order 2 (and consequently no elements of even order).

It also follows from equation (1) and from the preceding paragraphs that

$$p(G) = z \prod_{i=1}^{n} g_{2i-1} g_{2i} = z \prod_{i=1}^{n} z^2 = z^{2n+1} = z^{|G|}. \tag{2}$$

Since z was arbitrary, the theorem follows by taking $z = 1$. ■

The next corollary follows from equation (2) and Theorem 1.

Corollary (Euler's Theorem for a Group of Odd Order). *Let G be a group of odd order. For $z \in G$, $z^{|G|} = 1$.*

When the order of G is even

Now suppose $|G| = 2n$. Unfortunately, we cannot apply quite the same argument as above. For an arbitrary $z \in G$, we have an odd number of elements left in $G \setminus \{z\}$, so that there is some $x \in G \setminus \{z\}$ that does not get paired. If we try to circumvent this by isolating two elements x and y and then choosing pairs from $G \setminus \{x, y\}$, the generalized pigeonhole principle now only gives the existence of $n - 2$ equal products! In either case, it is hard to say anything definite since we cannot control which elements get left out of the pairing. We therefore appeal to the reduction lemma, and reduce $p(G)$ to the product of all the elements of order 2 in G.

Let us first show that G has an element of order 2. Since G has even order, by removing the identity we are left with an odd number of elements. Then elements not of order 2 may be removed in pair with their inverses, after which we are still left with an odd number of elements. Thus G has at least one element of order 2. Note that if G has a unique element z of order 2, then (by the reduction lemma) we must have $p(G) = z$.

In the following, let $A = \{g \in G : |g| = 2\}$ and let $|A| = 2m + 1$. Note that A is not a subgroup of G because $1 \notin A$. We apply the generalized pigeonhole principle as in the odd-order case above. In particular, choose an arbitrary element $z \in A$ and let

$$A_z = \{\{x, y\} : x, y \in A \text{ and } x, y, z \text{ all distinct}\}.$$

Note that because the product of two distinct elements of order 2 is again an element of order 2, we have $\phi(A_z) \subseteq A$, and we may apply the generalized pigeonhole principle to obtain m equal products:

$$a_1 a_2 = a_3 a_4 = \cdots = a_{2m-1} a_{2m} = a \in A, \tag{3}$$

where the a_i are distinct and not equal to z. As in the odd-order case, equation (3) gives all the factorizations of a as the product of two elements in A distinct from z.

It follows that $a = z$. If $a \neq z$, then a must equal some other element x of order 2. As in the odd-order case, x must appear in one of our m equal products, say xa_i, so that we would have $xa_i = a = x$, implying $a_i = 1$, contrary to our choosing a_i of order 2.

Using the relations in equation (3) and that $a = z$ we may write

$$p(G) = z^{m+1}.$$

If $m + 1$ is odd, then, because z is of order 2, we find $p(G) = z$ for arbitrary $z \in G$ of order 2. This is only possible if z is the unique element of order 2 in G. If $m + 1$ is even, we have $p(G) = 1$. Putting these results together with Theorem 1, we have proved the following theorem.

Theorem (Wilson's Theorem for Finite Abelian Groups). *For any finite Abelian group G, if G has a unique element z of order 2, then $p(G) = z$; otherwise, $p(G) = 1$.*

By adjoining the identity 1 to the set A above, we obtain the elementary Abelian group $H = \{g \in G : |g| \leq 2\}$ which has (by Cauchy's theorem; see [5]) order $|H| = 2^r$ for some integer $r \geq 1$. Since $|A| = 2m + 1$, we have

$$2(m+1) = 2m + 2 = |\{1\} \cup A| = |H| = 2^r.$$

Thus, $m + 1$ is odd if and only if $r = 1$, as confirmed above.

Finally, note that prime numbers are not mentioned in the statement of Wilson's theorem for finite Abelian groups: we have lost the ring-theoretic setting of the classical Wilson theorem. As a consequence, Wilson's theorem for finite Abelian groups does not allow one to immediately deduce the classical Wilson theorem. We must first check that the set of positive integers less than the prime p forms a multiplicative group modulo p. Then we must show that -1 is indeed the unique element of order 2. A proof of this is outlined below.

Proof. Let p be prime and note that an element x of order 2 in the group U_p of units modulo p satisfies the equation $x^2 \equiv 1 \pmod{p}$. This may be rewritten as $x^2 - 1 \equiv (x+1)(x-1) \equiv 0 \pmod{p}$. The last equality together with the hypothesis that p is prime implies that p must divide either $x + 1$ or $x - 1$. The only nontrivial solution is $x \equiv -1 \pmod{p}$. This shows -1 is the unique element of order 2 in U_p; together with the reduction lemma this gives the classical Wilson theorem. ■

An involutive approach

Let G be an arbitrary (perhaps non-Abelian) finite group, and suppose $h \in G$ has order 2. Note that for arbitrary $g_1, g_2 \in G$, we have that $hg_1 = g_2$ if and only if $g_1 = hg_2$. This shows that the map $f_h(g) = hg$ defines an *involution* of G, that is, $f_h^2(g) = g$ for all $g \in G$. Furthermore, f_h has no fixed points; $f_h(g) \neq g$ for all $g \in G$. Such an involution without fixed points defines a pairing of G (every element is paired with its image under f_h) so that G must have even order. Thus, an arbitrary group of odd order cannot have elements of even order. (Note that the above can be thought of as a very special case of Lagrange's theorem: the "pairing" established by the map f_h is a partition of G into the cosets of the subgroup $\langle h \rangle = \{1, h\}$.)

Now let G be Abelian. If $|G|$ is odd, we can construct the relations in equation (1) as follows: Choose arbitrary $z \in G$ and note that the map $\varphi_z(g) = g^{-1}z^2$ is an involution. Furthermore, we have $\phi_z(g) = z$ if and only if $g = z$. To show that φ_z establishes a pairing of the elements of $G \setminus \{z\}$, suppose that $g \neq z$. Then $\varphi_z(g) \neq g$, or else $1 = (g^{-1})^2 z^2 = (g^{-1}z)^2$, which implies that $g^{-1}z$ is an element of order 2. The pairs are then those appearing in equation (1).

If $|G|$ is even, we can construct the relations in equation (3). Again let $A = \{g \in G : |g| = 2\}$, choose $z \in A$, and recall that the map $f_z(a) = za$ is an involution without fixed points such that $f_z(A \setminus \{z\}) \subseteq A - \{z\}$. Thus f_z establishes a pairing of $A \setminus \{z\}$, whose pairs are those appearing in equation (3). To the author's knowledge, this is the simplest route to Wilson's theorem for finite Abelian groups.

A standard proof

Recalling the reduction lemma, we need only evaluate the product over the elementary Abelian group $H = \{g \in G : |g| \leq 2\}$. Clark [3] suggests that "... to the *cognoscenti*, it is irresistible to begin . . .” by appealing to the fundamental theorem of finite Abelian groups to write H as a direct product of cyclic groups of order 2. In fact, elementary Abelian groups may be written additively as a vector space over the field F_p, where p is the prime such that $h^p = 1$ for all $h \in H$. We note that this decomposition does not require the full power of the fundamental theorem of finite Abelian groups; see [5, chapter 1.8, exercise 6].

If $|H| = 2^n$, then

$$H \cong F_2 \times \cdots \times F_2 \cong F_2^n.$$

We define an involution $\sigma : F_2^n \to F_2^n$ without fixed points by $\sigma(h) = (1, 1, \ldots, 1) - h$, where the first term of the difference is understood to be the vector with all coordinates equal to 1. Now we sum over the pairs established by σ, of which there will be 2^{n-1}:

$$p(G) = \sum_{h\in H} h = \sum_{i=1}^{2^{n-1}} [h_i + \sigma(h_i)] = \sum_{i=1}^{2^{n-1}} (1, 1, \ldots, 1) = (2^{n-1}, 2^{n-1}, \ldots, 2^{n-1}).$$

If $n > 1$ the final term is the zero vector (mod 2); if $n = 1$ the sum is equal to 1 (mod 2), the unique element of order 2.

Other generalizations of Wilson's theorem

Gauss (see [9]) extended Wilson's theorem to the product over the group of units U_m for any integer $m > 2$. In particular, $p(U_m) \equiv -1 \pmod{m}$ if $m = 4$, p^b, or $2p^b$, where p is an odd prime and b is a positive integer; otherwise, $p(U_m) \equiv 1 \pmod{m}$.

Other generalizations were obtained in [7], such as the following theorem.

Theorem. *For any prime p and any natural number $n \geq p$, we have*

$$\sum_{k=1}^{t} \frac{n!}{p^k (n - kp)! k!} \equiv -1 \pmod{p}$$

where t is the largest natural number such that $tp \leq n$.

If we let $n = p$ in the above formula, we obtain the classical Wilson theorem. Also obtained in [7] is the following theorem, which reduces to Wilson's theorem when used for one prime $n = p$.

Theorem. *If $\frac{n}{2} < p_1 < p_2 < \cdots < p_t \leq n$, where each p_i is prime, then*

$$\sum_{k=1}^{t} \frac{n!}{p_k (n - p_k)!} \equiv -1 \pmod{p_1 p_2 \cdots p_t}.$$

Clark's paper [3] contains a trove of Wilson-type theorems, and there he states: “So far as I know, [Wilson's theorem for finite Abelian groups] was first proved by the early American group theorist G.A. Miller [8] . . . But I have not been able to find this result in any standard textbook.”

When G is non-Abelian

In the case when G is non-Abelian, $p(G)$ will in general depend on the ordering of the elements in the product; see [4] for details. However, we might consider trying to show (without Lagrange's theorem or involutions) that non-Abelian groups of odd order can have no elements of even order, with the idea mentioned in our alternate proof of Theorem 1. That is, if we can pair each nonidentity element with a distinct inverse, we will have shown there are no elements of order 2. We must now take into account that an element may appear more than once (but at most twice) in the pairs we obtain. We will have more pairs to work with since we count ordered pairs. Unfortunately, the generalized pigeonhole principle leaves us a pair short!

To be precise, let $|G| = 2n + 1$, and let $G_1 = \{(a, b) : a, b \in G, a \neq b$ and $a, b \neq 1\}$. Then $|G_1| = 2n(2n - 1)$ and the generalized pigeonhole principle gives us only $2n - 1$ equal products of distinct nonidentity elements, which is not enough to make up for the redundancy of elements appearing twice among these products. That is, some nontrivial $g \in G$ may not appear among our products, so we are unable to argue that they all must be equal to 1 as in the Abelian case. If we had $2n$ equal products, we should be able to argue that, because an element appears at most twice among them, every nonidentity element appears among them (by the standard pigeonhole principle) and (as before) this would force all the products to be equal to 1.

We ask: Is there a way to sharpen this argument (without appealing to Lagrange's theorem)?

Acknowledgment The author would like to thank his advisor Dr. Paul Robinson and the referees for their many helpful comments regarding the presentation of the content of this paper.

REFERENCES

[1] Andrews, G. E. (1994). *Number Theory*. New York: Dover Publications.
[2] Bóna, M. (2006). *A Walk Through Combinatorics: An Introduction to Enumeration and Graph Theory*. 2nd ed. Singapore: World Scientific.
[3] Clark, P. L. (n.d.) Wilson's theorem: An algebraic approach. `http://alpha.math.uga.edu/~pete/wilson_easy.pdf`
[4] Dénes, J., Hermann, P. (1982). On the product of all elements in a finite group. *North-Holland Math. Studies* 65: 105–109.
[5] Hungerford, T. (1974). *Algebra*. New York: Springer Science+Business Media, Inc.
[6] Isaacs, M. (1994). *Algebra: A Graduate Course*. Pacific Grove, CA: Brooks/Cole Publishing Company.
[7] Khurana, A., Khurana, D. (2005). A theorem of Frobenius and its applications. *Math. Mag.* 78: 220–225.
[8] Miller, G. A. (1903). A new proof of the generalized Wilson's theorem. *Ann. Math.* 4(2): 188–190.
[9] Ore, O. (1948). *Number Theory and Its History*. New York: McGraw-Hill Book Company, Inc.

Summary. The generalized pigeonhole principle is used to deduce an analog of Wilson's theorem for finite Abelian groups. An involutive approach and an approach using Euler's congruence theorem are also discussed. Other generalizations of Wilson's theorem are presented as well. We conclude with a discussion of how one might apply the generalized pigeonhole principle in the non-Abelian case.

CHASE SAUCIER (MR Author ID: 1229575) is a Ph.D. candidate and teaching assistant in the math department at the University of Florida, after having received a Bachelor's and Master's degrees in mathematics there. He enjoys long walks through nature and playing classical piano.

Cayley's Theorem Revisited: Embeddings of Small Finite Groups

ROBERT HEFFERNAN
Cork Institute of Technology
Bishopstown, Cork, Ireland
Robert.Heffernan@cit.ie

DES MACHALE
University College, Cork
Cork, Ireland
d.machale@ucc.ie

BRENDAN MCCANN
Waterford Institute of Technology
Waterford, Ireland
bmccann@wit.ie

One of the earliest and most significant theorems of group theory is Cayley's theorem which states:

> If a group G has order n, then G is isomorphic to a subgroup of S_n, the symmetric group on n symbols.

Cayley's theorem derives from the fact that left multiplication by a fixed element of a group permutes the elements of that group. This was noted for groups of permutations (then also known as substitutions) by Cauchy [**3**] in 1845, and for more abstract groups by Cayley [**4**] in 1854. Thus, in principle, any question regarding finite groups can be reduced to a question about permutation groups. For example, the problem of determining the possible groups of order n can be reduced to the apparently simpler problem of determining the set of nonisomorphic subgroups of order n in S_n. However, as Cayley himself noted when considering the same problem, this "does not in any wise show that the best or easiest mode of treating the general problem is thus to regard it as a problem of substitutions." Indeed it seemed clear to Cayley that the better course was "to consider the general problem in itself, and to deduce from it the theory of groups of substitutions" (Cayley [**5**], Kleiner [**9**]).

Cayley's theorem tells us that the symmetric groups are rich in subgroups, with S_n containing an isomorphic copy of each group of order n. In other words, every group of order n can be (isomorphically) embedded in S_n. Moreover, since S_n obviously contains a copy of S_{n-i} for each $i < n$, we see that S_n contains a copy of each finite group of order less than or equal to n. We pose five elementary questions concerning embeddings of groups and answer them for groups of order 15 or less. Questions 1, 2, and 4 concern minimal embeddings in symmetric groups, whereas questions 3 and 5, in keeping with Cayley's more abstract approach to group theory, relate to minimal embeddings in general. By placing an emphasis on the arrangement of subgroups in a group, our questions could form the basis for a novel (re)acquaintance with the theory of finite groups.

For groups of order n, we ask the following five questions.

Math. Mag. **91** (2018) 103–111. doi:10.1080/0025570X.2017.1419770
MSC: Primary: 20B35, Secondary: 20D99
Supplemental data for this article can be accessed on the publisher's website.

TABLE 1: Groups of order 15 or less

n	Groups of order n	n	Groups of order n
1	C_1	9	$C_9, C_3 \times C_3$
2	C_2	10	C_{10}, D_5
3	C_3	11	C_{11}
4	$C_4, C_2 \times C_2$	12	$C_{12}, C_2 \times C_2 \times C_3, D_6, Q_3, A_4$
5	C_5	13	C_{13}
6	C_6, D_3	14	C_{14}, D_7
7	C_7	15	C_{15}
8	$C_8, C_2 \times C_4, C_2 \times C_2 \times C_2, D_4, Q_2$		

1. For each group G, what is the least m such that G can be embedded in S_m?
2. For each n, what is the least m such that all groups of order n can be embedded in S_m?
3. For each n, what is the size of a smallest group K such that all groups of order n can be embedded in K? When is K unique?
4. For each n, what is the least m such that all groups of order less than or equal to n can be embedded in S_m?
5. For each n, what is the order of a smallest group K that contains an isomorphic copy of each group of order less than or equal to n? When is $|K| = |S_n|$?

To answer our questions, we first need to list the groups of order 15 or less. Denote the cyclic group of order n by C_n and let A_n denote the alternating group of degree n and order $\frac{n!}{2}$ for $n \geq 2$. Let S_n denote the symmetric group of degree n and order $n!$ for $n \geq 1$. For the subset X of a group G, we use $\langle X \rangle$ to denote the subgroup of G generated by X. The dihedral group of order $2n$ ($n \geq 3$), given in terms of generators and relations by $\langle a, b \mid a^n = 1 = b^2; b^{-1}ab = a^{-1} \rangle$, is denoted by D_n; while Q_n ($n \geq 2$) denotes the dicyclic group of order $4n$, given by $\langle a, b \mid a^{2n} = 1; b^2 = a^n; b^{-1}ab = a^{-1} \rangle$. In particular, Q_2 is the quaternion group of order 8 and Q_3 is the non-trivial semi-direct product of C_3 by C_4. Table 1 gives (up to isomorphism) the groups of order n for $n \leq 15$. We note the following isomorphisms: $C_6 \cong C_2 \times C_3$, $D_3 \cong S_3$, $C_{10} \cong C_2 \times C_5$, $C_{12} \cong C_3 \times C_4$, $D_6 \cong C_2 \times D_3$, $C_{14} \cong C_2 \times C_7$, and $C_{15} \cong C_3 \times C_5$.

We note that it was Cayley [6] who, in 1889, produced the first complete listing of the groups of order 12 or less, an accomplished effort that can be viewed as a forerunner of modern repositories such as the GAP Small Groups Library [2]. Cayley [5] had previously proposed in 1878 that the "general problem" of group theory was "to find all groups of a given order n." However, despite his early successes and the present-day availability of computer algebra software, the "general problem" has proved difficult to solve for values of n with many noncoprime factors, even when n is relatively small. This points to why we have not ventured beyond $n = 15$. Perhaps surprisingly, it turns out that questions 3 and 5 have yet to be answered for $n = 16$. This is due to the fact that there are 14 nonisomorphic groups of order 16. As 16 is a power of 2, a group of minimal order in which all these groups can be embedded will be a finite 2-group, that is, a group with 2^k elements for some k. But the number of nonisomorphic finite 2-groups increases rapidly with increasing k, as shown in Table 2 [1].

Although the *number* of groups of order 1,024 has been calculated, methods have yet to be developed that will allow us to search through them for the presence of particular subgroups. Thus, our seemingly innocuous questions have led us to an open problem of considerable complexity.

TABLE 2: Numbers of 2-groups for orders between 16 and 1024

n	Number of groups	n	Number of groups
16	14	256	56,092
32	51	512	10,494,213
64	267	1,024	49,487,365,422
128	2,328		

Answers to question 1

For small groups question 1 is comparatively easy to answer, with the possible exception of the quaternion group Q_2 which we will treat later. For each group of order 15 or less, Table 3 lists the smallest symmetric group into which that group can be embedded, and gives a minimal permutation representation of that group.

We remark that the permutation representations for C_n and D_n have instructive alternative versions in terms of subgroups of symmetry groups of regular n-sided polygons. We note further that $\langle(12), (34)\rangle$ and $\langle(12)(34), (14)(32)\rangle$ are nonconjugate embeddings of $C_2 \times C_2$ in S_4. This tells us that minimal embeddings of a given group

TABLE 3: Minimal permutation representations for groups of order 15 or less

Group	Smallest symmetric group	Permutation representation
C_1	S_1	$\langle e\rangle$
C_2	S_2	$\langle(12)\rangle$
C_3	S_3	$\langle(123)\rangle$
C_4	S_4	$\langle(1234)\rangle$
$C_2 \times C_2$	S_4	$\langle(12), (34)\rangle$
C_5	S_5	$\langle(12345)\rangle$
D_3	S_3	$\langle(123), (12)\rangle$
C_6	S_5	$\langle(123)(45)\rangle$
C_7	S_7	$\langle(1234567)\rangle$
C_8	S_8	$\langle(12345678)\rangle$
$C_2 \times C_4$	S_6	$\langle(12), (3456)\rangle$
$C_2 \times C_2 \times C_2$	S_6	$\langle(12), (34), (56)\rangle$
D_4	S_4	$\langle(1234), (12)(34)\rangle$
Q_2	S_8	$\langle(1234)(5678), (1638)(2547)\rangle$
C_9	S_9	$\langle(123456789)\rangle$
$C_3 \times C_3$	S_6	$\langle(123), (456)\rangle$
C_{10}	S_7	$\langle(12345)(67)\rangle$
D_5	S_5	$\langle(12345), (14)(32)\rangle$
C_{11}	S_{11}	$\langle(12345\ldots11)\rangle$
C_{12}	S_7	$\langle(1234)(567)\rangle$
$C_2 \times C_2 \times C_3$	S_7	$\langle(12), (34), (567)\rangle$
D_6	S_5	$\langle(123), (12), (45)\rangle$
Q_3	S_7	$\langle(123), (12)(4567)\rangle$
A_4	S_4	$\langle(123), (12)(34)\rangle$
C_{13}	S_{13}	$\langle(12345\ldots13)\rangle$
C_{14}	S_9	$\langle(1234567)(89)\rangle$
D_7	S_7	$\langle(1234567), (27)(36)(45)\rangle$
C_{15}	S_8	$\langle(12345)(678)\rangle$

TABLE 4: Symmetric group of least order containing all groups of order n, for $1 \leq n \leq 15$

n	Smallest symmetric group	n	Smallest symmetric group	n	Smallest symmetric group
1	S_1	6	S_5	11	S_{11}
2	S_2	7	S_7	12	S_7
3	S_3	8	S_8	13	S_{13}
4	S_4	9	S_9	14	S_9
5	S_5	10	S_7	15	S_8

in S_m need not necessarily be conjugate in S_m. As we have mentioned, only one of the results in Table 3 presents any serious challenge.

Theorem 1. *The smallest symmetric group that contains Q_2, the quaternion group of order 8, is S_8.*

Proof. By Cayley's theorem Q_2 can be embedded in S_8. We proceed to show that it cannot be embedded in S_7 and hence not in any smaller symmetric group. A Sylow 2-subgroup of S_7 has order 16 and is isomorphic to $D_4 \times C_2 \cong \langle(1234), (12)(34), (56)\rangle$. Now every 2-subgroup of a group is contained in a Sylow 2-subgroup, so the question is can $D_4 \times C_2$ have a subgroup isomorphic to Q_2? But $D_4 \times C_2$ has only four elements of order 4, whereas Q_2 has six such elements, so inclusion is impossible. Thus our result is established. ∎

Answers to question 2

From the results answering question 1, we develop Table 4, which presents the answers to question 2 for groups of order at most 15. Consideration of the embeddings of cyclic groups in Table 4 leads us to the following general results.

Theorem 2. *If n is a prime or prime power, then the smallest symmetric group in which all of the groups of order n can be embedded is S_n.*

Theorem 3. *If $n = 2p$, where p is a prime, then the smallest symmetric group in which all of the groups of order n can be embedded is S_{p+2}.*

(In fact there are just two nonisomorphic groups of order $2p$, namely, C_4 and $C_2 \times C_2$ for $p = 2$ and C_{2p} and D_p for p odd).

We note that the question of determining the smallest symmetric group in which certain classes of groups and their direct products can be embedded has been studied by Johnson [8], Wright [13], and Saunders [10, 11].

Answers to question 3

How efficient are the symmetric groups for embedding purposes? Could we perhaps find other, *smaller* groups that perform the same functions? This is the motivation for questions 3 and 5. To provide answers to question 3, we need to define three additional groups, which we denote by H_1, H_2, and H_3.

$H_1 = \langle x, y \mid x^8 = y^2 = 1, yxy = x^3\rangle$ is the quasidihedral (or semidihedral) group of order 16.

TABLE 5: Group(s) of least order containing all groups of order n, for $1 \leq n \leq 15$

n	Smallest group(s)	n	Smallest group(s)	n	Smallest group(s)
1	C_1	6	D_6	11	C_{11}
2	C_2	7	C_7	12	$S_3 \times S_4$
3	C_3	8	$C_2 \times H_1,\ H_2$	13	C_{13}
4	$C_2 \times C_4,\ D_4$	9	$C_3 \times C_9,\ H_3$	14	D_{14}
5	C_5	10	D_{10}	15	C_{15}

$H_2 = \langle x, y, z \mid x^8 = y^2 = z^2 = 1; yz = zy; yxy = x^{-1}; zxz = x^3 \rangle$ is the semidirect product of C_8 by its automorphism group. Thus H_2 is isomorphic to the holomorph of C_8 and $|H_2| = 32$.

$H_3 = \langle a, b \mid a^9 = 1 = b^3; b^{-1}ab = a^4 \rangle$ is the non-Abelian group of order 27 and exponent 9 (the exponent of a group G is the smallest k such that g^k is the identity for all $g \in G$).

For each value of n, with $1 \leq n \leq 15$, Table 5 lists the group(s) of smallest order into which all the groups of order n can be embedded.

Apart from the cases $n = 8$ and $n = 12$, the results in Table 5 are quite straightforward and indeed could be used as introductory exercises in group theory. It might also be instructive to consider how all groups of order 8 can be embedded in both $C_2 \times H_1$ and H_2, and how all groups of order 12 can be embedded in $S_3 \times S_4$. The results in Table 5 tell us that, for small values of n at least, the order of S_n is generally much greater than the order of a minimal group in which all groups of order n can be embedded.

However, we would note that the proof that our list is complete for $n = 8$ and $n = 12$ poses more of a problem. The full details can be found in [7]. Here, we content ourselves with an outline of the proof for the more elementary case $n = 8$. By Sylow's theorems, a group G of minimal order in which all groups of order 8 can be embedded must be a finite 2-group. It is straightforward to show that if $|G| = 16$ and G has a subgroup isomorphic to $C_2 \times C_2 \times C_2$, then G cannot have a cyclic subgroup of order 8. Since, by inspection, all groups of order 8 can be embedded in $C_2 \times H_1$ and H_2, we see that $|G| = 32$. To show that $C_2 \times H_1$ and H_2 are the only such groups, we first show that if our group G (now of order 32) has an Abelian subgroup K of order 16 such that $k^4 = 1$ for all $k \in K$, then not all groups of order 8 can be embedded in G. We combine this with the fact that every proper subgroup of a finite p-group is properly contained in its normalizer to show that G can be expressed in the form $G = \langle x_1, x_2, y \rangle$, for elements x_1, x_2, and y, with $o(x_1) = o(x_2) = 2$ and $o(y) = 8$, and such that $\langle y \rangle \trianglelefteq \langle x_1, y \rangle = C_G(y^2) \trianglelefteq G$. We note that $C_G(x) = \{g \in G \mid gx = xg\}$ denotes the centralizer of the element x in a group G.

Four cases then arise. First we have $\langle y \rangle \trianglelefteq G$ and $y^{x_1} = y$. In this case, we can show that $G \cong C_2 \times H_1$. In the second case, we have $\langle y \rangle \trianglelefteq G$ and $y^{x_1} \neq y$. Here we can show that $G \cong H_2$. In the remaining two cases, we have, respectively, $\langle y \rangle \ntrianglelefteq G$ and $y^{x_1} = y$; and $\langle y \rangle \ntrianglelefteq G$ and $y^{x_1} \neq y$. In both these cases it turns out that not all groups of order 8 can be embedded in such a group G. Thus $C_2 \times H_1$ and H_2 are the only groups of order 32 in which all groups of order 8 can be embedded.

The proof for the case $n = 12$ is no longer elementary. It involves two deep results from finite group theory, namely, Burnside's theorem, which states that if p and q are primes, then any group of order $p^a q^b$ (for $a \neq 0$ and $b \neq 0$) has a nontrivial normal subgroup; and Maschke's theorem, which allows us to decompose Abelian subgroups in a group that are normalized by other subgroups of coprime order. Thus, our elementary questions act as signposts pointing to some of the more advanced topics in group theory.

TABLE 6: Symmetric group of least order containing all groups of order n or less, for $1 \le n \le 15$

n	Smallest symmetric group	n	Smallest symmetric group	n	Smallest symmetric group
1	S_1	6	S_5	11	S_{11}
2	S_2	7	S_7	12	S_{11}
3	S_3	8	S_8	13	S_{13}
4	S_4	9	S_9	14	S_{13}
5	S_5	10	S_9	15	S_{13}

An alternative approach for small values of n, up to and including $n = 15$, is to consider the use of computer algebra software, which was used in tandem with "pen and paper" methods in our explorations of groups of small order. Problems such as the above could provide a stimulating introduction to the GAP computer algebra system [**12**] and, in particular, to the Library of Small Groups [**2**]. A brief example of how we explored the questions we have posed can be found as an online supplement.

Answers to question 4

Here we return to embeddings in symmetric groups. We can again read off the answers by consulting Table 3. Recall that we are asking, for each n, what is the smallest symmetric group that contains a copy of each group of order less than or equal to n?

We note that if p is a prime and $k > 0$, then S_{p^k-1} cannot contain a cycle of length p^k. This explains the "jumps" in Table 6 when n is a prime power.

Answers to question 5: The functions n^* and π_n

We come finally to the most difficult, and perhaps the most interesting, of our questions: for each n, determine the smallest order of a group in which every group of order less than or equal to n can be embedded and discuss its uniqueness. The following functions will be useful in this regard. For each $n \in \mathbb{N}$, we define n^* to be the least common multiple of $\{1, 2, 3, \ldots, n\}$. For example, $1^* = 1$; $2^* = 2$; $3^* = 6$; $4^* = 12$; $5^* = 60$; $6^* = 60$; and $7^* = 420$. We define π_n for each $n \in \mathbb{N}$ as the product of all the prime divisors of $n!$. Thus $\pi_1 = 1$; $\pi_2 = 2$; $\pi_3 = 6$; $\pi_4 = 6$; $\pi_5 = 30$; $\pi_6 = 30$; and $\pi_7 = 210$.

Alternative definitions of n^* and π_n are as follows. Let $\{p_1, \ldots, p_m\}$ be the set of distinct primes that are less than or equal to n. For $i = 1, \ldots, m$ we let k_i be maximal such that $p_i^{k_i} \le n$. Then $\pi_n = p_1 \cdots p_m$ and $n^* = p_1^{k_1} \cdots p_m^{k_m}$. If we seek a smallest group, G, that contains all groups of orders less than or equal to n, then by Lagrange's theorem, $|G|$ must be a multiple of n^*. However, we shall see that this lower bound for $|G|$ is unattainable for $n \ge 4$.

In order to find an improved lower bound for our minimal groups, we note that if the p-group P has a cyclic subgroup, P_1, of order p^k and an elementary Abelian subgroup of rank k, P_2 (where an elementary Abelian p-group of rank k is one that is isomorphic to the direct product of k copies of C_p), then $|P_1 \cap P_2| \le p$ and

$$|P| \ge |P_1P_2| = \frac{|P_1||P_2|}{|P_1 \cap P_2|} \ge \frac{p^k p^k}{p} = p^{2k-1}.$$

TABLE 7: Examples of groups of least order in which all groups of order n or less can be embedded, for $1 \leq n \leq 11$.

n	Examples of minimal groups	Order
1	C_1	1
2	C_2	2
3	$C_6,\ D_3$	6
4	$C_4 \times D_3,\ S_4$	24
5	$C_4 \times C_5 \times D_3,\ C_5 \times S_4, S_5$	120
6	$C_4 \times C_5 \times D_3,\ C_5 \times S_4, S_5$	120
7	$C_4 \times C_5 \times C_7 \times D_3, C_5 \times C_7 \times S_4,\ C_7 \times S_5$	840
8	$C_5 \times C_7 \times D_3 \times H_1,\ C_7 \times D_{15} \times H_1$	3,360
9	$C_5 \times C_7 \times C_9 \times D_3 \times H_1,\ C_7 \times C_9 \times D_{15} \times H_1$	30,240
10	$C_7 \times C_9 \times D_{15} \times H_1,\ C_9 \times D_{105} \times H_1$	30,240
11	$C_7 \times C_9 \times C_{11} \times D_{15} \times H_1,\ C_9 \times C_{11} \times D_{105} \times H_1$	332,640

Here AB denotes the product of the subgroups A and B of the group G and is given by $AB = \{ab \mid a \in A,\ b \in B\}$. While AB is evidently a subset of G, we note that AB is a *subgroup* of G if and only if $AB = BA$. From the foregoing it follows, in particular, that if all groups of order p^k can be embedded in the finite group G, then the Sylow p-subgroups of G will have order at least p^{2k-1}. This allows us to improve on the lower bound of n^* given by Lagrange's theorem.

Theorem 4. *If G is a finite group in which all groups of order n or less can be embedded, then $|G|$ is a multiple of $\frac{(n^*)^2}{\pi_n}$.*

Proof. Letting $p_1, \ldots, p_m$ and $k_1, \ldots, k_m$ be as above, we see from the foregoing that $|G|$ is a multiple of $p_1^{2k_1-1} \cdots p_m^{2k_m-1} = \frac{p_1^{2k_1} \cdots p_m^{2k_m}}{p_1 \cdots p_m} = \frac{(n^*)^2}{\pi_n}$. ■

Theorem 4 tells us, for example, that if all groups of order 8 or less can be embedded in the finite group G, then $|G|$ is a multiple of $2^5 \cdot 3 \cdot 5 \cdot 7 = 3{,}360$, while if all groups of order 9 or less can be embedded in G, then $|G|$ is a multiple of $2^5 \cdot 3^3 \cdot 5 \cdot 7 = 30{,}240$. Table 7 gives examples of groups of minimal order in which all groups of order n or less can be embedded for values of n up to $n = 11$.

These easily verified examples show that the lower bound given by Theorem 4 can be attained for $n = 1, \ldots, 11$. While Table 7 is not exhaustive, it shows that the groups in question are not unique for $2 < n \leq 11$. Table 7 further shows that the order of S_n is minimal with respect to the embedding of all groups of order n or less for $n = 1, \ldots, 5$, whereas this is no longer the case for $n = 6, \ldots, 11$.

We note that for $n = 12$, 13, 14, and 15, the bound given by Theorem 4 must be increased by a factor of 2 [7]. To show this we require not only Sylow's theorems and Maschke's theorem, but also generalizations of Sylow subgroups called Hall subgroups, after their discoverer Philip Hall. We further require the determination of all finite simple groups whose orders are divisors of $2^5 \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 = 4{,}324{,}320$. Thus, commencing with Cayley's theorem, our questions have led us to the foothills of one of the great endeavors of modern algebra, namely, the classification of the finite simple groups.

While we cannot easily show that the bound given by Theorem 4 must be increased for $n = 12, \ldots, 15$, we can at least provide an example of a minimal group for each of these values of n. In fact, the cases $n = 13$, 14, and 15 are covered by the same example, so just two distinct groups are required. To construct our examples, we first

TABLE 8: Minimal order of a group in which all groups of order n or less can be embedded, for $1 \leq n \leq 15$

n	Minimal order	n	Minimal order
1	1	9	$30{,}240 = 2^5 \cdot 3^3 \cdot 5 \cdot 7$
2	2	10	30 240
3	$6 = 2 \cdot 3$	11	$332{,}640 = 2^5 \cdot 3^3 \cdot 5 \cdot 7 \cdot 11$
4	$24 = 2^3 \cdot 3$	12	$665{,}280 = 2^6 \cdot 3^3 \cdot 5 \cdot 7 \cdot 11$
5	$120 = 2^3 \cdot 3 \cdot 5$	13	$8{,}648{,}640 = 2^6 \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 13$
6	120	14	8,648,640
7	$840 = 2^3 \cdot 3 \cdot 5 \cdot 7$	15	8,648,640
8	$3{,}360 = 2^5 \cdot 3 \cdot 5 \cdot 7$		

specify generators and relations for Q_2 and A_4:

$$Q_2 = \langle a_1, a_2 \mid a_1^4 = a_2^4 = 1,\ a_1^2 = a_2^2,\ a_1^{-1}a_2a_1 = a_2^3\rangle;$$
$$A_4 = \langle b_1, b_2, b_3 \mid b_1^2 = b_2^2 = b_3^3 = 1,\ b_1b_2 = b_2b_1,\ b_3^{-1}b_1b_3 = b_2,\ b_3^{-1}b_2b_3 = b_1b_2\rangle.$$

We further let $C_9 = \langle c\rangle$, $C_5 = \langle d\rangle$, and $C_7 = \langle e\rangle$. We define an automorphism θ of the direct product $Q_2 \times A_4 \times C_9 \times C_5 \times C_7$ by specifying the images of these generators:

$$a_1^\theta = a_2;\ a_2^\theta = a_1;\ b_1^\theta = b_2;\ b_2^\theta = b_1;\ b_3^\theta = b_3^{-1};\ c^\theta = c^{-1};\ d^\theta = d^{-1};\ e^\theta = e^{-1}.$$

To verify that θ defines an automorphism of $Q_2 \times A_4 \times C_9 \times C_5 \times C_7$, it suffices to show that $Q_2 \times A_4 \times C_9 \times C_5 \times C_7 = \langle\, a_1^\theta,\ a_2^\theta,\ b_1^\theta,\ b_2^\theta,\ b_3^\theta,\ c^\theta,\ d^\theta,\ e^\theta\rangle$ and that our 'new' generators satisfy relations identical to those satisfied by the original generating set. It is then easy to see that θ has order 2.

We form the semidirect product $H_4 = (Q_2 \times A_4 \times C_9 \times C_5 \times C_7) \rtimes \langle\theta\rangle$ and note that $|H_4| = 2^6 \cdot 3^3 \cdot 5 \cdot 7 = 60{,}480$. We can then verify that all groups of 12 or less can be embedded in the group $H_4 \times C_{11}$, while all groups of order 15 or less can be embedded in $H_4 \times C_{11} \times C_{13}$. It is relatively straightforward to produce variations on our examples to show that these minimal groups are not unique.

Table 8 summarizes our current knowledge with regard to the minimal order of a group in which all groups of order n can be embedded. Our results indicate that, as n increases, the order of a minimal group in which all groups of order n or less can be embedded does not increase as fast as the order of the corresponding minimal symmetric group. For example, the table shows that the minimal order of a finite group in which all groups of order 13 or less can be embedded is 8,648,640, whereas the minimal order of a symmetric group in which the same groups can be embedded is $|S_{13}| = 13! = 6{,}227{,}020{,}800$. We may thus conclude that, while Cayley's theorem provides us with an elegant embedding of all groups of order n or less, it does so at the cost of a certain degree of efficiency.

Acknowledgments The authors would like to thank the anonymous referees, whose helpful suggestions and thoughtful reviews of an earlier version of this article significantly improved both its style and content.

REFERENCES

[1] Besche, H., Eick, B., O'Brien, E. A. (2002). A millennium project: Constructing small groups. *Int. J. Algebra Comput.* 12: 623–644.

[2] Besche, H., Eick, B., O'Brien, E. A. (2001). The groups of order at most 2000. *Electron. Res. Announc. Amer. Math. Soc.* 7: 1–4.

[3] Cauchy, A. (1845). ANALYSE MATHÉMATIQUE.—Sur le nombre des valeurs égales ou inégales que peut acquérir une fonction de n variables indépendants, quand on permute ces variables entre elles d'une maniière quelconque. *Compte Rendu des Séances de l'Académie des Sciences* XXI(11): 593–607.
[4] Cayley, A. (1854). On the theory of groups, as depending of the symbolic equation $\theta^n = 1$. *Philos. Mag.* VII: 40–47. Reprinted in *The Collected Mathematical Papers of Arthur Cayley*, Vol. II. Cambridge, England: Cambridge Univ. Pr., 1898.
[5] Cayley, A. (1878). Desiderata and suggestions. No. 1. The theory of groups. *Amer. J. Math.* I: 50–52. Reprinted in *The Collected Mathematical Papers of Arthur Cayley*, Vol. X. Cambridge, England: Cambridge Univ. Pr., 1898.
[6] Cayley, A. (1889). On the theory of groups. *Amer. J. Math.* XI: 139–157. Reprinted in *The Collected Mathematical Papers of Arthur Cayley*, Vol. XII. Cambridge, England: Cambridge Univ. Pr., 1898.
[7] Heffernan, R., MacHale, D., McCann, B. Minimal embeddings of small finite groups. arXiv:1706.09286 [math.GR].
[8] Johnson, D. L. (1971). Minimal permutation representations of finite groups. *Amer. J. Math.* 93: 857–866.
[9] Kleiner, I. (1986). The evolution of group theory: A brief survey. *Math. Mag.* 59: 195–215.
[10] Saunders, N. (2008). Minimal faithful permutation degrees of finite groups. *Aust. Math. Soc. Gazette* 35(2): 332–338.
[11] Saunders, N. (2009). Strict inequalities for minimal degrees of direct products. *Bull. Aust. Math. Soc.* 79(1): 23–30.
[12] The GAP Group. (2016). *GAP – Groups, algorithms, and programming*, Version 4.8.6. gap-system.org.
[13] Wright, D. (1975). Degrees of minimal embeddings for some direct products. *Amer. J. Math.* 97: 897–903.

Summary. Motivated by Cayley's theorem, we pose five elementary questions concerning embeddings of finite groups and answer them for groups of order 15 or less. Most of our results are easily verified. They provide a novel perspective on group theory that has its roots in the historical origins of the subject.

ROBERT HEFFERNAN (MR Author ID: 852327) received his Ph.D. in mathematics from University College Cork, Ireland under the direction of Des MacHale in 2009. He is now a lecturer in mathematics at the Cork Institute of Technology, Ireland.

DESMOND MacHALE (MR Author ID: 117490) received his Ph.D. in mathematics (group theory) from the University of Keele in the UK under the supervision of Hans Liebeck. He is currently Emeritus Professor of Mathematics at University College Cork in Ireland where he taught for nearly 40 years. His research interests include groups and rings, number theory, history of mathematics (he has written a biography of George Boole), and mathematical humor.

BRENDAN McCANN (MR Author ID: 248119) received his doctorate from the University of Würzburg, Germany in 1985 under the supervision of Hermann Heineken. His research interests lie mainly in the area of finite soluble groups. Since 1989 he has been a lecturer at the Waterford Institute of Technology, where he particularly enjoys teaching mathematics to engineering students.

Some Taylor Series without Taylor's Theorem

MARK DALTHORP
Crescent Valley High School
Corvallis, OR 97333
ashorocetus@yahoo.com

The Taylor series for e^x, $\sin(x)$, and $\cos(x)$ are perhaps the most frequently used in all of mathematics (and give a nice proof of Euler's formula), but rigorously deriving them is nontrivial and requires advanced material such as Taylor's theorem. However, the following theorem yields a simple derivation of these series using only basic calculus.

Theorem. *Let $f(x)$ be any function which is bounded and integrable on any bounded set. Define a sequence of functions, $F_n(x)$, recursively by $F_n(x) = \int_0^x F_{n-1}(t)dt$ with $F_0(x) = \int_0^x f(t)dt$. (When in the complex plane, these are defined to be the integrals over the straight line from 0 to x.) Then, $\lim_{n\to\infty} F_n(x) = 0$ for all x.*

Proof. Since $f(x)$ is bounded on bounded sets, given $r > 0$, there exists $c > 0$ such that $|f(x)| < c$ when $|x| \leq r$. Thus, $|F_0(x)| = |\int_0^x f(t)dt| \leq c|x|$, $|F_1(x)| \leq \frac{c|x|^2}{2}$, and, by induction, $|F_{n-1}(x)| \leq \frac{c|x|^n}{n!}$. Clearly, $\frac{c|x|^n}{n!}$ goes to 0 as n goes to ∞, so by the squeeze theorem, $F_n(x)$ also converges to 0 when $|x| \leq r$. Since this works for all r, $F_n(x)$ goes to 0 for any x in the complex plane. ■

Corollary. *The Taylor series of e^x, $\sin x$, and $\cos x$ are as follows:*

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \ \sin x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}, \ \textit{and} \ \cos x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}.$$

Proof. Let $f(x) = e^x$ and apply the Theorem. We can use induction to find $F_n(x) = e^x - \sum_{k=0}^{n} \frac{x^k}{k!}$. This is clear for $n = 0$. Then observe that $F_n(x) = e^x - \sum_{k=0}^{n} \frac{x^k}{k!}$ implies $F_{n+1}(x) = e^x - 1 - \sum_{k=1}^{n+1} \frac{x^{k+1}}{k+1} = e^x - \sum_{k=0}^{n+1} \frac{x^k}{k!}$. Then, by the theorem,

$$\lim_{n\to\infty} \left(e^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right) = 0.$$

Equivalently, $e^x = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!} = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. Starting with $f(x) = \sin x$ instead, we can use the same method to prove the series for $\sin x$ and $\cos x$. ■

Summary. We present an elementary derivation of the Taylor series of e^x, $\sin x$, and $\cos x$.

MARK DALTHORP (MR Author ID: 1233425, ORCID 0000-0002-4825-5833) is a student at Crescent Valley High School in Corvallis, Oregon. He does math compulsively all the time, even when he ought to be doing homework.

Math. Mag. **91** (2018) 112. doi:10.1080/0025570X.2017.1408980 © Mathematical Association of America
MSC: Primary 41A58, Secondary 97I30

Triangles in Wonderland

JOSEPH P. PREVITE
MICHELLE PREVITE
Pennsylvania State University
Erie, PA 16563
jpp4@psu.edu
mlp17@psu.edu

Charles Lutwidge Dodgson , whose *nom de plume* was Lewis Carroll, first proposed the following problem in his work *Pillow Problems Thought Out During Wakeful Hours* in 1893 [**1**]:

> Three points are taken at random on an infinite plane. Find the chance of their being the vertices of an obtuse-angled triangle.

He argued that the probability of an obtuse triangle was $\frac{3\pi}{8\pi-6\sqrt{3}} \approx 63.9\%$. However, Dodgson's original question is ill-posed, since it is impossible to generate even one random point on the plane where all points are equally likely. To be more precise, there is no translationally invariant probability measure on Euclidean space. The problem that Dodgson actually solved is a conditional probability problem, where we are first given a line segment $\overline{AB}$ that is the longest side of the triangle and then select the third point C at random. This limits the possible location of the third point to a bounded region. Guy modified Dodgson's problem by computing the probability of an obtuse triangle assuming that we first fix $\overline{AB}$ so that it is the *second* longest side of the triangle and all possible choices for the third vertex are equally likely. (Again, this assumption limits the possible location of the third point C to a bounded region.) The case in which $\overline{AB}$ is first fixed to be the shortest side of the triangle leads to an unbounded region for the possible locations of point C. In this case, it is not possible to make all possible locations of C equally likely. In his paper, Guy overcame this hurdle by restricting the location of the third point to a certain aesthetically pleasing bounded set (an ellipse), and then he computed the probability that $\triangle ABC$ is obtuse. To summarize, Dodgson's (and Guy's) calculations involve conditional probability, where one is first given information about the relative size of one edge of the triangle.

Other authors have considered Dodgson's question, but the manners in which they generate random triangles are different from Dodgson's approach. Hall [**4**] generated three random points inside the unit disk and computed the likelihood that they form an obtuse triangle. He also considered the case for higher dimensional disks. Of course, the probability of obtaining an obtuse triangle in this set up is different from Dodgson's, as the points are chosen differently. Eisenberg and Sullivan [**2**] obtain three random points in the plane by generating the coordinates of each point from a standard normal distribution. Unlike Dodgson's case, this allows for the possibility of any point in the plane, but then not all points are equally likely. They also considered the problem in higher dimensions by choosing three random points in n-dimensional Euclidean space by using a standard normal distribution to generate each coordinate. In his paper, Portnoy [**6**] suggests several natural ways to generate random triangles. One such scheme involves fixing one side of the triangle of a certain length and then generating

Math. Mag. **91** (2018) 113–120. doi:10.1080/0025570X.2017.1419769
MSC: Primary 51M04, Secondary 60C05
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

two random angles from a uniform distribution, while restricting the sum of these angles to be less than π. Strang [7] has an interesting lecture, where random triangles are generated three different ways leading to three different answers to the problem. Because there are numerous ways to reasonably construct random triangles, then the relative percentage of obtuse triangles can be considered in any of these situations.

In this paper, unlike those to which we refer above, we specifically seek to generalize Dodgson's approach in the two-dimensional case to higher dimensions. Namely, we assume that we are given a specific line segment $\overline{AB}$ that is the longest side of a triangle in n-dimensional space, and that all remaining possible locations for the third point C are equally likely. After reproducing Dodgson's argument in dimension two, we consider the problem in dimension three (using techniques found in a typical sophomore-level multivariable calculus class). Surprisingly, in contrast to the planar case, acute triangles become more likely than obtuse ones in three dimensions. This remarkable turnabout prompts us to analyze the same problem in any dimension n. The three-dimensional computation sets the stage for computing the probability in dimension n using hyperspherical coordinates. We prove that as the dimension increases, the likelihood of obtuse triangles approaches zero. Therefore, we make the assertion that there are more acute triangles than obtuse ones in any Euclidean space with dimension greater than 2.

The planar case

We reproduce Dodgson's argument in the plane. Without loss of generality, assume that the longest side of the triangle $\triangle ABC$ is the line segment from $A = (-\frac{D}{2}, 0)$ to $B = (\frac{D}{2}, 0)$ in the plane, where $D > 0$ is any real number. Therefore, the third point C must lie somewhere in the eye-shaped region R formed by the intersection of two disks with radius D centered at A and B, respectively (see Figure 1).

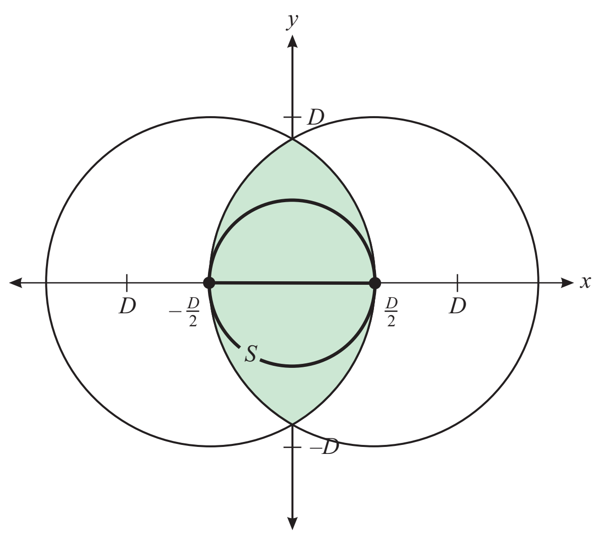


Figure 1 For $A = (-\frac{D}{2}, 0)$ and $B = (\frac{D}{2}, 0)$, a third point C must be chosen in the intersection R of the disks so that $\overline{AB}$ is the longest side of $\triangle ABC$. Choosing a third point C on the inner circle S produces a right triangle $\triangle ABC$. When the third point C is inside the inner circle S, then the triangle $\triangle ABC$ is obtuse. However, when C is a point outside the inner circle S (but inside the intersection R), then $\triangle ABC$ is acute.

Under the assumption that all points C in the region R are equally likely, one simply needs to find the area of the region that corresponds to all possible obtuse triangles and compute the ratio of this area to the area of R. This then will be the probability of an obtuse triangle $\triangle ABC$ if the third point C is chosen randomly, given that $\overline{AB}$ is the longest side.

A nice precalculus exercise using the Pythagorean theorem shows that the set of all points $C = (x, y)$ that make a right triangle $\triangle ABC$ with $\overline{AB}$ as hypotenuse is on the circle given by $x^2 + y^2 = (\frac{D}{2})^2$ (the inner circle in Figure 1). Moreover, by the law of cosines, all points C inside this circle create obtuse triangles $\triangle ABC$ with longest side $\overline{AB}$, while all points C chosen outside of this circle and inside the region R form acute triangles $\triangle ABC$ with $\overline{AB}$ as the longest side. The area of the inner circle shown in Figure 1 is $\frac{1}{4}D^2\pi$.

We now compute the area of the region R. Since the region R is symmetric about the y-axis, it suffices to compute the area to the right of the y-axis. If we translate the disk centered at A so that it is centered at the origin, the portion of R to the right of the y-axis becomes the shaded region in Figure 2 whose area can be easily computed using polar coordinates. Note that $x = \frac{D}{2}$ becomes $r = \frac{D}{2}\sec\theta$ in polar coordinates. So, the area of the region R is

$$2\int_{-\frac{\pi}{3}}^{\frac{\pi}{3}}\int_{\frac{D}{2}\sec\theta}^{D} r\,dr\,d\theta = D^2\left(\frac{2}{3}\pi - \frac{\sqrt{3}}{2}\right).$$

Taking the quotient of the areas of the inner circle in Figure 1 and the area of region R, the probability of an obtuse triangle is

$$\frac{\pi}{\frac{2}{3}\pi - \frac{\sqrt{3}}{2}} = \frac{3\pi}{8\pi - 6\sqrt{3}} \approx 0.6393825609,$$

implying that, in the plane, there are more obtuse than acute triangles.

Exercise. Reproduce this argument in the case when $\overline{AB}$ is the *second* longest side of the triangle. (This was done in [**3**].)

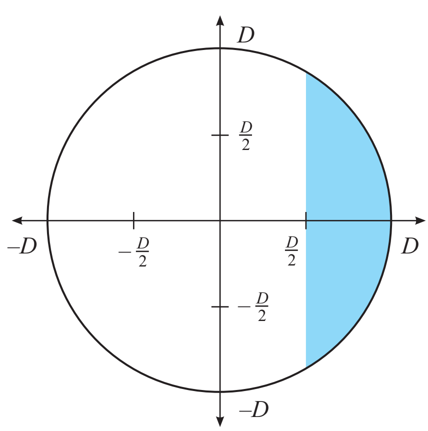


Figure 2 Half of the region R shifted to right. The area can be computed using polar coordinates.

The three-dimensional case

The situation is similar in higher dimensions. In order to make the n-dimensional case clearer, we handle the case in $\mathbb{R}^3$ separately before launching into higher dimensions. As before, we assume that the longest side of the triangle $\triangle ABC$ is positioned as the line segment from $A = (-\frac{D}{2}, 0, 0)$ to $B = (\frac{D}{2}, 0, 0)$. Then the allowable triangles are those whose third point C is in the intersection M of the two balls of radius D having centers A and B. See Figure 3.

One can again show via the Pythagorean theorem that the choices of $C = (x, y, z)$ that form a right triangle $\triangle ABC$ having longest side $\overline{AB}$ are on the sphere S given by $x^2 + y^2 + z^2 = (\frac{D}{2})^2$ and depicted as the dark sphere in Figure 3. Analogously to the two-dimensional case, obtuse triangles in $\mathbb{R}^3$ are formed by choosing the third point C of $\triangle ABC$ inside the sphere S. Observe that S has volume $V = \frac{4}{3}\pi(\frac{D}{2})^3 = \frac{1}{6}\pi D^3$.

Next, we compute the volume of the solid M associated with the allowable triangles. Here in three dimensions, the boundary of M is the union of two symmetric spherical caps. (A spherical cap is the smaller region bounded by a sphere and an intersecting plane.) Similar to the planar case, we shift the center of the left ball from Figure 3 to the origin in order to make the computation of the volume of M easier. See Figure 4. The volume inside the cap indicated in Figure 4, which is half the volume of M, is given by the spherical integral

$$\int_0^{2\pi} \int_0^{\frac{\pi}{3}} \int_{\frac{D}{2\cos\phi}}^{D} \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta = \frac{5}{24} D^3 \pi.$$

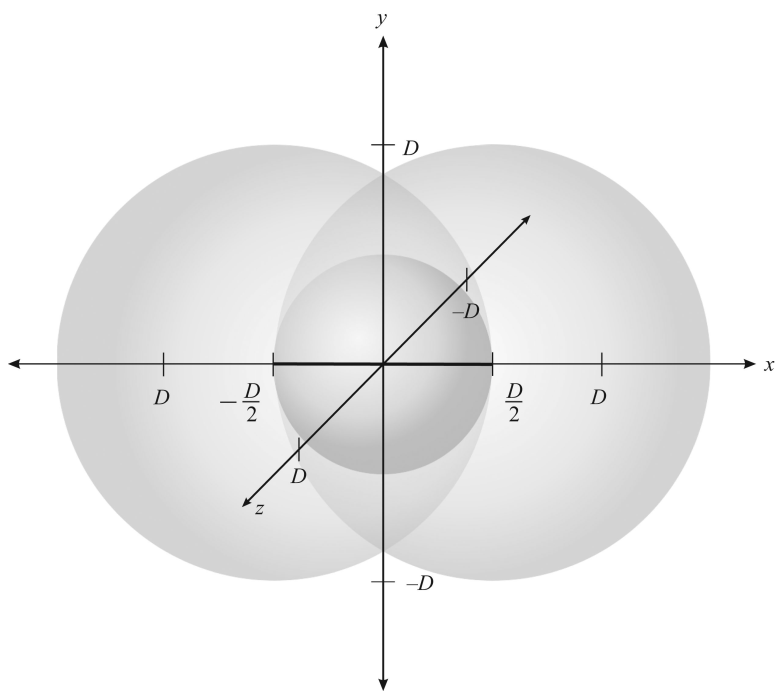


Figure 3 This figure depicts the intersection M of the balls of radius D centered at A and B. The line segment $\overline{AB}$ is bold and in the center of the figure. In order for $\overline{AB}$ to be the longest length of the triangle, C must be chosen in the intersection of these two balls. If C is chosen outside the inner sphere S, then $\triangle ABC$ is an acute triangle. If C is chosen inside S, then the triangle is obtuse.

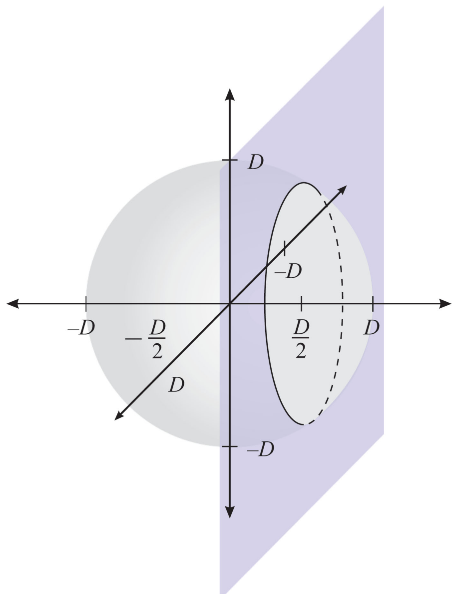


Figure 4 The spherical cap is formed by intersecting the plane $y = \frac{D}{2}$ and the sphere of radius D centered at the origin. It has half the volume of M.

So the volume associated with obtuse triangles divided by the volume associated with all allowable triangles is

$$\frac{\frac{1}{6}\pi D^3}{\frac{10}{24}D^3\pi} = \frac{2}{5}.$$

We conclude that, in three dimensions, acute triangles are more abundant than obtuse triangles, contrasting with the two-dimensional case.

Exercise. Compute the probability of an obtuse triangle in three dimensions, given that $\overline{AB}$ is the second longest side.

The *n*-dimensional case

As before, we can show that if the longest side of the triangle $\triangle ABC$ is positioned as the line segment from $A = (-\frac{D}{2}, 0, \ldots, 0)$ to $B = (\frac{D}{2}, 0, \ldots, 0)$, then the third point C must be inside the intersection of the hyperspheres $(x_1 - \frac{D}{2})^2 + x_2^2 + \cdots + x_n^2 = D^2$ and $(x_1 + \frac{D}{2})^2 + x_2^2 + \cdots + x_n^2 = D^2$. Once again, the Pythagorean theorem shows that the set of points $C = (x_1, \ldots, x_n)$ that form a right triangle $\triangle ABC$ with hypotenuse $\overline{AB}$ are on the hypersphere H given by $x_1^2 + x_2^2 + \cdots + x_n^2 = \frac{D^2}{4}$. As before, points C inside H form obtuse triangles, while points C outside H form acute triangles.

Next, we will use hyperspherical coordinates (see [8]), which generalize spherical coordinates to any dimension n. For $(x_1, \ldots, x_n) \in \mathbb{R}^n$, set $\rho = \sqrt{\sum_{i=1}^n x_i^2}$ and set

$$\begin{aligned}
x_1 &= \rho \cos\phi_1 \\
x_2 &= \rho \sin\phi_1 \cos\phi_2 \\
x_3 &= \rho \sin\phi_1 \sin\phi_2 \cos\phi_3
\end{aligned}$$

$$\begin{aligned}
\vdots &= \qquad \vdots \\
x_{n-1} &= \rho \sin\phi_1 \sin\phi_2 \cdots \sin\phi_{n-2} \cos\phi_{n-1} \\
x_n &= \rho \sin\phi_1 \sin\phi_2 \cdots \sin\phi_{n-2} \sin\phi_{n-1},
\end{aligned}$$

where $0 \le \phi_{n-1} < 2\pi$ and $0 \le \phi_j \le \pi$ for $1 \le j \le n-2$. Computing the determinant of the Jacobian transforms the volume element

$$dx_1\, dx_2 \cdots dx_n = \rho^{n-1} \sin^{n-2}(\phi_1) \sin^{n-3}(\phi_2) \cdots \sin(\phi_{n-2}) d\rho\, d\phi_1 \, \cdots \, d\phi_{n-1}.$$

Armed with this change of coordinates, we can compute the volumes of the hypersphere H, which corresponds to obtuse triangles, and of the larger region, which corresponds to all allowable triangles and consists of the interiors of two symmetric hyperspherical caps. The volume of H is

$$\begin{aligned}
&\int_0^{2\pi}\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{\frac{D}{2}} \rho^{n-1} \sin^{n-2}(\phi_1)\sin^{n-3}(\phi_2)\cdots\sin(\phi_{n-2}) d\rho\, d\phi_1 \cdots d\phi_{n-2}\, d\phi_{n-1} \\
&= \frac{2\pi D^n}{n 2^n}\left(\int_0^{\pi} \sin^{n-2}(\phi_1)\, d\phi_1\right)\left(\int_0^{\pi} \sin^{n-3}(\phi_2) d\phi_2\right)\cdots\left(\int_0^{\pi} \sin(\phi_{n-2}) d\phi_{n-2}\right).
\end{aligned}$$

The volume of the (shifted) hyperspherical cap (which is half of the required volume) is below. Since $x_1 = \frac{D}{2}$ transforms, as before, to $\rho = \frac{D}{2\cos\phi_1}$, we obtain

$$\begin{aligned}
&\int_0^{2\pi}\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{\frac{\pi}{3}}\int_{\frac{D}{2}\sec\phi_1}^{D} \rho^{n-1}\sin^{n-2}(\phi_1)\sin^{n-3}(\phi_2)\cdots \\
&\qquad\qquad \sin(\phi_{n-2}) d\rho\, d\phi_1\, d\phi_2 \cdots d\phi_{n-2}\, d\phi_{n-1} \\
&= 2\pi\left(\int_0^{\pi}\sin^{n-3}(\phi_2) d\phi_2\right)\cdots\left(\int_0^{\pi}\sin(\phi_{n-2}) d\phi_{n-2}\right) \\
&\quad \times\left[\int_0^{\frac{\pi}{3}}\int_{\frac{D}{2}\sec\phi_1}^{D} \rho^{n-1}\sin^{n-2}(\phi_1) d\rho\, d\phi_1\right] \\
&= 2\pi\left(\int_0^{\pi}\sin^{n-3}(\phi_2) d\phi_2\right)\cdots\left(\int_0^{\pi}\sin(\phi_{n-2}) d\phi_{n-2}\right) \\
&\quad \times\left[\int_0^{\frac{\pi}{3}} \frac{D^n}{n}\left(1 - \frac{1}{2^n}\sec^n(\phi_1)\right)\sin^{n-2}(\phi_1) d\phi_1\right].
\end{aligned}$$

Taking the ratio of the volumes and simplifying (and remembering to double the denominator), we obtain

$$\frac{\frac{1}{2^n}\left(\int_0^{\pi}\sin^{n-2}(\phi_1)\, d\phi_1\right)}{2\int_0^{\frac{\pi}{3}}\left(1 - \frac{1}{2^n}\sec^n(\phi_1)\right)\sin^{n-2}(\phi_1) d\phi_1}.$$

The integral $\int_0^{\frac{\pi}{3}} \sec^n(\phi_1)\sin^{n-2}(\phi_1) d\phi_1$ is equal to $\frac{(\sqrt{3})^{n-1}}{n-1}$ via straightforward u-substitution where $u = \tan(\phi_1)$. The remaining integrals are powers of sine for which we turn to the familiar recursive formula

$$\int_a^b \sin^n(x)\, dx = -\frac{1}{n}\sin^{n-1} x\cos x\Big|_a^b + \left(\frac{n-1}{n}\right)\int_a^b \sin^{n-2} x\, dx.$$

Table 1 shows the ratio of obtuse triangles to allowable triangles for $2 \le n \le 10$.

TABLE 1: Ratio of obtuse triangles in dimension n.

Dimension	Ratio
2	$\frac{3\pi}{8\pi-6\sqrt{3}} \approx 0.6393825609$
3	$\frac{4}{10} = 0.40$
4	$\frac{3\pi}{32\pi - 36\sqrt{3}} \approx 0.2468696972$
5	$\frac{8}{53} \approx 0.1509433962$
6	$\frac{15\pi}{640\pi - 864\sqrt{3}} \approx 0.09165800095$
7	$\frac{16}{289} \approx 0.05536332180$
8	$\frac{105\pi}{17920\pi - 26784\sqrt{3}} \approx 0.03329943276$
9	$\frac{128}{6413} \approx 0.01995945735$
10	$\frac{105\pi}{71680\pi - 114048\sqrt{3}} \approx 0.01192904992$

We next show that the ratio

$$\frac{\frac{1}{2^n}\left(\int_0^{\pi} \sin^{n-2}(\phi_1)\, d\phi_1\right)}{2\int_0^{\frac{\pi}{3}} \left(1 - \frac{1}{2^n}\sec^n(\phi_1)\right)\sin^{n-2}(\phi_1) d\phi_1}$$

tends to zero as $n \to \infty$. To see this, simply note that

$$\frac{\frac{1}{2^n}\left(\int_0^{\pi} \sin^{n-2}(\phi_1)\, d\phi_1\right)}{2\int_0^{\frac{\pi}{3}} \left(1 - \frac{1}{2^n}\sec^n(\phi_1)\right)\sin^{n-2}(\phi_1) d\phi_1} \le \frac{\frac{1}{2^n}\pi}{2\int_0^{\frac{\pi}{3}} \left(\cos(\phi_1) - \frac{1}{2^n}\sec^n(\phi_1)\right)\sin^{n-2}(\phi_1) d\phi_1}. \tag{1}$$

The right-hand side of inequality (1) equals

$$\frac{\pi(n-1)}{2(\sqrt{3})^{n-1}},$$

which tends to zero as $n \to \infty$. By the squeeze theorem, we deduce that the ratio of obtuse triangles to allowable triangles tends to zero as the dimension increases to infinity.

Final remarks, open problems, and projects

We have produced a recursive formula to compute the ratio of obtuse triangles to all allowable triangles in any dimension n, where it is assumed that we are first given the longest side $\overline{AB}$. This generalizes Dodgson's computation in two dimensions. We have also shown that as the dimension increases, the ratio of obtuse triangles to all allowable triangles tends to zero. As soon as we move to three dimensions, there is a minority of obtuse triangles, which is very unlike the planar case. In essence, we have that obtuse triangles in large dimensional Euclidean space are a rarity.

One could take this work forward in several directions. In [**5**], the question was solved for three random points forming a spherical triangle on the n-dimensional sphere. (Note that since the sphere is compact, then points can be generated randomly on the sphere with all points being equally likely.) A similar computation could be done on other compact surfaces and manifolds (e.g., the torus). For noncompact surfaces (e.g., the hyperbolic disk), an approach similar to the one taken in this paper could be studied. Namely, fix a geodesic segment $\overline{AB}$ on an n-dimensional manifold having fixed length and consider all points C on the manifold that form a geodesic triangle $\triangle ABC$ with longest side $\overline{AB}$. What is the probability that $\triangle ABC$ is an obtuse triangle?

Acknowledgments The authors would like to thank the referees for their helpful suggestions and guidance. One referee even provided us with code to generate the beautiful figures that appear in this manuscript.

REFERENCES

[1] Carroll, L. (1893). *Curiosa Mathematica, Part II: Pillow Problems*. London: MacMillan.

[2] Eisenberg, B., Sullivan, R. (1996). Random triangles n dimensions. *Amer. Math. Monthly.* 103: 308–318.

[3] Guy, R. (1993). There are three times as many obtuse-angled triangles as there are acute-angled ones. *Math. Mag.* 66(3): 175–179.

[4] Hall, G. R. (1982). Acute triangles in the n-ball. *J. Appl. Prob.* 19: 712–715.

[5] Maeda, Y. (2008). Seven types of random spherical triangle in $\mathbf{S}^n$ and their probabilities. In: Ito, H., Kano, M., Katoh, N., Uno, Y., eds. *Computational Geometry and Graph Theory*, *Lecture Notes In Computer Science*. Vol. 4535, Berlin: Springer-Verlag, pp. 119–126.

[6] Portnoy, S. (1994). A Lewis Carroll pillow problem: Probability of an obtuse triangle. *Stat. Sci.* 9(2): 279–284.

[7] Strang, G. (2009). Are random triangles acute or obtuse? http://blossoms.mit.edu/videos/lessons/are_random_triangles_acute_or_obtuse

[8] Weisstein, E. W. (2002). "Hypersphere." From MathWorld–A wolfram web resource. http://mathworld.wolfram.com/Hypersphere.html

Summary. In this paper, we deduce that most triangles in n-dimensional Euclidean space are acute for $n \geq 3$, unlike the case in the plane which was originally considered by Lewis Carroll in 1893 and further developed in THIS MAGAZINE by Richard Guy in 1993. Further, we produce a formula that computes the percentage of obtuse triangles in any dimension $n \geq 3$ using generalized spherical coordinates. These results are novel for $n \geq 3$ and are attainable by standard techniques taught in a typical undergraduate multivariable calculus course.

JOSEPH P. PREVITE (MR Author ID: 635962) is an associate professor of mathematics at Penn State Erie, The Behrend College. He did his graduate work at the University of Maryland under the direction of Michael Brin. He married fellow graduate student Michelle Previte (née Lubinsky) in 1996 and received his doctorate in 1997. His research interests include mathematical biology, topology, geometry, and dynamical systems. He and his wife have four children. Besides mathematics, he is active in advising the Christian group on campus.

MICHELLE PREVITE (MR Author ID: 714799) is an associate professor of mathematics at Penn State Erie, The Behrend College. She received her doctorate at the University of Maryland in 2000 under the direction of Henry King. Her research interests include topology and geometry. Other than teaching mathematics, doing research, and cleaning up after her husband, she homeschools their four children Joseph, Josiah, Leah, and Boaz.

Bisecting the Perimeter of a Triangle

ALLAN BERELE
STEFAN CATOIU
DePaul University
Chicago, IL 60614
aberele@depaul.edu
scatoiu@depaul.edu

It is well known that a median of a triangle divides it into two triangles of equal area, and that the three medians meet at the centroid. Hence, every triangle has at least one point through which pass three area-bisecting lines. Are there always others? The answer is yes: In 1972, J. A. Dunn and J. E. Pretty proved in [**10**] that the area-bisecting lines of a triangle remain tangent to a three-cusped closed curve, their envelope; see Figure 1. The subject remained dormant for almost 30 years until it was tackled again by an internet newsgroup [**14**]. They named the bisecting envelope a *deltoid*, and it has many striking properties, which we now list as (AB1)–(AB6). The label (AB*n*) refers to the *n*th property of the (area-bisecting) deltoid, a known object that is included here only for motivational purpose. It is different from the *perimeter-bisecting deltoid*, which is introduced in this article, and the proof of whose properties listed as (PB1)–(PB10) will be the main focus of this article. Both of these deltoids are different from the classical deltoid or the hypocycloid with three cusps that was discovered by Steiner in 1856. For more on Steiner's deltoid, see [**13**], [**17**], [**19**] or [**20**].

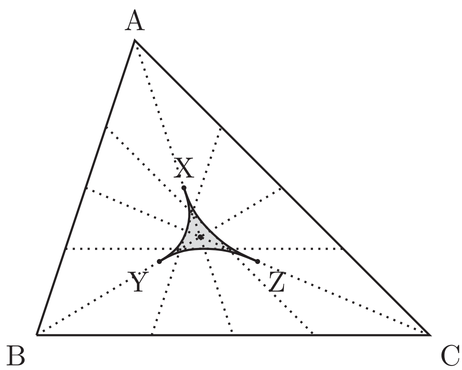


Figure 1 The area-bisecting deltoid of a triangle. The dotted lines bisect the area and are tangent to the deltoid.

Here are the promised properties of the area-bisecting deltoid.

(AB1) A line bisects the area of the triangle precisely when it is tangent to the deltoid. Of course, this is the definition of an envelope.

(AB2) The vertices of the deltoid are the midpoints of the medians; see [**2**].

(AB3) The sides of the deltoid are arcs of hyperbolas having the sides of the triangle as their asymptotes; see [**10**].

Math. Mag. **91** (2018) 121–133. doi:10.1080/0025570X.2017.1418589
MSC: Primary 51N20, Secondary 52A10; 53A04

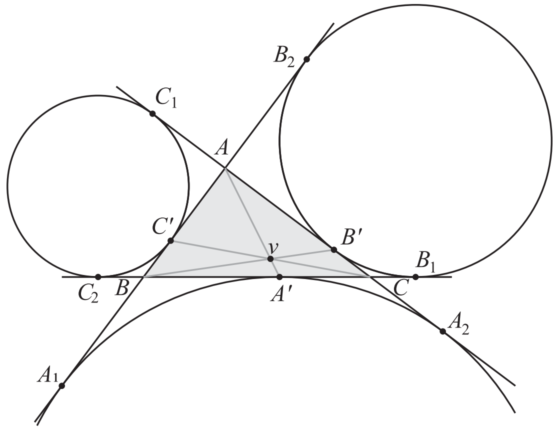


Figure 2 Triangle $\triangle ABC$ together with its escribed circles, their tangency points, and the cevians concurrent at the Nagel point ν.

(AB4) Each interior point of the deltoid is on three area-bisecting lines; each point on the boundary, except for the vertices, is on two; and every other point is on one; see [**10**].

(AB5) The area of the deltoid equals the area of the triangle times the constant $\frac{3}{4} \cdot \ln 2 - \frac{1}{2} = 0.019860\ldots$; see [**14**].

(AB6) The deltoid is also the locus of the midpoint of the area-bisecting lines; see [**2**] or [**7**].

For a triangle $\triangle ABC$, $a = BC$, $b = AC$ and $c = AB$ are the side lengths, and $s = (a+b+c)/2$ is the semiperimeter. The escribed circles, as shown in Figure 2, are those tangent to one side and to the extensions of the remaining two sides; see [**9**]. Let A', B', C' be their contact points with the sides. The following are known properties of this setting; see [**1**], [**11**] or [**12**]:

- the cevians AA', BB', CC' are concurrent at the Nagel point ν;
- each of the cevians AA', BB', CC' is a perimeter bisector; in particular, $A'C = s - b$ and $A'B = s - c$. These three perimeter-bisecting cevians of a triangle are called *splitters*, see [**11**, Chapter 1];
- if $A_1, A_2, B_1, B_2, C_1, C_2$ are the contact points of the escribed circles with the extensions of the sides, as shown in Figure 2, then $AA_1 = AA_2 = BB_1 = BB_2 = CC_1 = CC_2 = s$.

This notation will be kept throughout the article. A consequence of the first two properties is that the splitters of any triangle meet at the Nagel point, suggesting that the theory of area-bisecting lines in a triangle should have an analogous theory of perimeter-bisecting lines. The subject of the current article is to show that this is the case.

In Figure 3, we show a triangle with a large number of equally spaced perimeter-bisecting lines. It is evident that the perimeter-bisecting lines admit a deltoid-like envelope similar to the one for area-bisecting lines. Figure 4 shows the same triangle and the same deltoid. There is an essential difference between this deltoid and the one formed by the area-bisecting lines. Consider, for example, the side A_bC_a. It consists of two parts: from A_b to A_c is a line segment and from A_c to C_a is a segment of a parabola.

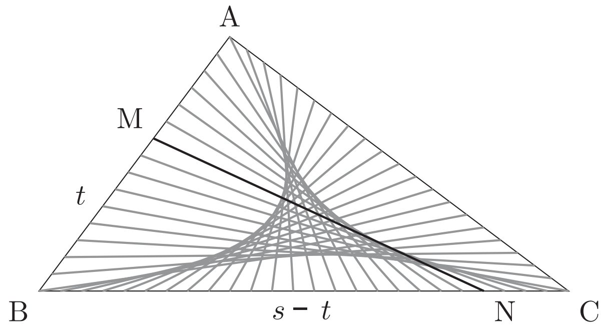


Figure 3 The perimeter-bisecting deltoid of a triangle. For a generic perimeter-bisecting line MN, with M on AB and N on BC, if $BM = t$ then $BN = s - t$.

Throughout the article, the perimeter-bisecting deltoid is simply referred to as "the deltoid" and denoted by $\mathcal{D}$. The following is the list of interesting properties of $\mathcal{D}$ that we will be proving in this article. Most properties can be visualized by looking at Figure 4.

(PB1) The envelope of the perimeter-bisecting lines of a triangle is a six-sided deltoid whose sides alternate between line segments and segments of parabolas. In Figure 4, the line segments are A_bA_c, B_aB_c and C_aC_b, and the parabolic segments are A_cC_a, B_cC_b and A_bB_a. Although the deltoid consists of segments of parabolas alternating with line segments, its "sides" behave differently according to the side lengths of the triangle: that is, if triangle $\triangle ABC$ has long side BC and short side AB, the deltoid has sides A_bB_a that is only a segment of a parabola; B_aC_a that is a smooth curve consisting of a line segment B_aB_c, an arc of parabola B_cC_b and another line segment C_bC_a; and finally, the last side is a parabola segment C_aA_c connected smoothly to a line segment A_cA_b.

(PB2) The three line segments lie along the splitters and their positions can be described precisely. For example, the position of A_b and A_c on AA' is given by the relations

$$\frac{A_bA}{A_bA'} = \frac{CA'}{CA} = \frac{s-b}{b} \quad \text{and} \quad \frac{A_cA}{A_cA'} = \frac{BA'}{BA} = \frac{s-c}{c}.$$

Property (PB2′) at the end of Section 2 provides an alternative way of locating the same points.

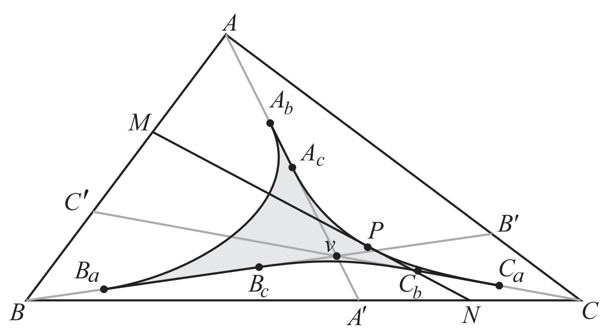


Figure 4 The deltoid labeled as $\mathcal{D} = A_bB_aB_cC_bC_aA_c$. Here $BC > AC > AB$. Label A_b is given to the vertex on splitter AA' that is connected to a vertex on splitter BB'.

(PB3) The sides of the deltoid make cusps at points A_b, B_a, C_a that are closer to A, B, C than they are to ν, and they connect smoothly at points A_c, B_c, C_b that are closer to ν than they are to A, B, C.

(PB4) The parabolas extending the parabolic segments have the angle bisectors as their symmetry axes. They pass through the contact points of the escribed circles with the extensions of the sides, and they are tangent to these extensions at these points; see Figure 5.

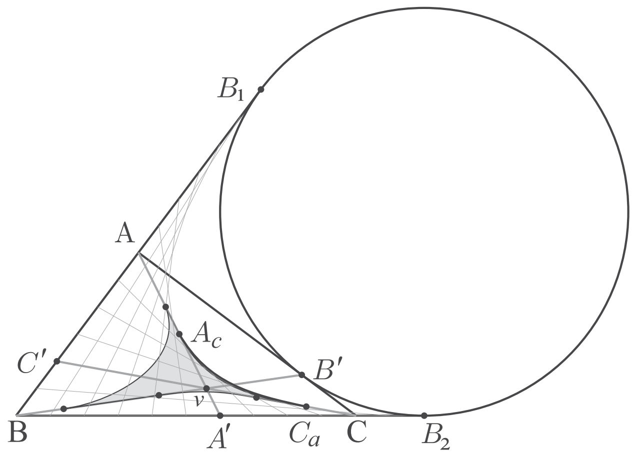


Figure 5 Property (PB4): The parabola extending arc $\widetilde{A_cC_a}$ is tangent to the sides AB and BC at the contact points with the escribed circle.

(PB5) The arc of parabola that opens toward the longest side of the triangle connects smoothly at both endpoints. The arc of parabola corresponding to the shortest side makes cusps at both endpoints. The third parabolic arc connects the remaining two points, hence it has both smooth and cusp connections at its endpoints; see Figure 4.

(PB6) The original triangle being isosceles is equivalent to one of the three line segment sides of the deltoid collapsing to a point. If two of these line segment sides collapse to a point, then so does the third. This is the case when the triangle is equilateral.

(PB7) In each direction, there is a unique perimeter-bisecting line MN. Segment MN divides the triangle into a smaller triangle and a quadrilateral. The deltoid is the locus described by the point P on MN that divides it in a ratio inversely proportional to the ratio of the other two sides of the smaller triangle; in Figure 4, we have $\frac{PM}{PN} = \frac{BN}{BM}$. Segments A_bA_c, B_aB_c and C_aC_b complete the contour of $\mathcal{D}$.

(PB8) Each point in the interior of the deltoid lies on exactly three perimeter-bisecting lines; each point on the boundary, except for the three cuspoidal vertices, lies on two; and every other point lies on just one.

(PB9) With the usual notation for sides, angles, and the semiperimeter of triangle $\triangle ABC$, the area of its deltoid is given by the formula

$$\frac{1}{6s} \cdot [(s-a)^3 \sin A + (s-b)^3 \sin B + (s-c)^3 \sin C].$$

(PB10) Unlike the deltoid of the area-bisecting lines, the fraction of the total area enclosed by $\mathcal{D}$ is not a constant. We have

$$\frac{1}{12} \leq \frac{\text{Area}(\mathcal{D})}{\text{Area}(ABC)} < \frac{1}{3},$$

where equality with the lower bound occurs if and only if the triangle is equilateral. The upper bound is the best possible for proper triangles; it is attained only at degenerated triangles.

The first section contains the basic geometric and analytic properties of the parabola that will be used in the subsequent sections. The geometric properties (PB1)–(PB8) of the perimeter-bisecting deltoid $\mathcal{D}$ are proved in the second section. The remaining two sections are respectively devoted to proving (PB9) and (PB10).

Recent developments The area-bisecting deltoid has been employed in [**5**] to give a simple proof to a classical theorem of convex geometry, while the perimeter-bisecting deltoid was found useful in [**4**] in giving a more precise answer to another classical theorem of the same subject. Article [**6**] solves a related open question dating back to 1949. Both deltoids were extended to area and perimeter deltoids of general convex polygons in [**7**]. A new reflection of the general theory of the quadratic equation into geometry that involves the simultaneous bisection of the area and the perimeter is given in [**3**]. All triangular area and perimeter bisecting deltoids are classified in [**8**].

Single angle envelopes

When studying the envelope of the perimeter-bisecting lines of a triangle $\triangle ABC$, we imagine a generic such line segment MN, with its endpoints traveling along the perimeter; see first Figure 4 and then Figure 6. The journey is divided into three portions by three bumps that occur when one of M, N is a vertex of the given triangle. During each of the three portions, M and N travel on either side of a fixed angle of the triangle. Our problem is then reduced to studying the envelope of segments MN when M, N travel along the sides of a fixed angle with vertex at O and such that the sum $OM + ON = \sigma$ is constant. We call this the *single angle envelope* $\mathcal{E}(\angle O, \sigma)$. Since envelopes of congruent angles are congruent, we denote by $\mathcal{E}(\theta, \sigma)$ the envelope of any angle of measure θ and sum σ.

The envelope as a geometric locus Let $M'N'$ be another instance of the same variable segment MN considered above (see Figure 7), and let P be the point of intersection of the two segments. Menelaus' Theorem (see [**1**] or [**12**]) applied to triangle $\triangle OMN$ and

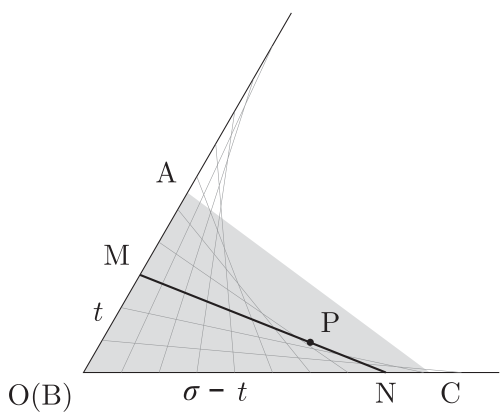


Figure 6 The single angle envelope $\mathcal{E}(\angle O, \sigma)$ of segments MN with $OM + ON = \sigma$. A generic such segment is parametrized by $t = OM$, with $0 \le t \le \sigma$. When $\sigma = s$ and $\angle O = \angle B$, the triangle $\triangle ABC$ fits inside the angle.

transversal $M'PN'$ yields

$$\frac{PM}{PN} \cdot \frac{OM'}{MM'} \cdot \frac{NN'}{ON'} = 1.$$

The condition $\sigma = OM + ON = OM' + ON'$ implies $MM' = NN'$. The above relation is then equivalent to

$$\frac{PM}{PN} = \frac{ON'}{OM'}.$$

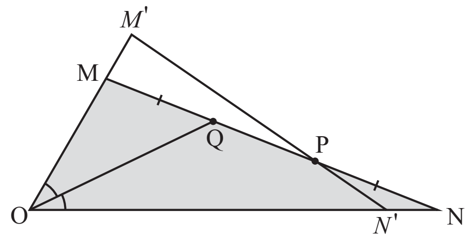


Figure 7 Triangle $\triangle OMN$ with transversal $M'PN'$ and angle bisector OQ.

When M', N' are close to M, N we have $ON' \approx ON$ and $OM' \approx OM$. Then the envelope of all segments MN with $OM + ON = \sigma$ is (approximated by, hence it coincides with) the locus of points $P \in MN$, with

$$\frac{PM}{PN} = \frac{ON}{OM}. \tag{1}$$

In other words, a variable point P of the envelope determines on the tangent MN two segments, PM and PN, that are inversely proportional to the segments OM and ON determined by the tangent on the sides of the angle.

If Q is the intersection of MN with the bisector of angle $\angle O$, then the angle bisector theorem makes $QM/QN = OM/ON$ so, by (1), P and Q will be equidistant from the midpoint of MN, or *isotomic* points of MN. For more on isotomic points, see [**1**] or the last part of Section 2.

The equation of the envelope We set up a coordinate system centered at O such that the y-axis is a symmetry axis of the given angle of measure 2α; see Figure 8 which represents a rotation about O of the picture in Figure 6. If $\mathbf{i}$, $\mathbf{j}$ are the usual unit vectors of the axes and $t \in [0, \sigma]$ is a parameter, then

$$\overrightarrow{OM} = t(-\sin\alpha\mathbf{i} + \cos\alpha\mathbf{j}) \text{ and } \overrightarrow{ON} = (\sigma - t)(\sin\alpha\mathbf{i} + \cos\alpha\mathbf{j}).$$

The point P given by (1) has position vector $\overrightarrow{OP} = (1-\lambda)\overrightarrow{OM} + \lambda\overrightarrow{ON}$, where $\lambda = PM/MN = ON/(OM + ON) = (\sigma - t)/\sigma$. Consequently,

$$\begin{aligned}\overrightarrow{OP} &= \frac{t^2}{\sigma}(-\sin\alpha\mathbf{i} + \cos\alpha\mathbf{j}) + \frac{(\sigma - t)^2}{\sigma}(\sin\alpha\mathbf{i} + \cos\alpha\mathbf{j}) \\ &= (\sigma - 2t)\sin\alpha\mathbf{i} + \left(\sigma - 2t + \frac{2t^2}{\sigma}\right)\cos\alpha\mathbf{j}.\end{aligned}$$

Then, the parameterization of the curve, for $t \in [0, \sigma]$, is

$$\begin{cases} x = (\sigma - 2t)\sin\alpha, \\ y = \left(\sigma - 2t + \frac{2t^2}{\sigma}\right)\cos\alpha. \end{cases} \tag{2}$$

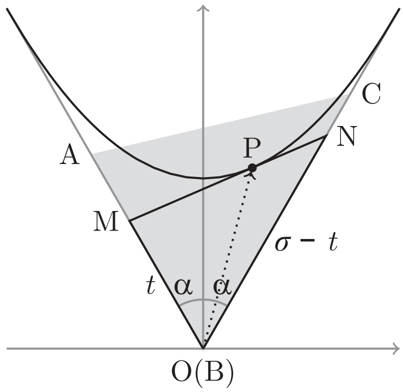


Figure 8 The graph of the angle envelope $\mathcal{E}(2\alpha, \sigma)$. When $\sigma = s$ and $\angle B = \angle O = 2\alpha$, the triangle $\triangle ABC$ fits inside the angle as shown.

Eliminating t/σ between x and y yields the Cartesian equation of the curve

$$y = \frac{\cos\alpha}{2\sigma \sin^2\alpha} \cdot x^2 + \frac{\sigma\cos\alpha}{2}, \qquad \text{for } x \in [-\sigma \sin\alpha, \sigma \sin\alpha]. \tag{3}$$

This is an arc of a parabola and P is the point of contact of the tangent line with the parabola. Taking $t = 0, \sigma$ shows that the parabola is tangent to the sides of the angle.

The particular envelope of the angle whose sides are the coordinate axes and $\sigma = 1$ was considered in [**20**]; see also [**16**] and [**18**].

Describing the deltoid

By the previous section, the envelope of the perimeter-bisecting lines of a triangle $\triangle ABC$ is given by the union of three arcs of parabolas. These arcs are cut from the single angle envelopes $\mathcal{E}(\angle A, s)$, $\mathcal{E}(\angle B, s)$, and $\mathcal{E}(\angle C, s)$ by the splitters AA', BB', and CC'. (PB1) is now clear.

Our next goal is to locate the endpoints of the three arcs on the splitters. We keep the setting of the last section and refer to Figure 4, with MN a variable perimeter-bisecting line, and P is the point on MN that generates the envelope. If M approaches A from the direction of C, then N approaches A' from the direction of B. By the previous section, the point P of envelope $\mathcal{E}(\angle C, s)$ approaches one end of its arc, the point A_b on AA'. The first equation in (PB2) follows from (1); the second is similar. Since the tangents at the endpoints of the parabolic arcs meet at ν, the arcs form cusps with the segments from ν to their endpoints. (PB3) follows from here.

The parabola of Figures 6 and 8 is symmetric with respect to the bisector of angle $\angle O = \angle B$. The distances from the vertex of the angle to the points, where it touches the angle are both equal to s; this corresponds to the extreme values $t = 0$ and $t = s$ of the parameter t. By the third bullet property in the Introduction, the two points are the points where the escribed circle touches the sides of the angle. This finishes (PB4).

(PB5) is a rephrase of (PB3) that is independent of picture. Indeed, by (PB2) the inequality $A_bA \leq A_cA$ is equivalent to $b \geq c$, with two similar equivalences. Thus, the order in which the points A_b and A_c are located on $A\nu$ depends on the inequalities between the sides adjacent to A. The correspondence of equalities and transitivity of the two equivalent equality relations prove (PB6).

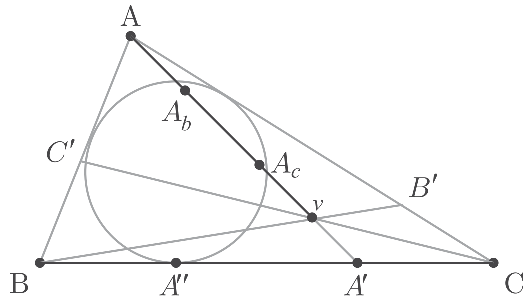


Figure 9 The situation in Property (PB2′).

(PB7) is a consequence of (1). For example, if P is as shown on Figure 4 then $PM/PN = BN/BM$, and if P is either of A_b or A_c, then $MN = AA'$ and the ratio PM/PN follows from the two equations in (PB2).

The proof of (PB8) is based on a case-by-case discussion depending on the regions, arcs, and points determined by the deltoid and the splitters. One important observation is that the perimeter-bisecting lines through a point are precisely the tangents from the point to the perimeter of the deltoid. For example, the interior of the deltoid is a union of interiors of three parabolic triangles one of which is $\widetilde{\Delta}\nu A_c C_a$ (see Figure 4), that have two straight sides and one side that is a segment of a parabola, three line segments one of which is νA_c, and one point, the Nagel point ν. From each interior point of $\widetilde{\Delta}\nu A_c C_a$, one can draw two tangents to the arc $\widetilde{A_c C_a}$ and one tangent to exactly one of the other two parabolic arcs, depending on which half-plane determined by BB' the point belongs to. From points on νA_c one can draw three tangents to $\mathcal{D}$: AA' and one other tangent to each of $\widetilde{A_b B_a}$ and $\widetilde{A_c C_a}$. Since AA', BB' and CC' are the only tangents from ν to the deltoid, we deduce that the interior points of $\mathcal{D}$ are intersections of three perimeter bisecting lines. The rest of (PB8) is proved in a similar manner.

Recall that the property (PB2) locates the vertices of the deltoid on the splitters and involves the contacts of the escribed circles with the sides. The last part of the section is devoted to an alternative way (PB2′) of locating the same points. It involves the contacts of the incircle with the sides.

For simplicity, we only locate the points A_b and A_c on the splitter AA'. Let A'' be the point where the inscribed circle touches the side BC; see Figure 9. Recall that AA' being a splitter means that $A'B = s - c$ and $A'C = s - b$. It is well-known (see [**1**, Theorem 160]) that A' and A'' are isotomic points of the side BC, and $A''B = s - b$, $A''C = s - c$. The point ν is still the Nagel point. The following property of the deltoid says that the configuration of points A, A_b, A_c, ν is a scaling of the configuration of points B, A'', A', C. In particular, A_b and A_c are isotomic points of segment $A\nu$.

(PB2′) The vertices A_b and A_c of the deltoid are located on the splitter AA' so that they divide the segment $A\nu$ in the same ratios as the points A'' and A' divide the side BC. Specifically, we have

$$\frac{A_bA}{A_b\nu} = \frac{A_c\nu}{A_cA} = \frac{s-b}{s-c} = \frac{A'C}{A'B} = \frac{A''B}{A''C}.$$

To prove this, we begin by making derived proportions in the first relation in (PB2) and deduce that

$$\frac{A_bA}{AA'} = \frac{A_bA}{A_bA' + A_bA} = \frac{s-b}{(s-b)+b} = \frac{s-b}{s}. \tag{4}$$

Van Aubel's relation (see [1, Section 342]) says that the ratio each of three concurrent cevians is divided by their intersection point is equal to the sum of the ratios the other two cevians divide their opposite sides. In particular, the Nagel point ν divides AA' in ratio

$$\frac{\nu A}{\nu A'} = \frac{B'A}{B'C} + \frac{C'A}{C'B} = \frac{s-c}{s-a} + \frac{s-b}{s-a} = \frac{a}{s-a}.$$

Equivalently,

$$\frac{\nu A}{AA'} = \frac{\nu A}{\nu A' + \nu A} = \frac{a}{(s-a)+a} = \frac{a}{s}. \tag{5}$$

Division of the equations of the first and last terms in (4) and (5) yields

$$\frac{A_bA}{\nu A} = \frac{s-b}{a},$$

which by derived proportions is equivalent to

$$\frac{A_bA}{A_b\nu} = \frac{A_bA}{A\nu - A_bA} = \frac{s-b}{a-(s-b)} = \frac{s-b}{s-c}.$$

Similarly, we have $\frac{A_cA}{A_c\nu} = \frac{s-c}{s-b}$. The last two equalities of (PB2$'$) are clear.

Area of the deltoid

In order to deduce the expression (PB9) for the area of the deltoid, we first look at the area of the parabolic triangle $\widetilde{\triangle}\nu A_cC_a$ (see Figure 4), whose one side is the arc $\widetilde{A_cC_a}$. We arrange triangle $\triangle ABC$ to fit angle $\angle O$ as shown in Figure 8, with vertex B in place of O and points M on AB and N on BC. Then, $\alpha = \angle B/2$. When the value of the parameter t is $c = AB$, then $M = A$, $N = A'$, and $P = A_c$, where A' and A_c are the same as before; we record this in Figure 10. The Cartesian coordinates of any point P in the plane are

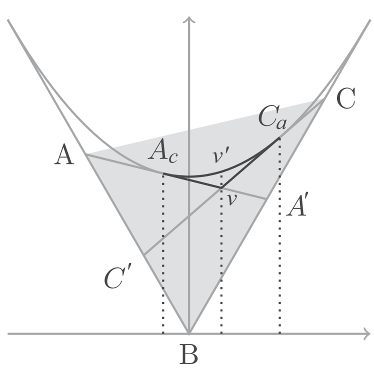


Figure 10 Triangle $\triangle ABC$ with a part of its deltoid, the parabolic triangle $\widetilde{\triangle}\nu A_cC_a$.

denoted by the pair (x_P, y_P). By (2), the point A_c has coordinates

$$\begin{cases} x_{A_c} = (s-2c)\sin\alpha, \\ y_{A_c} = \left(s - 2c + \frac{2c^2}{s}\right)\cos\alpha. \end{cases} \tag{6}$$

Similarly, when $t = s - a$, $M = C'$, $N = C$, and $P = C_a$ has coordinates

$$\begin{cases} x_{C_a} = (2a-s)\sin\alpha, \\ y_{C_a} = \left(s - 2a + \frac{2a^2}{s}\right)\cos\alpha. \end{cases} \tag{7}$$

On the other hand, differentiating with respect to x in equation (3) gives the slope m_P of the parabola at its variable point $P = P(x, y)$,

$$m_P = \frac{dy}{dx} = \frac{\cos\alpha}{s\sin^2\alpha} \cdot x. \tag{8}$$

In particular, by (6) and (7),

$$\begin{aligned} m_{A_c} &= \frac{s-2c}{s} \cdot \cot\alpha, \\ m_{C_a} &= \frac{2a-s}{s} \cdot \cot\alpha. \end{aligned} \tag{9}$$

From here, one can deduce the equations of the tangent lines to the parabola at A_c and C_a

$$\begin{aligned} y &= \frac{\cot\alpha}{s} \cdot [(s-2c)x + 2c(s-c)\sin\alpha] \\ y &= \frac{\cot\alpha}{s} \cdot [(2a-s)x + 2a(s-a)\sin\alpha] \end{aligned} \tag{10}$$

and their intersection point ν has coordinates

$$\begin{cases} x_\nu = (a-c)\sin\alpha, \\ y_\nu = \left(a + c - \frac{2ac}{s}\right)\cos\alpha. \end{cases} \tag{11}$$

The simple observation that

$$x_\nu = \frac{x_{A_c} + x_{C_a}}{2} \tag{12}$$

allows the area computation for $\widetilde{\triangle}\nu A_c C_a$ by evaluation rather than integration. Indeed, we define

$$\Delta_x = x_\nu - x_{A_c} = (a + c - s)\sin\alpha \tag{13}$$

and let ν' be the point on arc $\widetilde{A_c C_a}$ such that $x_{\nu'} = x_\nu$. Substituting $x = x_\nu$ in (3) yields

$$y_{\nu'} = \frac{a^2 - 2ac + c^2 + s^2}{2s} \cdot \cos\alpha. \tag{14}$$

We can now deduce the area of the parabolic triangle $\widetilde{\triangle}\nu A_c C_a$.

$$\text{Area}(\widetilde{\triangle}\nu A_c C_a) = \text{Area}(\text{below}\widetilde{A_c C_a}) - \text{Area}(\text{below} A_c \nu) - \text{Area}(\text{below } \nu C_a). \tag{15}$$

We compute Area(below$\widetilde{A_c C_a}$) by Simpson's Rule, see [**15**, Section 8.7], which is exact for any parabola. The remaining terms are areas of trapezoids. Putting it together, we get

$$\text{Area}(\widetilde{\triangle}\nu A_c C_a) = \frac{\Delta_x}{3}[y_{A_c} + 4y_{\nu'} + y_{C_a}] - \frac{\Delta_x}{2}[y_{A_c} + y_\nu] - \frac{\Delta_x}{2}[y_\nu + y_{C_a}]$$
$$= \frac{\Delta_x}{6}[-y_{A_c} + 8y_{\nu'} - y_{C_a} - 6y_\nu]. \quad (16)$$

Using (6), (7), (11), and (14), this simplifies to

$$\text{Area}(\widetilde{\triangle}\nu A_c C_a) = \frac{(a+c-s)^3}{3s}\sin\alpha\cos\alpha = \frac{(s-b)^3}{6s}\sin B. \quad (17)$$

Similar formulae are deduced for the areas of $\widetilde{\triangle}\nu A_b B_a$ and $\widetilde{\triangle}\nu B_c C_b$. Addition of all three yields the formula for the area of the deltoid $\mathcal{D}$ given in (PB9):

$$\text{Area}(\mathcal{D}) = \frac{(s-a)^3\sin A + (s-b)^3\sin B + (s-c)^3\sin C}{6s}. \quad (18)$$

The range of the area ratio

We employed formula (18) to compute the ratio $\text{Area}(\mathcal{D})/\text{Area}(ABC)$ for various triangles $\triangle ABC$. Our findings are included in the following table.

ΔABC type	$\frac{\text{Area}(\mathcal{D})}{\text{Area}(ABC)}$
equilateral	$\frac{1}{12} \approx .083333$
isosceles 5-5-6	$\frac{53}{600} \approx .088333$
isosceles right	$\frac{19-13\sqrt{2}}{6} \approx .102537$
3-4-5 right	$\frac{59}{540} \approx .109259$
isosceles 13-13-24	$\frac{7492}{50700} \approx .147771$
5-12-13 right	$\frac{5428}{35100} \approx .154644$

In particular, unlike the area-bisecting deltoid, the perimeter-bisecting deltoid of a triangle does not enclose a fixed fraction of the whole area. If one leg of a right triangle ΔABC is constant, then

$$\lim_{s\to\infty} \frac{\text{Area}(\mathcal{D})}{\text{Area}(ABC)} = \frac{1}{3}.$$

If triangle ΔABC is assumed isosceles type 1-1-2t, then

$$\lim_{t\to 1} \frac{\text{Area}(\mathcal{D})}{\text{Area}(ABC)} = \frac{1}{6} \quad \text{and} \quad \lim_{t\to 0} \frac{\text{Area}(\mathcal{D})}{\text{Area}(ABC)} = \frac{1}{3}.$$

Based on the above computations, it would seem that for any triangle ΔABC, we have

$$\frac{1}{12} \leq \frac{\text{Area}(\mathcal{D})}{\text{Area}(ABC)} < \frac{1}{3}, \tag{19}$$

with equality if and only if the triangle is equilateral.

Indeed, this is actually not hard to prove. We multiply both numerator and denominator of equation (18) by abc and factor out the area of the triangle. The resulting identity is

$$\text{Area}(\mathcal{D}) = \text{Area}(ABC) \cdot \frac{a(s-a)^3 + b(s-b)^3 + c(s-c)^3}{3abcs}. \tag{20}$$

Substituting $x = s - a$, $y = s - b$, and $z = s - c$ here yields the area ratio

$$\frac{\text{Area}(\mathcal{D})}{\text{Area}(ABC)} = \frac{x^3(y+z) + y^3(x+z) + z^3(x+y)}{3(y+z)(x+z)(x+y)(x+y+z)}. \tag{21}$$

The fraction is symmetric in x, y, z and subject to the conditions $x, y, z > 0$. The maximum and minimum can be computed using standard techniques from multivariable calculus. Here is a more elegant proof using the theory of symmetric functions. The right side in (21) expressed in terms of the elementary symmetric polynomials $e_1 = x + y + z$, $e_2 = xy + xz + yz$, and $e_3 = xyz$ is

$$\frac{1}{3} - \frac{2e_2^2}{3e_1(e_1e_2 - e_3)}. \tag{22}$$

The second inequality in (19) is then equivalent to $e_1e_2 - e_3 > 0$, which is true due to e_3 being one of the terms in the expansion of e_1e_2. The first inequality in (19) is equivalent to $\frac{2e_2^2}{3e_1(e_1e_2-e_3)} \leq \frac{1}{4}$, which is the same as

$$8e_2^2 + 3e_1e_3 \leq 3e_1^2e_2. \tag{23}$$

The *means inequality* $\dfrac{x+y+z}{3} \geq \dfrac{3}{\frac{1}{x}+\frac{1}{y}+\frac{1}{z}}$ implies

$$9e_3 \leq e_1e_2 \tag{24}$$

and the *sum of squares inequality* $(x-y)^2 + (y-z)^2 + (z-x)^2 \geq 0$ implies

$$3e_2 \leq e_1^2. \tag{25}$$

Multiplying (24) by $\frac{1}{3}e_1$ and (25) by $\frac{8}{3}e_2$ and adding the results yields (23). Both the means and the sum of squares inequalities give equality precisely when $x = y = z$, that is when the triangle is equilateral, and so does (23). This finishes the proof of (PB10).

Acknowledgment We thank the referees for their careful reading and suggestions for improvement.

REFERENCES

[1] Altshiller-Court, N. (1952). *College Geometry. An Introduction to the Modern Geometry of the Triangle and the Circle*. Mineola, NY: Dover Publications, Inc., Reprint of the second edition published by Barnes and Noble, Inc., New York, 1952.

[2] Ball, D. (1980). Halving envelopes. *Math. Gaz.* 64(429):166–173. MR0594844

[3] Berele, A., Catoiu, S. (2016). Area and perimeter bisecting lines of a triangle. *College Math. J.* 47(1):19–28. MR3482365

[4] Berele, A., Catoiu, S. (2017). The perimeter sixpartite center of a triangle. *J. Geom.* 108(3):861–868.

[5] Berele, A., Catoiu, S. (2018). The centroid as a nontrivial area bisecting center of a triangle. *College Math. J.* 49(1):27–34.
[6] Berele, A., Catoiu, S. (2018). Nonuniqueness of sixpartite points. *Amer. Math. Monthly*, to appear.
[7] Berele, A., Catoiu, S. (under review). *Bisecting Envelopes of Convex Polygons*.
[8] Berele, A., Catoiu, S. (in progress). *The Classification of Triangular Deltoids*.
[9] Berele, A., Goldman, J. (2001). *Geometry Theorems and Constructions*. Prentice Hall.
[10] Dunn, J. A., Pretty, J. E. (1972). Halving a triangle. *Math. Gaz.* 56(396):105–108.
[11] Honsberger, R. (1975). *Episodes in Nineteenth and Twentieth Century Euclidean Geometry*. New Mathematical Library **37**, The Mathematical Association of America.
[12] Johnson, R. A. (2007). *Advanced Euclidean Geometry*. Mineola, NY: Dover Publications. Reprint of the 1960 Dover edition of the original publication: *Modern Geometry*, Houghton Mifflin, Boston, 1929.
[13] Lockwood, E. H. (1967). *A Book of Curves*. New York: Cambridge University Press. MR0126191 Reprinted by Cambridge Univ. Press, Cambridge, 2007. MR2431517
[14] *Topics: 3-Points of a Triangle, Euler's Line, Bisecting Lines*, geometry.puzzles, The Math Forum@Drexel, October–December 2001.
[15] Stewart, J. (2012). *Calculus*. 7th ed. , Brooks/Cole.
[16] Weisstein, E. W. "Envelope" From MathWorld-A Wolfram Web Resource. Available at: `http://mathworld.wolfram.com/Envelope.html`.
[17] Weisstein, E. W. "Deltoid" From MathWorld-A Wolfram Web Resource. Available at: `http://mathworld.wolfram.com/Deltoid.html`.
[18] Envelope (mathematics) - Wikipedia, the free encyclopedia. Available at: `http://en.wikipedia.org/wiki/Envelope_(mathematics)`.
[19] Deltoid curve (mathematics) - Wikipedia, the free encyclopedia. Available at: `http://en.wikipedia.org/wiki/Deltoid_curve`.
[20] Yates, R. C. (1947). A handbook on curves and their properties. J. W. Edwards, Ann Arbor, Mich. MR0023075

Summary. We introduce a new curve: the "perimeter-bisecting deltoid of a triangle" is the envelope of all lines that bisect its perimeter. This is a six-sided curve in the shape of the Greek letter delta consisting of three line segments and three segments of parabolas. We describe this curve both as analytic and as geometric locus, compute the area enclosed by it, and classify the points of the triangle according to the number of distinct perimeter-bisecting lines through them.

ALLAN BERELE (MR Author ID: 35020) received his PhD from the University of Chicago and is a professor at DePaul University in Chicago. His main research interest is in algebras with polynomial identities, although he fell in love with Euclidean geometry many years before he ever heard of p. i. algebras.

STEFAN CATOIU (MR Author ID: 632038) received his PhD from the University of Wisconsin-Madison and is an associate professor at DePaul University in Chicago. His research interest includes non-commutative algebra, real analysis, geometry, number theory, and elementary mathematics.

The Reciprocal Fibonacci Function

MICHAEL LORD
Intera Inc.
Boulder, CO 80301
MLord@Intera.com

In the late 1960s, NASA Langley was interested in solving the communication problem between mission control and spacecraft during reentry. This required designing an antenna which could communicate a radio signal through a plasma of ionized air . The mathematical model for the antenna's conductance was expressed as a series of Bessel functions. For this problem, an efficient summation algorithm uses a recurrence relation and an identity involving Bessel functions to speed up the numerical calculations [**1**, p. 385]. The analysis in this paper has a similar goal of accelerated convergence, but for a different series: the power series whose coefficients are the reciprocal of the Fibonacci numbers, which we'll call the reciprocal Fibonacci function.

Using properties of the reciprocal Fibonacci function, two other series representations for the reciprocal Fibonacci function are found. These representations converge more quickly than the original series. As in the treatment of the series used in NASA's communication problem, the new series representations demand less computational effort than the original series. A numerical comparison of the rates of convergence for each series representation quantifies the gains in computational effort.

Reciprocal Fibonacci function and properties

The Fibonacci sequence $\{f_n\}$ is defined by the recursion equation $f_{n+2} = f_{n+1} + f_n$, for $n = 1, 2, 3, \ldots$ and starting values $f_1 = f_2 = 1$. A known result needed for the subsequent analysis is the Euler–Binet formula, which expresses the Fibonacci numbers in closed form; specifically,

$$f_n = \frac{\phi^n - \hat{\phi}^n}{\sqrt{5}} \text{ where } \phi = \frac{1+\sqrt{5}}{2} \text{ and } \hat{\phi} = \frac{1-\sqrt{5}}{2}.$$

The number ϕ is referred to as the Golden Mean. It follows that $\phi + \hat{\phi} = 1$, $\phi - \hat{\phi} = \sqrt{5}$, and $\phi\hat{\phi} = -1$. The limiting value for the ratio of successive Fibonacci numbers, $\lim_{n\to\infty} f_{n+1}/f_n = \phi$, is attributed to Johannes Kepler.

The power series for which the nth coefficient is the nth Fibonacci number is considered in [**2**]. In what follows, we'll consider the power series in which the nth coefficient is the reciprocal of f_n. This power series evaluated at $x = 1$ yields the series considered in [**3**]. In this sense, the power series generalizes the series of reciprocal Fibonacci numbers [**3**].

Definition. The reciprocal Fibonacci function is

$$F(x) = \sum_{n=1}^{\infty} \frac{x^n}{f_n} = \sqrt{5} \sum_{n=1}^{\infty} \frac{x^n}{\phi^n - \hat{\phi}^n}. \tag{1}$$

Since $\lim_{n\to\infty}(f_n/f_{n+1})|x| = |x|/\phi$, the ratio test implies the series converges for $|x| < \phi$, in which case the radius of convergence is ϕ. Subsequent analysis provides two

Math. Mag. **91** (2018) 134–137. doi:10.1080/0025570X.2018.1426946
MSC: Primary 40A30

other series representations for the reciprocal Fibonacci function. After establishing alternate series representations, calculations determine the convergence rate for each series representation. The resulting convergence rates determine each series numerical effectiveness for the calculation of the reciprocal Fibonacci function.

The nth term of the series for the reciprocal Fibonacci function approaches the terms of a geometric series for n large. This suggests the comparison of $F(x)$ with the geometric series

$$\sqrt{5}\sum_{n=1}^{\infty}\left(\frac{x}{\phi}\right)^n = \sqrt{5}\frac{x/\phi}{1-x/\phi}.$$

This yields

$$\begin{aligned} F(x) - \sqrt{5}\frac{x/\phi}{1-x/\phi} &= F(x) - \sqrt{5}\sum_{n=1}^{\infty}\left(\frac{x}{\phi}\right)^n = \sqrt{5}\left[\sum_{n=1}^{\infty}\frac{x^n}{\phi^n-\hat{\phi}^n} - \sum_{n=1}^{\infty}\left(\frac{x}{\phi}\right)^n\right] \\ &= \sqrt{5}\sum_{n=1}^{\infty}\left(\frac{1}{\phi^n-\hat{\phi}^n} - \frac{1}{\phi^n}\right)x^n = \sqrt{5}\sum_{n=1}^{\infty}\frac{\left((\hat{\phi}/\phi)x\right)^n}{\phi^n-\hat{\phi}^n} = F\left(\frac{\hat{\phi}}{\phi}x\right). \end{aligned}$$

Because the radius of convergence of the two series is ϕ, one can use term-wise subtraction in the above calculation. The result is the following theorem.

Theorem. *The reciprocal Fibonacci function can be represented by*

$$F(x) = \sqrt{5}\frac{x}{\phi - x} + F\left(\frac{\hat{\phi}}{\phi}x\right), \quad \text{for } -\phi < x < \phi. \tag{2}$$

Note that $\hat{\phi} < (\hat{\phi}/\phi)x < -\hat{\phi}$. It follows that equation (2) represents the reciprocal Fibonacci function as a geometric series in closed form plus the value of the function with argument scaled by $\hat{\phi}/\phi$. This representation determines the behavior of the reciprocal Fibonacci function at the endpoints of the interval of convergence, ϕ and $-\phi$. As x approaches ϕ from the left, the function approaches $+\infty$ and is asymptotic to the function $\sqrt{5}[x/(\phi - x)]$. As x approaches $-\phi$ from the right, the limit exists and the Fibonacci function equals $F(-\phi) = -\sqrt{5}/2 + F(-\hat{\phi})$.

Equation (2) provides a convenient method for the numerical evaluation of the Fibonacci function since the series expansion for $F((\hat{\phi}/\phi)x)$ converges faster than the series expansion given by equation (1). The numerical advantage provided by equation (2) improves as x approaches ϕ. More efficient computational forms than equation (2) are considered next.

Alternate series representations for the reciprocal Fibonacci function

The result given by equation (2) can be iterated using the ratio $\beta = \hat{\phi}/\phi = (-3 + \sqrt{5})/2 \approx -0.3820$ to give

$$F(x) = \sqrt{5}\sum_{n=0}^{\infty}\frac{x\beta^n}{\phi - x\beta^n} \text{ for } -\phi \le x < \phi, \tag{3}$$

because for continuous F, it follows that $\lim_{n\to\infty} F(\beta^n x) = F(\lim_{n\to\infty}\beta^n x) = F(0) = 0$. The series for the reciprocal Fibonacci function in equation (3) can be expressed as

$$F(x) = \sqrt{5}\left[\frac{\hat{x}}{1-\hat{x}} + \sum_{n=1}^{\infty}\frac{\hat{x}\beta^n}{1-\hat{x}\beta^n}\right], \qquad \text{where } \hat{x} = x/\phi \text{ and } -1 \le \hat{x} < 1. \tag{4}$$

In [3], the summation in equation (4) is identified as a generalized Lambert series. It follows from [3, Eq. 5], a special case of the Rogers-Fine identity [4, relation 14.1], that the representation in equation (4) is identical to the series, for $-1 \le \hat{x} < 1$,

$$F(x) = \sqrt{5}\left[\frac{\hat{x}}{1-\hat{x}} + \sum_{n=1}^{\infty}\frac{1-\hat{x}\beta^{2n}}{(1-\hat{x}\beta^n)(1-\beta^n)}\hat{x}^n\beta^{n^2}\right]. \tag{5}$$

Numerical summation and rate of convergence

For a convergent series $S = \sum_{k=1}^{\infty} a_k$, the sequence of partial sums $\sum_{k=1}^{n} a_k$ has error $e_n = \sum_{k=n+1}^{\infty} a_k$. The series converges linearly if $\lim_{n\to\infty} |e_{n+1}/e_n| = r$, where $0 < r < 1$, and the value of r is called the asymptotic rate of convergence. The sequence of partial sums converges superlinearly if $\lim_{n\to\infty} |e_{n+1}/e_n| = 0$.

In [3], the series representation provided by equation (4) is used to provide a fast computation for the sum of the inverse Fibonacci numbers, $F(1) = \sum_{n=1}^{\infty}(1/f_n) \approx 3.35988566624318$. In what follows, the numerical behavior describing convergence for the three summations of equations (1), (4), and (5) are compared. The numerical convergences for each of the three series appear in Table 1.

TABLE 1: Convergence behavior of the series representations for $F(1)$

Convergence of partial summation for Equation (1), $x = 1$

n	nth Partial sum	Error	$\lvert e_{n+1}/e_n \rvert$
10	3.330469041	2.942E-02	0.6180
20	3.359646489	2.392E-04	0.6180
30	3.359883722	1.945E-06	0.6180
40	3.359885650	1.581E-08	0.6180
47	3.359885666	5.446E-10	

Convergence of partial summation for Equation (4), $x = 1$

n	nth Partial sum	Error	$\lvert e_{n+1}/e_n \rvert$
5	3.356770386	3.115E-03	0.3863
10	3.359910916	2.525E-05	0.3819
15	3.359885461	2.053E-07	0.3820
20	3.359885668	1.669E-09	0.3820
21	3.359885666	6.376E-10	

Convergence of partial summation for Equation (5), $x = 1$

n	nth Partial sum	Error	$\lvert e_{n+1}/e_n \rvert$
1	3.336881039	2.300E-02	3.626E-03
2	3.359969090	8.342E-05	8.309E-04
3	3.359885597	6.932E-08	1.020E-04
4	3.359885666	7.067E-12	

Numerical results in Table 1 for equation (1) show $\lim_{n\to\infty} |e_{n+1}/e_n| = 1/\phi < 1$, so that the series in equation (1) converges linearly with asymptotic rate $\phi^{-1} \approx 0.6180$ and converges to 10-digit accuracy for $n = 47$.

For the series in equation (4), $\lim_{n\to\infty} |e_{n+1}/e_n| = |\beta| < 1$. The series representation in equation (4) converges linearly with asymptotic rate $|\beta| \approx 0.3820$ and converges to 10-digit accuracy for $n = 21$.

For the series in equation (5), $\lim_{n\to\infty} |e_{n+1}/e_n| = 0$. Hence, the series representation of equation (5) possesses superlinear convergence and converges with 10-digit accuracy for $n = 4$.

The numerical results in Table 1 were calculated for $x = 1$. Note that for $-\phi < x < \phi$, the asymptotic convergence rate for equation (1) is $|x|/\phi$. As x approaches ϕ from the left, the asymptotic convergent rate approaches 1, which implies that the convergence rate slows as x approaches 1. The asymptotic convergence rate for the series in equations (4) and (5) does not depend on x. The disparity between the rate of convergence of the series in equation (1) and either series from equations (4) or (5) becomes greater as x approaches ϕ. In summary, the series representation in equation (4) is numerically superior to the series in equation (1). The most powerful computationally is equation (5).

The analysis for the summation of the inverse Fibonacci series demonstrates that the goal of the numerical analyst is to do the analysis first and then do the computation. In this there is much to be gained. Suggested for further study is the problem: for any power series $\sum a_n x^n$, determine conditions on the coefficients for which the method described in this paper provides accelerated convergence.

As a historical note to the reader, NASA never developed an antenna that can communicate during space craft reentry. Communication with astronauts during reentry is still an unsolved problem.

Acknowledgement The author thanks Shelly Gessing of Foothills Elementary, Boulder, CO for the opportunity to discuss with her students the contributions of Fibonacci. The author's presentation to the 5th grade mathematicians was the catalyst for this paper.

REFERENCES

[1] Abramowitz, M., Stegun, A., eds. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Mineola, NY: Dover Publications, Inc., 1965.

[2] Glaister, P. (1995). Fibonacci power series. *Math. Gazette* 79(486): 521–525.

[3] Arndt, J. (2012). On computing the generalized Lambert series. arXiv: 1202.6525c3

[4] Fine, N.J. (1988). *Basic Hypergeometric Series and Applications, Mathematical Surveys and Monographs*, No. 27. Providence, RI: American Mathematical Society.

Summary. Properties of the power series whose coefficients are the reciprocals of the Fibonacci numbers are derived. It is shown that this function is equivalent to a generalized Lambert series. An application of the Rodgers–Fine identity provides a fast convergent series representation of the function.

MICHAEL LORD (MR Author ID: 115880) has Baccalaureate and Master's degrees in engineering with gas dynamics experience in the aerospace industry. Following a PhD in mathematics, his research in academia included stability analysis of nonlinear dynamic systems. His tenure in the petroleum industry focused on development of numerical models for simulation of petroleum production. He contributed to the numerical modeling of ground water flow and contaminant transport in support of the risk assessment for the licensing of the Waste Isolation Pilot Plant (WIPP) and the high level nuclear waste repository at Yucca Mountain. Since retirement his efforts have included analysis for the remediation of nuclear waste materials at the Hanford production site.

A Note on the AM-GM Inequality

JÓZSEF SÁNDOR
Universitatea Babes-Bolyai Cluj-Napoca
Romania
jsandor@math.ubbcluj.ro

In THIS MAGAZINE, Freedman [1] used Lagrange multipliers to prove a generalization of arithmetic mean-geometric mean (AM-GM) inequality. In what follows, we point out that Freedman's result is equivalent to an earlier result from [2].

Theorem (Freedman [1]). *Let* $r_j, p_j, q_j > 0$ *be given for* $1 \leq j \leq n$; *for each* j *let* $k_j = \frac{p_j}{q_j}$, *and let* $K = \sum_{j=1}^{n} k_j$. *Then*

$$\frac{1}{K}\sum_{j=1}^{n} r_j^{q_j} \geq \left(\frac{r_1^{p_1} r_2^{p_2} \cdots r_n^{p_n}}{k_1^{k_1} k_2^{k_2} \cdots k_n^{k_n}}\right)^{\frac{1}{K}}, \tag{1}$$

with equality if and only if $\frac{r_1^{q_1}}{k_1} = \frac{r_2^{q_2}}{k_2} = \cdots = \frac{r_n^{q_n}}{k_n}$.

The AM–GM inequality is obtained upon taking $p_j = q_j = 1$ for all j in (1).

In [2], by using the inequality $\sum_{i=1}^{n} \inf_{x \in E} f_i(x) \leq \inf_{x \in E} \sum_{i=1}^{n} f_i(x)$, where $f_i : E \subset \mathbb{R} \to \mathbb{R}$ are arbitrary functions, the following theorem was proved where $E := (0, \infty)$, $f_i(x) = a_i x^{\alpha} - b_i \log x$, $(a_i, b_i > 0)$, $\alpha > 0$. Note that Hölder's, Minkowski's and other classical inequalities are proved in [2] by using this method.

Theorem (Sándor and Szabó [2]). *If* $a_i, b_i > 0$ *are arbitrary real numbers, then*

$$\prod_{i=1}^{n}\left(\frac{b_i}{a_i}\right)^{b_i} \geq \left(\sum_{i=1}^{n} b_i \Big/ \sum_{i=1}^{n} a_i\right)^{\sum_{i=1}^{n} b_i}, \tag{2}$$

with equality if and only if $a_j b_l - a_l b_j = 0$ *for all* j, l.

Let $p_j = k_j q_j$, so that after a simple transformation, (1) becomes

$$\frac{k_1^{k_1} \cdots k_n^{k_n}}{(r_1^{q_1})^{k_1} \cdots (r_n^{q_n})^{k_n}} \geq \left(\frac{k_1 + \cdots + k_n}{r_1^{q_1} + \cdots + r_n^{q_n}}\right)^{k_1 + \cdots + k_n}. \tag{1'}$$

By letting $k_i = b_i$ and $r_i^{q_i} = a_i$ for all i, inequality (1′) is in fact inequality (2). As $a_j b_l = a_l b_j$ only if $r_j^{q_j} k_l = r_l^{q_l} k_j$, the cases of equality in (1′) are established, too.

REFERENCES

[1] Freedman, W. (2012). Minimizing areas and volumes and a generalized AM–GM inequality. *Math. Mag.* 85(2):116–123. doi.org/10.4169/math.mag.85.2.116

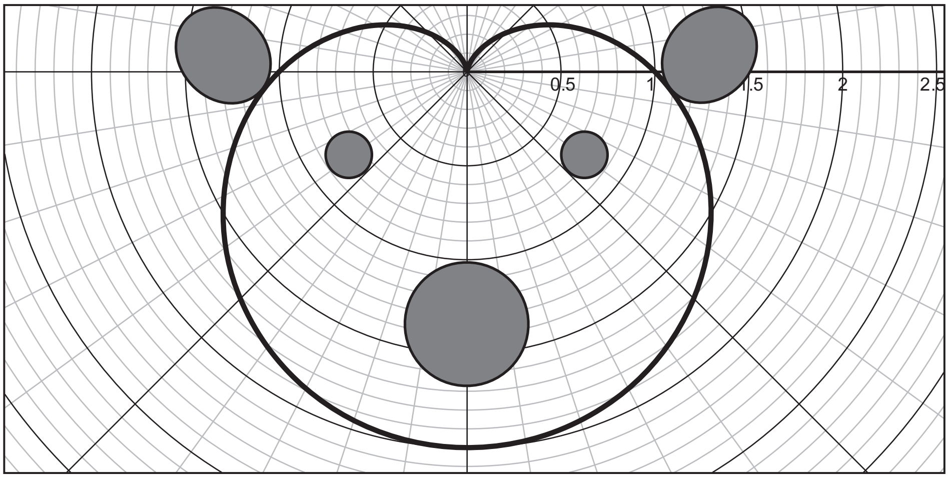
[2] Sándor, J., Szabó, V. E. S. (1996). On an inequality for the sum of infimums of functions. *J. Math. Anal. Appl.* 204(0459):646–654. doi.org/10.1006/jmaa.1996.0459

Math. Mag. **91** (2018) 138–139. doi:10.1080/0025570X.2017.1409533
MSC: Primary 26D15

Summary. We show that, a generalization of the AM-GM inequality, proved by Lagrange multipliers, can be proved by an elementary method.

JÓZSEF SÁNDOR (MR Author ID: 217908) is an associate professor at Babes-Bolyai University, Romania. He received his Ph.D. in 1998, in the theme of generalized convexity and their applications. He is the editor or associate editor of 20 international journals. His research is primarily in number theory, mathematical analysis, and classical geometry.

Polar Bear



$$r = 1 - \sin\theta$$

— Submitted by Robert Haas
Cleveland Heights, OH
rhaas3141@yahoo.com

Math. Mag. **91** (2018) 139. doi:10.1080/0025570X.2017.1415574

Even Perfect Numbers End in 6 or 28

ROGER B. NELSEN
Lewis & Clark College
Portland, OR
nelsen@lclark.edu

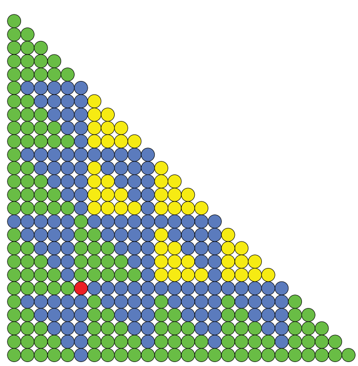
Perfect numbers—positive integers equal to the sum of their proper divisors, such as $6 = 1 + 2 + 3$ and $28 = 1 + 2 + 4 + 7 + 14$—are a staple of every elementary number theory course. Most textbooks for such a course include examples or exercises asking students to show that (a) every even perfect number is a triangular number, and (b) every even perfect number ends in 6 or 28. In this short note, we illustrate how one can employ (a) and simple congruences to prove (b).

Euclid proved that if p and $q = 2^{p-1}$ are prime, then $N_p = 2^{p-1}q$ is perfect, and Euler showed that every even perfect number must have this form. If $T_n = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ denotes the nth triangular number, then simple algebra establishes $N_p = T_{2^p-1}$. For a visual proof of this fact, see [1].

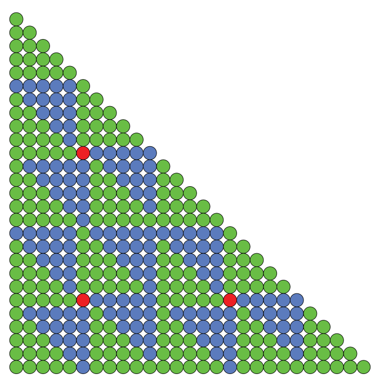
We begin with congruence results for T_{5n+1} and T_{5n+2}.

Lemma. $T_{5n+1} \equiv 1 \pmod 5$ and $T_{5n+2} \equiv 3 \pmod{25}$.

Proof. Below we partition T_{5n+1} and T_{5n+2} to show that $T_{5n+1} = 20T_n + 5T_{n-1} + 1$ and $T_{5n+2} = 25T_n + 3$.



$T_{5n+1} = 20T_n + 5T_{n-1} + 1$



$T_{5n+2} = 25T_n + 3$

■

We now proceed to our main result.

Theorem. *Every even perfect number ends in 6 or 28.*

Proof. Since $N_2 = 6$, we need only consider N_p for p odd. When p is odd, there are two cases: $p = 4k + 1$ and $p = 4k + 3$. If $p = 4k + 1$ then $2^p - 1 = 2 \cdot 16^k - 1 \equiv 1 \pmod 5$, i.e., $2^{p-1} = 5n + 1$ for some positive integer n. Hence, $N_p = T_{2^p-1} = T_{5n+1} \equiv 1 \pmod 5$, so that in base 10, N_p ends in 1 or 6. Since N_p is even, it ends in 6. If $p = 4k + 3$ then $2p - 1 = 8 \cdot 16k - 1 \equiv 2 \pmod 5$, i.e., $2p - 1 = 5n + 2$ for some positive integer n. Hence, $N_p = T_{2^p-1} = T_{5n+2} \equiv 3 \pmod{25}$, so that in base 10, N_p ends in 03, 28, 53, or 78. Since N_p is a multiple of 4 for $p \geq 3$, it ends in 28. ■

Math. Mag. **91** (2018) 140–141. doi:10.1080/0025570X.2017.1409558
MSC: Primary 11A07
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

REFERENCE

[1] Nelsen, R. B. (2016). Proof without words: Perfect numbers and triangular numbers. *College Math. J.* 47: 171. doi.org/10.4169/college.math.j.47.3.171.

Summary. We partition triangular numbers to show that even perfect numbers end in 6 or 28.

ROGER NELSEN (MR Author ID: 237909) is a professor emeritus at Lewis & Clark College, where he taught mathematics and statistics for 40 years.

Solution to the Partiti Puzzle

13 1237	9 45	16 79	11 56	14 347	11 128
15 69	8 8	3 3	3 12	9 9	11 56
10 235	8 17	6 6	11 47	8 8	13 1237
12 48	9 9	5 23	6 15	9 9	15 456
7 16	7 7	8 8	4 4	5 23	16 178
20 389	11 245	13 139	18 567	9 9	15 456

Math. Mag. **91** (2018) 141. doi:10.1080/0025570X.2017.1421372

Counting Interior Roots of Trinomials

MICHAEL A. BRILLESLYPER
LISBETH E. SCHAUBROECK
United States Air Force Academy
CO 80840
mike.brilleslyper@usafa.edu
beth.schaubroeck@usafa.edu

The history of locating roots of polynomials is a vast and interesting field, spanning algebraic, geometric, topological, and numerical techniques. It dates back to ancient times and has received considerable attention in every era of mathematics. It is common to specialize to specific classes of polynomials, often by restricting the coefficients to certain subsets of the integers. There are far too many references to list, but we refer the reader to the detailed paper on the geometry of trinomial zeros by Melman [4] and the references therein as a starting point. In this paper, we focus on a particularly simple class of trinomials and we investigate how many roots occur inside, outside, or on the unit circle. We begin with an example to illustrate the particulars of the problem.

Example. Let $p(z) = z^{10} + z^7 - 1$. The roots $z = x + iy$ of $p(z)$ are plotted in the complex plane in Figure 1. Five roots are inside the unit circle (*interior roots*) and five roots are outside the unit circle (*exterior roots*). For this trinomial, no roots occur on the unit circle (*unimodular roots*).

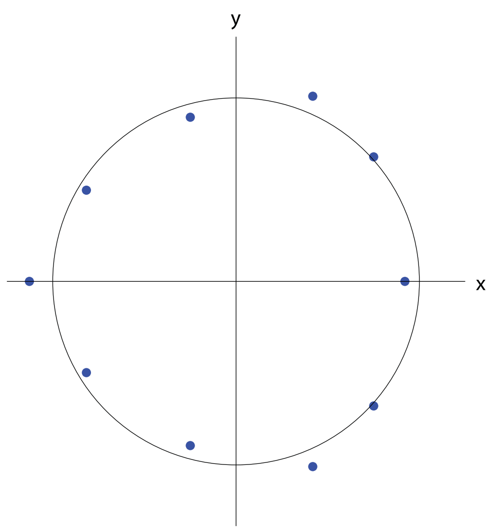


Figure 1 The distribution of roots of $p(z) = z^{10} + z^7 - 1$ relative to the unit circle.

Generalizing the above example, we explore the family of trinomials given by $p(z) = z^n + z^k - 1$, where n and k are positive integers with $n \geq 2$ and $1 \leq k \leq n-1$. In an earlier paper [1], we explored this family of trinomials and settled the question

Math. Mag. **91** (2018) 142–150. doi:10.1080/0025570X.2017.1420332.
MSC: Primary 30C15
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

of when they have unimodular roots. We restate the result here since we will refer to it later in the paper.

Theorem 1 (Brilleslyper and Schaubroeck [1]). *Let n and k be positive integers with $n \geq 2$, $1 \leq k \leq n-1$, and let $g = \gcd(n, k)$. Then the polynomial $p(z) = z^n + z^k - 1$ has unimodular roots if and only if $6g$ divides $n + k$. Moreover, when the condition holds, there are $2g$ unimodular roots given by z_m and $\bar{z}_m$, where*

$$z_m = \exp\left(i\left(\frac{\pi}{3g} + \frac{2\pi m}{g}\right)\right), \quad 0 \leq m \leq g-1.$$

Note that z_m and $\bar{z}_m$ are the gth roots of $e^{i\pi/3}$ and $e^{-i\pi/3}$, respectively.

In the course of establishing Theorem 1, we observed the number of interior and exterior roots did not seem to follow a predictable pattern. We left it as an open question to derive formulas that count these values as functions of the exponents, n and k. Here, we conjecture a formula for the number of interior roots and provide a proof that the formula is correct in the case $k = 1$. A corresponding formula for the number of exterior roots follows easily.

Counting interior roots

Consider again the introductory example, $p(z) = z^{10} + z^7 - 1$. By Theorem 1, we know that $p(z)$ does not have unimodular roots because 6 does not divide $10 + 7$. Thus, as shown in Figure 1, all 10 roots are interior or exterior, with five of each type. But change the polynomial to $z^{10} + z^9 - 1$, and the numbers change to seven interior roots and three exterior roots. Is there a discernible pattern here? Table 1 shows the complexity of the situation for the complete family of degree-10 trinomials, $p(z) = z^{10} + z^k - 1$, where $1 \leq k \leq 9$.

The pattern in Table 1 is difficult to analyze. This is partly because n and k are not always relatively prime. For a specific polynomial in which $\gcd(n, k) = g > 1$, there is a simple relationship between the roots of $p(z) = z^n + z^k - 1$ and the roots of the lower degree *reduced* polynomial $\widetilde{p}(z) = z^{n/g} + z^{k/g} - 1$. As explained in Lemma 2 in [1], the roots of $p(z)$ are in g-to-1 correspondence with the roots of $\widetilde{p}(z)$. Moreover, the correspondence preserves the division into interior, exterior, and unimodular categories. As an example, the polynomial $p(z) = z^{10} + z^4 - 1$ has six interior roots. Because $\gcd(10, 4) = 2$, the reduced polynomial $\widetilde{p}(z) = z^5 + z^2 - 1$ has three interior roots. Thus it suffices to study only trinomials in which $\gcd(n, k) = 1$.

TABLE 1: The number of interior, exterior, and unimodular roots for the family $p(z) = z^{10} + z^k - 1$

k	Interior roots	Exterior roots	Unimodular roots
1	3	7	0
2	2	4	4
3	5	5	0
4	6	4	0
5	5	5	0
6	6	4	0
7	5	5	0
8	6	4	0
9	7	3	0

TABLE 2: The number of interior, exterior, and unimodular roots for the family $p(z) = z^{11} + z^k - 1$

k	Interior roots	Exterior roots	Unimodular roots
1	3	6	2
2	5	6	0
3	5	6	0
4	5	6	0
5	5	6	0
6	5	6	0
7	5	4	2
8	7	4	0
9	7	4	0
10	7	4	0

To ensure $\gcd(n, k) = 1$ for all k, we consider a family where n is prime. Table 2 shows the root distribution for the family of trinomials when $n = 11$.

The pattern in Table 2 is more easily analyzed. We observe the values in the second and third columns of the table are constant as k increases until we reach a row where 6 divides $n + k$. In each such row, there are two unimodular roots and the number of exterior roots decreases by two. In the next row, there are no unimodular roots, the number of interior roots increases by two, and the number of interior and exterior roots remain constant for five rows. It is possible to capture this pattern with a relatively simple formula, which we state in the following conjecture. We have confirmed through numerical calculations that our conjectured formula agrees with the actual root distribution for many combinations of n and k. Here, we will prove that the formula gives the number of interior roots for the trinomials when $k = 1$.

Conjecture. *Let $p(z) = z^n + z^k - 1$, with $n, k \in \mathbb{N}$, and $1 \leq k \leq n - 1$. Let $\sigma(n, k)$ denote the number of interior roots. If $\gcd(n, k) = 1$, then*

$$\sigma(n, k) = 2\left\lfloor \frac{n + k - 1}{6} \right\rfloor + 1,$$

where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x.

In the cases where $\gcd(n, k) = g > 1$, we extend the above formula to

$$\sigma(n, k) = 2g\left\lfloor \frac{n + k - g}{6g} \right\rfloor + g.$$

This extension follows from the previous discussion concerning the relationship between the roots of the polynomial $p(z)$ and the reduced polynomial $\widetilde{p}(z)$.

Root localization

Plotting the set of roots of $p(z) = z^n + z^k - 1$ for any combination of n and k shows that they tend to be distributed around the unit circle with one root near each of the nth roots of unity (the roots of $z^n - 1$). This observation led us to seek a set of disjoint angular sectors, each containing exactly one root of $p(z)$. Studying the angular separation of roots is common in the research area of *root localization* and was extensively used in the study of trinomials by Melman [4]. Here, we use the classical result due to Rouché to accomplish the task. For completeness, we state Rouché's theorem for analytic functions and refer the reader to any standard text in complex analysis for a proof and further applications [3].

Rouché's Theorem. *Let C be a simple closed contour in $\mathbb{C}$ and f and g be functions that are analytic in a simply connected domain that contains C. If $|f(z) - g(z)| < |g(z)|$ for all $z \in C$, then f and g have the same number of roots, counted with multiplicity, inside C.*

Rouché's theorem applied to $p(z) = z^n + z^k - 1$ We begin by defining a set of n disjoint angular sectors,

$$R_a = \left\{ \rho e^{i\theta} \ : \ 0 < \rho < 2 \text{ and } \frac{2a\pi}{n} - \frac{\pi}{2n} < \theta < \frac{2a\pi}{n} + \frac{\pi}{2n} \right\},$$

where $a = 0, 1, \ldots, n-1$. The sectors R_a are centered on the nth roots of unity. We refer to these sectors as *Rouché sectors* and prove the following lemma.

Lemma 1. *Each Rouché sector contains exactly one root of $p(z)$.*

Proof. We apply Rouché's theorem letting $f(z) = p(z)$ and $g(z) = z^n - 1$. Then $|f - g| = |z^k|$. Observe the roots of $g(z)$ are the nth roots of unity, and they lie on the unit circle in the center of each angular sector. We consider the rays ρe^{it}, where $0 < \rho < 2$ and $t = \frac{2a\pi}{n} \pm \frac{\pi}{2n}$. Along these rays, $|z^k| = \rho^k$, and

$$\begin{aligned} |z^n - 1|^2 &= |\rho^n e^{itn} - 1|^2 = (\rho^n e^{itn} - 1)(\rho^n e^{-itn} - 1) \\ &= \rho^{2n} + 1 - \rho^n(e^{itn} + e^{-itn}) = \rho^{2n} + 1 - 2\rho^n \cos(tn). \end{aligned}$$

Since $\cos(tn) = 0$, we have $|z^n - 1| = \sqrt{\rho^{2n} + 1}$. Now if $\rho < 1$, then $|z^k| = \rho^k < 1$, and if $\rho = 1$, then $\rho^k = 1$. Also, if $\rho > 1$, then $\rho^k < \rho^n$ since $k < n$. In all cases we have that $|z^k| < \sqrt{\rho^{2n} + 1} = |z^n - 1|$ which satisfies the hypotheses of Rouché's theorem. Furthermore, along the circular arc $\rho = 2$, $|z^n - 1| \geq 2^n - 1 > 2^k = |z^k|$. Therefore the conditions of Rouché's theorem are satisfied along the boundary of R_a, and since $z^n - 1$ has only one root within R_a, so does $p(z)$. ■

To illustrate, Figure 2 shows the Rouché sectors and the six roots for the trinomial $p(z) = z^6 + z - 1$. Note that each root occurs in exactly one Rouché sector, along with a 6th root of unity.

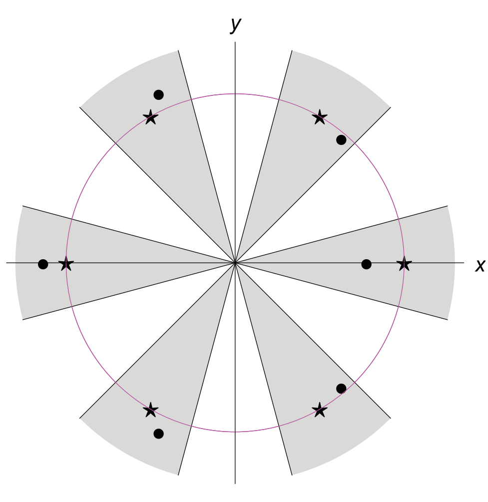


Figure 2 The Rouché sectors, roots of $p(z) = z^6 + z - 1$, and the 6th roots of unity, shown as stars.

The special case $k = 1$

Assuming $k = 1$, our conjecture for the number of interior roots reduces to an expression that depends only on n. Specifically,

$$\sigma(n, 1) = 2\left\lfloor \frac{n}{6} \right\rfloor + 1.$$

Proving this formula is correct requires introducing a related trinomial, whose roots are in one-to-one correspondence with those of $p(z) = z^n + z^k - 1$ via reflection across the unit circle. We define

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

We then observe that $q(z) = z^n - z^{n-k} - 1$, and $p(z_0) = 0$ if and only if $q(1/\bar{z}_0) = 0$. The polynomial $q(z)$ is the negation of what is commonly called the *reverse polynomial* of $p(z)$. For what follows, we note that for any complex number $z \neq 0$, both z and $1/\bar{z}$ have the same argument.

Disjoint angular sectors Our first observation uses properties of the polynomials $p(z) = z^n + z - 1$ and $q(z) = z^n - z^{n-1} - 1$ along with basic geometry in the complex plane to establish that all interior roots of $p(z)$ must lie in the interior of the sector with angular boundaries $-\pi/3$ and $\pi/3$, while all the exterior roots lie in the complementary sector with angular boundaries $\pi/3$ and $5\pi/3$. We recall by Theorem 1 that unimodular roots can only occur on the boundaries of these angular sectors.

To establish this angular separation, let $z_0 = r_0 e^{i\theta_0}$. Assume $p(z_0) = 0$ and $r_0 < 1$. Let $w_0 = \frac{1}{\bar{z}_0}$. Observe that $|w_0| > 1$, $\text{Arg}(w_0) = \text{Arg}(z_0)$, and $q(w_0) = 0 = w_0^n - w_0^{n-1} - 1$, so that $w_0^{n-1}(w_0 - 1) = 1$. Taking the modulus of both sides and observing that $|w_0^{n-1}| > 1$, it follows that $|w_0 - 1| < 1$. Therefore, w_0 lies outside the circle $|z| = 1$ and inside the circle $|z - 1| = 1$. These two circles intersect at the angular values of $\pm\frac{\pi}{3}$. Since $\text{Arg}(w_0) = \text{Arg}(z_0)$, then $-\frac{\pi}{3} < \theta_0 < \frac{\pi}{3}$.

Now assume $r_0 > 1$ and let $w_0 = \frac{1}{\bar{z}_0}$. As before, note that $q(w_0) = 0$, $|w_0| < 1$, and $\text{Arg}(w_0) = \text{Arg}(z_0)$. Since $q(w_0) = 0$, we have $w_0^{n-1}(w_0 - 1) = 1$. Taking the modulus of both sides and observing that $|w_0^{n-1}| < 1$, it follows that $|w_0 - 1| > 1$. Therefore, w_0 lies inside the circle $|z| = 1$ and outside the circle $|z - 1| = 1$. But $z_0 = 1/\overline{w_0}$, so z_0 must lie in the region given by $\frac{\pi}{3} < \theta_0 < \frac{5\pi}{3}$.

The preceding argument establishes the following result.

Lemma 2. *Let z_0 be a root of $p(z) = z^n + z - 1$. If z_0 is an interior root, then $-\pi/3 < \arg(z_0) < \pi/3$. If z_0 is an exterior root, then $\pi/3 < \arg(z_0) < 5\pi/3$.*

A counting strategy To count the number of interior roots of $p(z)$, we count the number of intersections of the Rouché sectors with the angular sector

$$I = \left\{ re^{i\theta} \ : \ r > 0, -\frac{\pi}{3} < \theta < \frac{\pi}{3} \right\}.$$

Our analysis above shows that any Rouché sector that is a subset of I must contain an interior root. Similarly, any Rouché sector disjoint from the sector I must contain an exterior root. Thus, the argument comes down to showing what happens to any Rouché sector that intersects the boundary rays of the sector I.

We first consider which Rouché sectors can intersect the boundary of I. The result depends on the value of n mod 6.

Lemma 3. *The points on the unit circle with arguments $\pm\pi/3$ are contained in a Rouché sector, R_a, if and only if $n \equiv 0, 1,$ or $5 \bmod 6$.*

Proof. Assume there exists a, such that $e^{i\pi/3} \in R_a$. Thus, we have the inequality

$$\frac{2a\pi}{n} - \frac{\pi}{2n} < \frac{\pi}{3} < \frac{2a\pi}{n} + \frac{\pi}{2n}.$$

Multiplying through by $3n/\pi$ gives

$$6a - \frac{3}{2} < n < 6a + \frac{3}{2}.$$

The inequality defines an open interval of length 3, centered on a multiple of 6. Thus, any n that satisfies the inequality must be -1, 0, or 1 when reduced modulo 6. Therefore, this inequality can only hold when $n \equiv 0, 1,$ or $5 \bmod 6$. The same holds for $e^{-i\pi/3}$ because the Rouché sectors are symmetric via reflection with respect to the real axis. ∎

Lemma 3 shows when $n \equiv 2, 3,$ or $4 \bmod 6$, that no Rouché sector contains $e^{\pm i\pi/3}$. In these cases, none of the Rouché sectors overlap the boundary of I. Hence, for these three cases, the number of interior roots of $p(z)$ is simply equal to how many Rouché sectors are subsets of the sector I.

To count how many Rouché sectors are subsets of I, first note that the sector R_0 is symmetric about the positive real axis and we always have $R_0 \subset I$. Furthermore, the other Rouché sectors are symmetric with respect to the real axis. Thus, it suffices to count only those sectors that lie entirely in the upper half plane and intersect I. To do so we count how many positive integer values of a satisfy

$$\frac{2a\pi}{n} - \frac{\pi}{2n} < \frac{\pi}{3}. \tag{1}$$

We then multiply this result by 2 to account for the symmetry and add 1 to account for R_0. Solving equation (1) for a gives

$$a < \frac{n}{6} + \frac{1}{4}. \tag{2}$$

The right side of equation (2) is never an integer since it equals $(2n+3)/12$, which has an odd numerator and an even denominator. Hence, the number of integers a that satisfy equation (2) is given by

$$\left\lfloor \frac{n}{6} + \frac{1}{4} \right\rfloor. \tag{3}$$

Expression (3) is equal to $\lfloor n/6 \rfloor$ when $n \not\equiv 5 \bmod 6$. Multiplying by 2 and adding 1 confirms our formula for $\sigma(n, 1)$ when $n \equiv 2, 3,$ or $4 \bmod 6$.

The case $n \equiv 5 \bmod 6$ When $n \equiv 5 \bmod 6$, expression (3) becomes $\lfloor n/6 \rfloor + 1$. However, in this case $n + 1$ is divisible by 6 and Theorem 1 applies to $p(z)$. Hence, there are two unimodular roots at angles of $\pm\pi/3$. Recalling that each Rouché sector contains exactly one root of $p(z)$, we can subtract 2 to account for these two intervals. Thus, we obtain $\sigma(n, 1) = 2(\lfloor n/6 \rfloor + 1) + 1 - 2 = 2\lfloor n/6 \rfloor + 1$.

The case $n \equiv 0 \bmod 6$ When $n \equiv 0 \bmod 6$ or $n \equiv 1 \bmod 6$, the angles $\pm\pi/3$ are always contained in some Rouché sector. However, in these cases we will show that no exterior root can be contained in those particular Rouché sectors. We show the detailed analysis for $n \equiv 0 \bmod 6$, and leave the similar case of $n \equiv 1 \bmod 6$ to the reader.

Lemma 4. *If* $n \equiv 0 \bmod 6$ *then no exterior root of* $p(z) = z^n + z - 1$ *is in the Rouché sectors that contain the angles* $\pm\pi/3$.

Proof. We first demonstrate that if there is an exterior root of $p(z)$, then the root must lie within one of the sectors

$$E_j = \left\{ re^{i\theta} : \frac{2\pi}{3(n-1)} + \frac{2\pi j}{n-1} < \theta < \frac{4\pi}{3(n-1)} + \frac{2\pi j}{n-1} \right\},$$

where $j = 0, 1, \ldots, n-2$. To that end, assume $p(z_0) = 0$ with $|z_0| > 1$. Since $z_0^n + z_0 = 1$, we observe that $|z_0||z_0^{n-1} + 1| = 1$. Thus $|z_0^{n-1} + 1| < 1$. Coupled with the observation that $|z_0^{n-1}| > 1$, we conclude that z_0^{n-1} lies outside of the unit circle and inside of the circle of radius 1 centered at -1. Hence, the argument of z_0^{n-1} lies within the interval $(\frac{2\pi}{3}, \frac{4\pi}{3})$. Taking the $(n-1)$st roots of the interval, we conclude that z_0 lies within E_j.

For simplicity of language, we refer to $\exp[i(\frac{2\pi}{3(n-1)} + \frac{2\pi j}{n-1})]$ and $\exp[i(\frac{4\pi}{3(n-1)} + \frac{2\pi j}{n-1})]$ as the left and right endpoints, respectively, of E_j.

When $n \equiv 0 \bmod 6$, we note that $e^{i\pi/3}$ is contained in R_a precisely when $a = n/6$. We show that none of the E_j can intersect $R_{n/6}$.

Now if the left endpoint of E_j is in $R_{n/6}$, then we have the inequality

$$\frac{2a\pi}{n} - \frac{\pi}{2n} < \frac{2\pi}{3(n-1)} + \frac{2\pi j}{n-1} < \frac{2a\pi}{n} + \frac{\pi}{2n}.$$

Multiplying both sides by $\frac{6(n-1)}{\pi}$ and isolating the j terms in the middle yields

$$\frac{12a(n-1) - 7n + 3}{n} < 12j < \frac{12a(n-1) - n - 3}{n}.$$

Replacing n with $6a$ gives

$$12a - 9 + \frac{1}{2a} < 12j < 12a - 3 - \frac{1}{2a}.$$

Note that the left-hand side is greater than $12a - 9$ while the right-hand side is less than $12a - 3$. Such a j would be expressed as a multiple of 12 between $12a - 9$ and $12a - 3$, which is impossible.

Now we consider if the right endpoint of some E_j could be contained in $R_{n/6}$. In particular, because the argument of an exterior root must be between $\pi/3$ and $5\pi/3$, we only need to consider if it is possible for

$$\frac{\pi}{3} < \frac{4\pi}{3(n-1)} + \frac{2\pi j}{n-1} < \frac{2a\pi}{n} + \frac{\pi}{2n}.$$

Again multiplying through by $\frac{6(n-1)}{\pi}$ and isolating the j terms in the middle yields

$$\frac{2n(n-1) - 8n}{n} < 12j < \frac{12a(n-1) - 5n - 3}{n}.$$

Letting $n = 6a$, we have

$$12a - 10 < 12j < 12a - 7 - \frac{1}{2a}.$$

Again, such an integer j is impossible.

We conclude that $R_{n/6}$ cannot contain any exterior roots, and thus the one root within $R_{n/6}$ is an interior root. ■

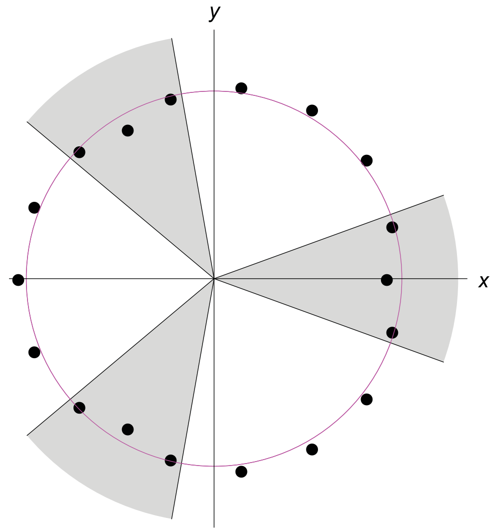


Figure 3 The roots of $p(z) = z^{18} + z^3 - 1$ and the cube roots of the sector I.

The case for $n \equiv 1 \bmod 6$ is algebraically similar and is left as an exercise. Thus, our method of counting interior roots for $n \equiv 2, 3, 4 \bmod 6$ can be extended to $n \equiv 0 \bmod 6$ and $n \equiv 1 \bmod 6$. Therefore, we have proved our conjectured formula for the case $k = 1$.

Theorem 2. *The number of interior roots of the polynomial* $p(z) = z^n + z - 1$ *is given by* $\sigma(n, 1) = 2\lfloor n/6 \rfloor + 1$.

For an example, Theorem 2 tells us that $p(z) = z^{60} + z - 1$ has 21 interior roots. Notice that the interior roots account for roughly one-third of all the roots. It is easily confirmed that $\lim_{n\to\infty} \sigma(n, 1)/n = 1/3$.

Theorem 2 immediately extends to any trinomial whose corresponding reduced polynomial has $k = 1$. Such a trinomial has the form $p(z) = z^{ng} + z^g - 1$, for $g \geq 2$. For example, if we let $n = 6$ and $g = 3$, then we obtain $p(z) = z^{18} + z^3 - 1$. Figure 3 shows all 18 roots of the polynomial along with the regions whose boundary rays correspond to the cube roots of $e^{\pm i\pi/3}$. We see that all the interior roots occur in the gth roots of the sector I. Or, putting it differently, the gth power of any interior root must land in the sector I.

After this paper was submitted for publication, the general conjecture was proved by an undergraduate student, David Kyle, and his advisor Russell Howell in [2] using ideas and techniques developed in an earlier preprint of this paper. We encourage the interested reader to see their work.

REFERENCES

[1] Brilleslyper, M., Schaubroeck, L. (2014). Locating unimodular roots. *Coll. Math. J.* 45(3): 162–168.

[2] Howell, R., Kyle, D. (2018). Locating trinomial zeros. *Involve*. 11(4): 711–720.

[3] Mathews, J. H., Howell, R. W. (2012). *Complex Analysis for Mathematics and Engineering.* 6th ed. Burlington, MA: Jones and Bartlett.

[4] Melman, A. (2012). Geometry of trinomials. *Pac. J. Math.* 259(1): 141–159.

Summary. We consider a simple family of trinomials. The roots of this family, when plotted in the complex plane, form intriguing patterns which motivate some natural questions about their distribution. In particular, we conjecture a formula that counts the number of roots that lie inside the unit circle. We provide a proof of this conjecture for a special case. The proof makes use of basic geometry, number theoretic calculations, and an application of Rouché's theorem. The result partially answers a previous open question from an earlier paper that investigated when this family of trinomials has roots on the unit circle.

MICHAEL A. BRILLESLYPER (MR Author ID: 994411) received his Ph.D. at the University of Arizona in 1994 and is now a professor of mathematical sciences at the United States Air Force Academy. His professional interests focus on problems emerging from undergraduate courses that are suitable for student research projects. Recent topics include roots of polynomials, prime graphs, and combinatorics. He is a co-PI on the StatPREP project and has been very active in the MAA throughout his career. He and his wife have two daughters, one in college and one looking at colleges.

LISBETH E. SCHAUBROECK (MR Author ID: 663310) earned her Ph.D. at the University of North Carolina at Chapel Hill in 1998 and is now a professor of mathematical sciences at the United States Air Force Academy. Her professional interests include complex variables, undergraduate knot theory, mentoring new faculty members, and searching for interesting mathematical coloring book images. She lives with her husband and two sons in Colorado Springs, where they especially enjoy bowling and target archery.

PROBLEMS

Proposals

To be considered for publication, solutions should be received by September 1, 2018.

2041. *Proposed by Vadim Mitrofanov, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine.*

Let $ABCD$ be a quadrilateral that circumscribes a circle of radius r and is also inscribed in a circle of radius R. Let s be the semiperimeter of $ABCD$. Prove the inequality $s^2 \leq 6R^2 + 4r^2$.

2042. *Proposed by Rick Mabry and Debbie Shepherd, Louisiana State University Shreveport, Shreveport, LA.*

Recursively define random variables $X_0, X_1, \ldots, X_n, \ldots$ and $Y_0, Y_1, \ldots, Y_n, \ldots$ taking values in $[0, 1]$ as follows:

- $X_0 = 0$ and $Y_0 = 1$ are constants;
- for $n = 0, 1, 2, \ldots$, X_{n+1} and Y_{n+1} are chosen uniformly and independently in the closed interval with endpoints X_n, Y_n.

Prove that, with probability 1, the limits $\widetilde{X} = \lim_{n\to\infty} X_n$ and $\widetilde{Y} = \lim_{n\to\infty} Y_n$ both exist and are equal, and find their common distribution.

2043. *Proposed by Greg Oman, University of Colorado, Colorado Springs and Adam Salminen, University of Evansville, Evansville, IN.*

Find all commutative rings R with unity such that:

(i) R contains some element x that is neither nilpotent nor a unit (i.e., $x^n \neq 0$ for all $n \geq 1$ and $xy \neq 1$ for all $y \in R$), and
(ii) every proper nonzero ideal of R is maximal.

Math. Mag. **91** (2018) 151–158. doi:10.1080/0025570X.2018.1429759

We invite readers to submit original problems appealing to students and teachers of advanced undergraduate mathematics. Proposals must always be accompanied by a solution and any relevant 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.

Proposals and solutions should be written in a style appropriate for this MAGAZINE.

Authors of proposals and solutions should send their contributions using the Magazine's submissions system hosted at `http://mathematicsmagazine.submittable.com`. *More detailed instructions are available there. We encourage submissions in PDF format, ideally accompanied by LATEX source. General inquiries to the editors should be sent to* `mathmagproblems@maa.org`.

2044. *Proposed by George Stoica, Saint John, New Brunswick, Canada.*

Find all continuous functions $f : [0, 1] \to [0, \infty)$ such that

$$\lim_{x\to 0^+} e^{1/x} f(x) = 0 \qquad \text{and} \qquad f(x) \le \int_0^x \frac{f(t)}{t^2}dt \quad \text{for all } x \in [0, 1].$$

2045. *Proposed by Kenneth Levasseur and Nicholas Raymond (student), University of Massachusetts Lowell, Lowell, MA.*

Let n be a positive integer. For any base b (a positive integer greater than 1) consider the set $R_{n,b}$ consisting of all 2^n nonnegative integers r whose base-b expansion has (at most) n digits, each either 0 or 1. Given $n > 1$, for what bases b is $R_{n,b}$ a complete system of residues to the modulus 2^n?

Quickies

1079. *Proposed by Howard Morris, Southwest Tennessee Community College, Memphis, TN.*

Given any positive integer n show that there exist nonnegative integers a, b, c, each no larger than $\sqrt{n} + 1$, and such that $n = a^2 + b^2 - c^2$.

1080. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Call a function $f : \mathbb{R} \to \mathbb{R}$ *proto-monotone* if

$$\liminf_{t\to x} f(t) \le f(x) \le \limsup_{t\to x} f(t) \qquad \text{for all } x \in \mathbb{R}.$$

Show that the sum of a continuous function and a proto-monotone function is proto-monotone, but the sum of two proto-monotone functions need not be proto-monotone.

Solutions

A regular octagon with one-third the area of another **April 2017**

2016. *Proposed by Rick Mabry, LSU Shreveport, Shreveport, LA.*

Let $\mathcal{P} = A_1A_2A_3A_4A_5A_6A_7A_8$ be a regular octagon. Let $A_0 = A_8$, $A_9 = A_1$, and $A_{10} = A_2$. For $j = 1, 2, \ldots, 8$, let M_j be the midpoint of $\overline{A_jA_{j+1}}$, and let A'_j be the point of intersection of the segments $\overline{A_jM_{j+2}}$ and $\overline{A_{j-1}M_{j+1}}$. Prove that the inner octagon $\mathcal{P}' = A'_1A'_2A'_3A'_4A'_5A'_6A'_7A'_8$ has one-third the area of $\mathcal{P}$.

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

Consider the primitive root of unity $\omega = e^{i\pi/4}$ in the complex plane. Since ratios of areas are invariant under rigid motions and similarities, without loss of generality we may take the vertices of $\mathcal{P}$ to be the eighth roots of unity on the complex plane, say $A_j = \omega^j$ ($j = 1, \ldots, 8$). The line segment from A_j to M_{j+2} is parameterized by $z(t) = (1 - t)A_j + tM_{j+2}$ ($0 \le t \le 1$), so A'_1 is equal to the common value of both sides of the equation

$$(1 - t)A_0 + tM_2 = (1 - u)A_1 + uM_3 \tag{1}$$

when t, u are real solutions thereof. These solutions can be obtained as follows: Since $t = \bar{t}$ and $u = \bar{u}$ are real, by taking equation (1) together with its complex-conjugate equation we obtain a system of two linear equations in t, u, whose solution is

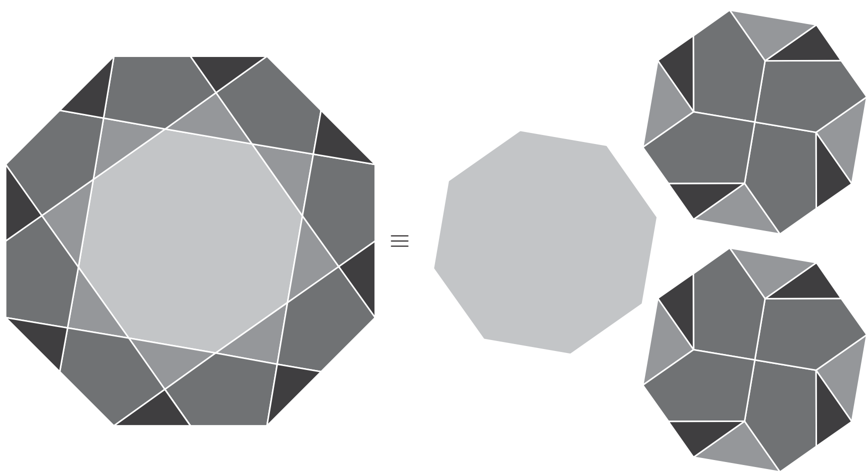
$$t = \frac{\Im[(A_1 - A_0)(\overline{M_3 - A_1})]}{\Im[(M_2 - A_0)(\overline{M_3 - A_1})]}, \quad u = \frac{\Im[(A_1 - A_0)(\overline{M_2 - A_0})]}{\Im[(M_2 - A_0)(\overline{M_3 - A_1})]},$$

where $\overline{z}$ and $\Im(z)$ denote the complex conjugate and the imaginary part of any complex number z, respectively. Straightforward calculation using the values $A_j = \omega^j$, $M_j = (A_j + A_{j+1})/2 = (\omega^j + \omega^{j+1})/2$ and $\omega = (1+i)/\sqrt{2}$ gives the solution $t = 2/3$, and hence

$$A_1' = \frac{1}{3}A_0 + \frac{2}{3}M_2 = \frac{1 + (\omega^2 + \omega^3)}{3} = \frac{(2-\sqrt{2}) + (2+\sqrt{2})i}{6}.$$

From considerations of symmetry, it is clear that $\mathcal{P}'$ is the image of $\mathcal{P}$ under the mapping $z \mapsto A_1' z$, namely, the composition of some rotation with a homothety by factor $|A_1'| = 1/\sqrt{3}$ (both having center 0). It follows that $\mathcal{P}'$ has area equal to $(1/\sqrt{3})^2 = 1/3$ that of $\mathcal{P}$.

A solution without words by Dixon J. Jones, Fairbanks, AK.



Also solved by Michel Bataille, Robert Calcaterra, Callie Drage, Luke Wilson & Jonny Dehaan, Habib Y. Far, Dmitry Fleischman, Michael Goldenberg & Mark Kaplan, Peter McPolin (UK), Jerry Minkus, Missouri State University Problem Solving Group, Mountain Lakes High School Problem Solving Group, Kangrae Park (South Korea), Edward Schmeichel, Achilleas Sinefakopoulos (Greece), John H. Smith, Michael Vowe, and the proposer. There was 1 incomplete or incorrect solution.

A variant of the game of Nim **April 2017**

2017. *Proposed by Sung Soo Kim, Hanyang University, Ansan, Korea.*

Consider the following modification of the classical game of Nim. Initially, there are one or more heaps, each consisting of one or more stones. Players Alice and Bob take turns (Alice plays first). A player's move consists of choosing one or more heaps and removing exactly one stone from each of them. The player who takes the last stone loses. Determine all initial states for which Alice has a winning strategy.

Solution by Michael Reid, University of Central Florida, Orlando, FL.

We claim that the positions for which Alice has a winning strategy are precisely those of the following two types:

(*A1*) there is an even number (zero or more) of stones altogether, all in the same heap;
(*A2*) there are at least two heaps, and at least one heap has an odd number of stones.

Let $\mathcal{A}$ denote the set of positions of these two types, and let $\mathcal{B}$ be its complement, namely, the set of positions of either of the following types:

(*B1*) there is only one heap, and it has an odd number of stones;
(*B2*) there are at least two heaps, all of which have an even number of stones.

Note that the empty position (when no stones are left) is of type A1; it is a winning position since it is reached after the other player removed every stone and lost. Hence, to prove that $\mathcal{A}$ is the set of winning positions for Alice it suffices to show that from every nonempty position P in $\mathcal{A}$, there is at least one move to a position P' in $\mathcal{B}$, and that from each position P in $\mathcal{B}$, every move results in a position P' in $\mathcal{A}$.

Consider any position P in $\mathcal{B}$. If P is of type B1, the only possible move is to remove a single stone from the only heap, leaving an even number of stones all in the same heap, hence leading to a position P' of type A1. If P is of type B2, every stone (at least one) that is removed from any heap will leave the heap with an odd number of stones, and furthermore the number of (nonempty) heaps will not change since each heap has an even number, 2 or more, of stones prior to the move, so it is still nonempty after possibly removing a single stone from it. Therefore, the resulting position P' is of type A2.

Next, let P be any nonempty position in $\mathcal{A}$. If P is of type A1, the only possible move is to take one stone from the unique nonempty heap, which will have an odd number of stones after the move, thus a position P' in $\mathcal{B}$, of type B1. If P is of type A2, the move leading to a position P' in $\mathcal{B}$ must be chosen according to the number of heaps having more than one stone as follows:

Case 1: There are no heaps with more than one stone. Take the unique stone from all heaps except one. By the assumption that P is of type A2, this is a valid move since there are at least two heaps left. Afterward, a single heap having a single stone is left, which is a position P' of type B1.

Case 2: There is exactly one heap (the "large heap") with more than one stone. Remove the stone from each single-stone heap; this results in removal of at least one stone since the large heap cannot be the only heap, because P is of type A2. In case the large heap has an even number of stones, take one stone from this heap also. The resulting position P', of type B1, has a single heap with an odd number of stones.

Case 3: There are at least two heaps with more than one stone. Take a stone from each heap that has an odd number of stones. Since P is of type A2 by assumption, there is at least one such heap. Afterward, all remaining heaps have an even number of stones, and at least two heaps remain, namely, a position P' of type B2.

Remark. Alice's next move to ensure victory is always unique, except in Case 1 for positions of type A2, when the choice of which single-stone heap not to remove is arbitrary. We offer a one-sentence summary of the winning strategy:

"Make all heaps even, unless that would result in a single heap."

Also solved by Pedro Acosta, Robert Calcaterra, Robin Chapman (UK), John Christopher, Joseph DiMuro, Wenwen Du & Paul Peck, Tom Jager, George Wahington University Math Problems Group, Reiner Martin (Germany), Northwestern University Math Problem Solving Group, Paul Stockmeyer, Sarah Swenson & James Swenson, and the proposer. There was 1 incomplete or incorrect solution.

G_δ subsets of $\mathbb{R}$ and derivatives **April 2017**

2018. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

A *derivative function* is a function $f : \mathbb{R} \to \mathbb{R}$ such that f is the derivative of some differentiable function on $\mathbb{R}$.

(i) If f is a derivative function and $c \in \mathbb{R}$, show that the set $\{x \in \mathbb{R} \mid f(x) \geq c\}$ is a countable intersection of open subsets of $\mathbb{R}$.

(ii) Find a derivative function f and $c \in \mathbb{R}$ such that $\{x \in \mathbb{R} \mid f(x) \geq c\}$ is not a closed subset of $\mathbb{R}$.

Solution by Don L. Hancock, Pepperdine University, Malibu, CA.

(i) Let f be a derivative, say $f = g'$. For all $x \in \mathbb{R}$ we have

$$f(x) = \lim_{m\to\infty} h_m(x), \qquad \text{where} \quad h_m(x) = \frac{g\left(x+\frac{1}{m}\right)-g(x)}{\frac{1}{m}}. \tag{1}$$

For fixed real c and any positive integers m, n, let $S = \{x \in \mathbb{R} : f(x) \geq c\}$, $E = \{x \in \mathbb{R} : f(x) < c\}$, and $E_{n,m} = \{x \in \mathbb{R} : h_m(x) \leq c - 1/n\}$. If $x \in E$ then $f(x) < c - 1/n$ for some n; in this case, equation (1) implies that $h_m(x) \leq c - 1/n$ for all sufficiently large m, say for $m \geq k$. It follows that E is a subset of the set

$$F = \bigcup_{n,k=1}^{\infty} \bigcap_{m=k}^{\infty} E_{n,m}.$$

Conversely, if $x \in F$ there exist n, k such that $h_m(x) \leq c - 1/n$ for all $m \geq k$, hence $f(x) = \lim_{m\to\infty} h_m(x) \leq c - 1/n < c$, so $x \in E$. This shows that $E = F$. Note that each function h_m is continuous since g is differentiable and hence continuous. Each set $E_{n,m}$ is closed, being the inverse image under the continuous function h_m of the closed interval $(-\infty, c - 1/n]$, hence the intersection $\bigcap_{m=k}^{\infty} E_{n,m}$ is also closed for any k. It follows that $E = F$ is a countable union of closed sets, so its complement $S = \mathbb{R} \setminus E$ is a countable intersection of open sets, by De Morgan's law.

(ii) Let $g(x) = x^2 \sin(1/x)$ for real $x \neq 0$ and $g(0) = 0$. Routine calculation shows that g has derivative $f = g'$ given by $f(x) = 2x \sin(1/x) - \cos(1/x)$ for $x \neq 0$ and $f(0) = 0$. We claim that $S = \{x \in \mathbb{R} : f(x) \geq 1/2\}$ is not closed. Indeed, for any positive integer n let $x_n = 1/[(2n + 1)\pi]$. We have $f(x_n) = 1 \geq 1/2$, hence $\{x_n\}$ is a sequence in S. Since $f(0) = 0 < 1/2$, the set S is not closed since it does not contain its limit point $0 = \lim_{n\to\infty} x_n$.

Also solved by Robert Calcaterra, Souvik Dey (India), Russell Gordon, Northwestern University Math Problem Solving Group, Celia Schacht, and the proposer. There were 2 incomplete or incorrect solutions.

A continuum of translation-invariant orderings of $\mathbb{Z} \times \mathbb{Z}$ **April 2017**

2019. *Proposed by Greg Oman, University of Colorado, Colorado Springs, CO.*

Recall that a (strict) linear ordering of a set X is any binary relation $\prec$ on X such that:

- $x \prec y$ and $y \prec z$ implies $x \prec z$ for all $x, y, z \in X$ (transitivity), and
- exactly one of the three possibilities $x \prec y$, $y \prec x$, $x = y$ holds for all $x, y \in X$ (trichotomy).

A strict linear ordering $\prec$ of an additive abelian group $(G, +)$ is said to be *translation-invariant* when, for all $x, y, z \in G$, if $x \prec y$, then $x + z \prec y + z$. Consider the group $(\mathbb{Z} \times \mathbb{Z}, +)$ of pairs of integers under the operation of coordinate-wise addition $(a, b) + (c, d) = (a + c, b + d)$. Prove or disprove: There exist infinitely many distinct translation-invariant linear orderings of $(\mathbb{Z} \times \mathbb{Z}, +)$.

Solution by Pedro Acosta (student), West Morris Mendham High School, Mendham, NJ.

We prove that there are infinitely many distinct translation-invariant linear orderings of $(\mathbb{Z} \times \mathbb{Z}, +)$. Let α be any irrational number. Define $(a, b) \prec_\alpha (c, d)$ to mean that $a + b\alpha < c + d\alpha$ (in the standard ordering of the reals). Transitivity of $\prec_\alpha$ follows from transitivity of $<$ on real numbers. Next, by trichotomy of the ordering on $\mathbb{R}$, at most one of the relations $(a, b) \prec_\alpha (c, d)$, $(c, d) \prec_\alpha (a, b)$, $(a, b) = (c, d)$ may hold. In fact, if $(a, b) \not\prec_\alpha (c, d)$ and $(c, d) \not\prec_\alpha (a, b)$, we must have $a + b\alpha = c + d\alpha$, so $a - c = (d - b)\alpha$. Since $a - c$ is an integer, so is the multiple $(d - b)\alpha$ of the irrational number α, hence we must have $0 = (d - b)\alpha = a - c$, so $(a, b) = (c, d)$ in this case. We conclude that $\prec_\alpha$ is trichotomic and hence it is a linear ordering of $(\mathbb{Z} \times \mathbb{Z}, +)$. Now we prove the translation-invariance of $\prec_\alpha$ as follows: for $(e, f) \in \mathbb{Z} \times \mathbb{Z}$ and $(a, b) \prec_\alpha (c, d)$ we have

$$\begin{aligned}(a, b) \prec_\alpha (c, d) &\Rightarrow a + b\alpha < c + d\alpha \Rightarrow a + e + (b + f)\alpha < c + e + (d + f)\alpha \\ &\Rightarrow (a + e, b + f) \prec_\alpha (c + e, b + f).\end{aligned}$$

Finally, we show that if α and β are two distinct irrational numbers, then the orderings $\prec_\alpha$ and $\prec_\beta$ are different. Assume without loss of generality that $\alpha < \beta$. Since the rational numbers are dense in the real numbers, there exist integers m, n such that $n > 0$ and $\alpha < m/n < \beta$. It follows that $n\alpha < m$ and $m < n\beta$, so $(0, n) \prec_\alpha (m, 0)$ and $(m, 0) \prec_\beta (0, n)$; therefore, $\prec_\alpha$ and $\prec_\beta$ are distinct orderings. Since there are infinitely many irrational numbers α, there exist infinitely many distinct translation-invariant orderings of $(\mathbb{Z} \times \mathbb{Z}, +)$.

Editor's Note. The solution above shows that the cardinality of the set $\mathcal{O}$ of translation-invariant orderings of $(\mathbb{Z} \times \mathbb{Z}, +)$ is no less than the cardinality of the set of irrational numbers, which is equal to the cardinality $\mathfrak{c}$ of the continuum (i.e., the cardinality of the set of real numbers). Since $\mathfrak{c}$ is also the cardinality of the power set $\mathcal{P}$ of the countable set $(\mathbb{Z} \times \mathbb{Z}) \times (\mathbb{Z} \times \mathbb{Z})$ and every relation $\prec$ on $\mathbb{Z} \times \mathbb{Z}$ may be regarded as a member of $\mathcal{P}$, the set $\mathcal{O}$ has the same cardinality $\mathfrak{c}$ as $\mathcal{P}$.

Also solved by Robert Calcaterra, Robin Chapman, Souvik Dey (India), Joseph DiMuro, Abhay Goel, Tom Jager, Reiner Martin (Germany), Northwestern University Math Problem Solving Group, Greg Oman, Michael Reid, Celia Schacht, Joseph Walsh, Woobin Yang (South Korea), and the proposer. There was 1 incomplete or incorrect solution.

Transcendental moments of the arctangent **April 2017**

2020. *Proposed by Julien Sorel, PNI, Piatra Neamt, Romania.*

Find all natural numbers n such that the integral $I_n := \int_0^1 x^n \arctan x\, dx$ is a rational number.

Solution by Hongwei Chen, Christopher Newport University, Newport News, VA.

We show that I_n is a rational number precisely when $n = 4k + 3$ with $k = 0, 1, 2, \ldots$. For any nonnegative integer n, using integration by parts, we have

$$\begin{aligned}
I_n &= \frac{1}{n+1}\int_0^1 \arctan x\, d(x^{n+1}) = \frac{1}{n+1}\left(x^{n+1} \arctan x|_0^1 - \int_0^1 \frac{x^{n+1}}{1+x^2}\, dx\right) \\
&= \frac{1}{n+1}\left(\frac{\pi}{4} - \int_0^1 \frac{x^{n+1}}{1+x^2}\, dx\right) = \frac{1}{n+1}\int_0^1 \frac{1-x^{n+1}}{1+x^2}\, dx \\
&= \frac{1}{n+1}\int_0^1 \frac{1-(x^4+1-x^4)x^{n+1}}{1+x^2}\, dx \\
&= \frac{1}{n+1}\left(\int_0^1 \frac{1-x^{(n+4)+1}}{1+x^2}\, dx - \int_0^1 (1-x^2)x^{n+1}\, dx\right) \\
&= \frac{1}{n+1}\left((n+5)I_{n+4} - \frac{2}{(n+2)(n+4)}\right).
\end{aligned}$$

Hence, we have a recursion

$$I_{n+4} = \frac{1}{n+5}\left((n+1)I_n + \frac{2}{(n+2)(n+4)}\right) \qquad \text{for } n = 0, 1, 2 \ldots.$$

Write $n = 4k + r$ with k, r nonnegative integers and $r \leq 3$. It follows by induction from the identity above that I_n is rational if and only if I_r is rational. Routine computation gives

$$I_0 = \frac{1}{4}\pi - \frac{1}{2}\ln 2, \quad I_1 = \frac{1}{4}\pi - \frac{1}{2}, \quad I_2 = \frac{1}{2}\pi + \ln 2 - 1, \quad I_3 = \frac{1}{6}.$$

Obviously, I_3 is rational. We show that I_0, I_1, and I_2 are irrational. Since $\ln 2$ and $\ln(-1) = i\pi$ are nonzero logarithms of algebraic numbers, and since $i\pi / \ln 2$ is not a rational (nor even real) number, it follows from Baker's theorem en.wikipedia.org/wiki/Baker's_theorem that any linear combination $ai\pi + b\ln 2 + c$ with algebraic coefficients a, b, c such that a, b are not both zero is transcendental. Note that I_0, I_1, I_2 are of the form just described (e.g., $I_0 = (-i/4)i\pi + (-1/2)\ln 2$ where $-i/4$ and $-1/2$ are algebraic), hence they are transcendental and, *a fortiori,* also irrational numbers. Therefore, I_n is a rational number iff $n = 4k + 3$ with $k = 0, 1, 2, \ldots$ as claimed. (It is clear that I_n is actually transcendental when irrational.)

Also solved by Herb Balley, Brian D. Beasley, Brian Bradie, Robert Calcaterra, Robin Chapman (UK), Joseph DiMuro, Gregory Dresden, Dmitry Fleischman, Nilotpal Ghosh, Michael Goldenberg and Mark Kaplan, Russell Gordon, Raymond N. Greenwell, Don Hancock, Eugene A. Herman, Dale Hughes and Nancy Wang, Seonho Hwang (Korea), Weiping Li, Daniel López Aguayo (Mexico), Rick Mabry, Missouri State University Problem Solving Group, Moubinool Omarjee (France), Satyanand Singh, Albert Stadler (Switzerland), Jon Staggs, Nora Thornber, John Zacharias, and the proposer. There was 1 incomplete or incorrect solution.

Answers

Solutions to the Quickies from page 152.

A1079. Let $m = \lfloor\sqrt{n}\rfloor$ be the unique nonnegative integer satisfying $m^2 \leq n < (m+1)^2$. Since $((m+1)^2 - n) + (n - m^2) = (m+1)^2 - m^2 = 2m+1$ is odd, one of the nonnegative integers $(m+1)^2 - n$, $n - m^2$ must be odd and necessarily of the form $2k+1$ with k a nonnegative integer such that $k \leq m$. Observe that $2k+1 = (k+1)^2 - k^2$. To finish the proof when $n - m^2$ is odd, let $a = m$, $b = k+1$, and $c = k$; when $(m+1)^2 - n$ is odd, let $a = m+1$, $b = k$, and $c = k+1$. Since $k \leq m \leq \sqrt{n}$ by choice of m, such a, b, c are no greater than $\sqrt{n}+1$ and satisfy $n = a^2 + b^2 - c^2$.

A1080. Let f be the indicator function of the set of irrational numbers and let g be the indicator function of the set of nonzero rational numbers. (The indicator function I of a set $A \subseteq \mathbb{R}$ is defined by $I(x) = 1$ for $x \in A$, $I(x) = 0$ for $x \in \mathbb{R} \setminus A$.) Clearly f, g are both proto-monotone but their sum $f + g$ is the indicator function of the set of nonzero reals, which is not proto-monotone.

Next, let f be proto-monotone and g be continuous. For all $x \in \mathbb{R}$ we have

$$\begin{aligned}(f+g)(x) = f(x) + g(x) &\leq \limsup_{t\to x} f(t) + g(x)\\ &= \limsup_{t\to x}\,[f(t) + g(t)] = \limsup_{t\to x}(f+g)(t),\end{aligned}$$

where the equality passing from the first to the second line above is justified by the continuity of g: $\lim_{t\to x} g(t) = g(x)$. Similarly, we have $(f+g)(x) \geq \liminf_{t\to x}(f+g)(t)$, so $f + g$ is proto-monotone.

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.

Scheinerman, Edward, *The Mathematics Lover's Companion: Masterpieces for Everyone*, Yale University Press, 2017; xx + 274 pp, $28. ISBN 978-0-300-22300-2.

Author Scheinerman's intention is to share some of his favorite "gems" of mathematics, with the goal of eliciting joy in the reader. "Certainly, mathematics has great practical applications, but it also has profound beauty." This book would not enhance a reader's "math skills," but it will enlarge understanding of the nature, scope, and—yes—beauty of mathematics. The chapters are arranged according to the themes of number, space, and uncertainty, and they contain many of the topics that you might expect: prime numbers, irrational numbers, imaginary numbers, π, e, Fibonacci numbers, transfinite numbers, Benford's law; Pythagorean triples and Fermat's last theorem, packing circles, platonic solids, fractals, hyperbolic geometry; nontransitive dice, medical probability, chaos, Arrow's theorem, Newcomb's paradox. The treatment and pace are gentle, and the writing and motivation are engaging. All told, the book is a mind-broadening experience, of the sort that could be a basis for a liberal arts course in mathematics, and one that would also be particularly beneficial for potential elementary teachers.

Stewart, Seán M., *How to Integrate It: A Practical Guide to Finding Elementary Integrals*, Cambridge University Press, 2018; xi + 368 pp, $29.99 (P). ISBN 978-1-108-40819-6.

As I embark on teaching Calculus II again, I must question the importance and time involved in teaching techniques of integration. Author Stewart gives short shrift to motivating such study, his main rationale being intellectual curiosity. The book is thorough, from a mathematical perspective, treating in detail all the common methods (and some not so common), offering practice exercises and challenging problems (with answers). I wish that there had been also an algorithmic perspective, including citations to papers about integration leading out of various classes of elementary functions. I would also have enjoyed seeing explicitly the formula below, which avoids integration by parts; my students in probability find it extremely convenient:

Let P be a polynomial and m a nonzero constant. Then

$$\int e^{-mx}P(x)\,dx = \frac{-e^{-mx}}{m}\left(P(x) + \frac{P'(x)}{m} + \frac{P''(x)}{m^2} + \frac{P'''(x)}{m^3} + \cdots\right) + C.$$

Acheson, David, *The Calculus Story: A Mathematical Adventure*, Oxford University Press, 2017; xv + 188 pp, $16.95. ISBN 978-0-19880454-3.

Author Acheson offers a much-needed short account of the "big picture" of calculus as a whole, illustrated with examples and reproductions from historic publications. The first half of the book follows the themes of slope of a curve, area enclosed by a curve, infinite series, and motion. The second half starts from Leibniz's 1684 paper on rules for differentiation, discusses the controversies about who invented calculus and Berkeley's objections, explores series in more depth (e.g., for $\pi/4$, e, trig functions, and Fourier series), and concludes with chaos. Short pages, many illustrations, and a sense of telling a big story contribute to the success of the book.

Math. Mag. **91** (2018) 159–160. doi:10.1080/0025570X.2018.1432204

Thomas, Lyn, Jonathan Crook, and David Edelman, *Credit Scoring and Its Applications*, 2nd ed., Society for Industrial and Applied Mathematics (SIAM), 2017; xiv + 357 pp, \$102. ISBN 978-1-611974-55-3.

A moment's thought will result in the realization that scoring for consumer credit could involve a great deal of statistical theory. It does; and this book gives many of the details of the methods involved: data mining, various regression methods, linear and integer programming, neural networks, support vector machines, survival methods, Markov chains, and much more. And in light of the global financial crisis in 2008, the authors offer a moral: "[I]f you are or want to be a credit scorer, do it well, otherwise you could trash the world economy for a decade!"

Stipp, David, *A Most Elegant Equation: Euler's Formula and the Beauty of Mathematics*, Basic Books, 2017; viii + 223 pp, \$27. ISBN 978-0-46509378-6.

The author asserts that "great mathematics is as provocative, beautiful, and deep as great art or literature"; and the instance here is $e^{i\pi} + 1 = 0$, the equation referred to in the title. However, this equation, which "effectively compresses about two millennia's worth of big ideas in mathematics into a fantastically small package," did not show up in the author's mathematical education, nor in his son's high school mathematics courses. This presentation about the equation, written in a leisurely informal and conversational style, is at a more elementary level than Paul Nahin's *Dr. Euler's Fabulous Formula...* (2006). One appendix gives Euler's own derivation of the equation (using trigonometric identities and series), and the other uses the equation to show that i^i is a real number.

Donoho, David, From blackboard to bedside: High-dimensional geometry is transforming the MRI industry, *Notices of the American Mathematical Society* 65 (1) (January 2018) 40–44.

This article consists of notes for a briefing to members of Congress on one instance of how federal funding of mathematical research (about \$250M/yr) pays off in benefit to the public. Federally-funded pure mathematics, dating back to the 1960s, has now made possible speeding up magnetic resonance imaging (MRI) scans by a factor of 10. The basis for the improvement is to take fewer measurements and select the reconstruction that minimizes the Manhattan metric. Author Donoho emphasizes that it is the certainty and breadth of the *guarantees* offered by mathematical theorems that persuaded technologists to try the new algorithms, termed (felicitously) *compressed sensing*.

"Gifted." Film, 2017. Written by Tom Flynn, directed by Marc Webb. Produced by Day-day Films and others. 101 mins. Subtitles: English, French, Spanish Dubbed: French, Spanish. DVD, \$29.98, Blu-ray \$34.99.

This feature film dramatizes a challenge that you or friends may have faced as parents or relatives—or that may even have affected you: how to manage the growth of a "gifted" child, exploring and enhancing the child's gift while maintaining an otherwise-normal upbringing. The film features a spirited first-grade girl who is a mathematical prodigy, and whose deceased mother had been a mathematician working on the Navier–Stokes Millennium Prize Problem. The film depicts a custody battle between an uncle favoring a "normal" social upbringing for the girl and a grandmother intent on elite schooling, and the impact of the situation and the duel itself on the girl. The acting is excellent, and the film is enjoyable.

Pitici, Mircea (ed.), *The Best Writing on Mathematics 2017*, Princeton University Press, 2018; xvi + 224 pp, \$24.95 (P). ISBN 978-0-691-17863-9.

I am familiar with the "moon illusion" (the moon near the horizon seems larger than when in the heavens) but not with the "moon tilt illusion," whose explanation by Marc Frantz in terms of projective geometry is included in this year's harvest of fine mathematical writing. Other notable essays discuss "threshold" concepts in mathematics (e.g., reifying a function as itself an object in a space), investigate the rhombicuboctahedron in the famous portrait of Luca Pacioli, and speculate about who would have been likely to win Fields Medals in 1866 (probably not whom you would imagine).