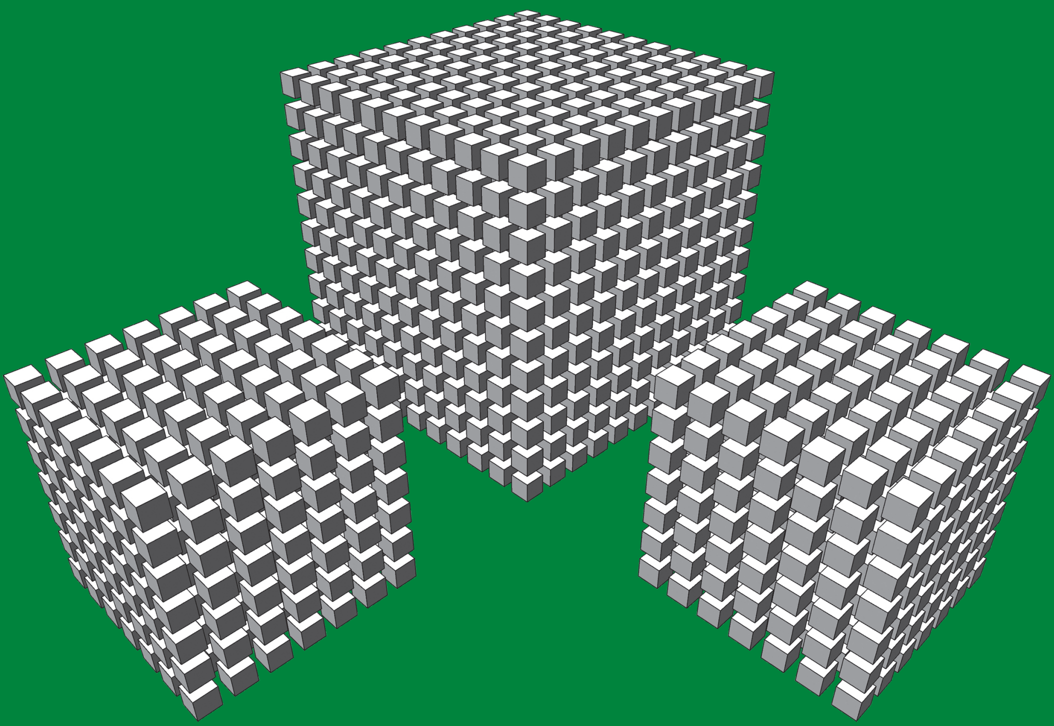
Vol. 90, No. 1, February 2017

MAA

MATHEMATICS MAGAZINE



- Hammers and feathers on the moon
- How do you define a spiral?
- Dividing and conquering the 15 puzzle
- Unifying the Pythagorean theorem in Euclidean, spherical, and hyperbolic geometries

EDITORIAL POLICY

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

Submissions of articles are required via the *Mathematics Magazine*'s Editorial Manager System. The name(s) of the author(s) should not appear in the file. Initial submissions in pdf or LaTeX form can be sent to the editor at www.editorialmanager.com/mathmag/.

The Editorial Manager System will cue the author for all required information concerning the paper. Questions concerning submission of papers can be addressed to the editor at mathmag@maa.org. Authors who use LaTeX are urged to use the Magazine article template. However, a LaTeX file that uses a generic article class with no custom formatting is acceptable. The template and the Guidelines for Authors can be downloaded from www.maa.org/pubs/mathmag.

MATHEMATICS MAGAZINE (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, NW, Washington, DC 20036 and Lancaster, PA, in the months of February, April, June, October, and December.

Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Address advertising correspondence to

MAA Advertising
1529 Eighteenth St. NW
Washington, DC 20036
Phone: (202) 319-8461
E-mail: advertising@maa.org

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

Change of address, missing issue inquiries, and other subscription correspondence can be sent to:

The MAA Customer Service Center
P.O. Box 91112
Washington, DC 20090-1112
(800) 331-1622
(301) 617-7800
maaservice@maa.org

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

Postmaster: Send address changes to Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, NW, Washington, DC 20036-1385.

Printed in the United States of America.

COVER IMAGE

2017 = $7^3 + 7^3 + 11^3$ © 2016 David A. Reimann (*Albion College*). Used by permission.

The number 2017 can be expressed as the sum of three cubes as depicted. It remains an open problem to determine if every integer $n \not\equiv 4,5((\text{mod } 9)$ can be expressed as a sum of three cubes, with 33 being the smallest positive integer where no solution is known (https://arxiv.org/abs/1604.07746).

Vol. 90, No. 1, February 2017

MATHEMATICS MAGAZINE

LETTER FROM THE EDITOR

One of the pleasures of editing THIS MAGAZINE is putting together an issue. This issue has diverse articles that should offer something for every reader. To start, Charles Groetsch's article is motivated by an experiment in which a hammer and a falcon's feather are dropped simultaneously on the moon. He uses calculus and elementary differential equations to examine the relationship between mass and fall time under two models of resistance treated in Newton's *Principia*.

The next article also has a historical bent. Although it is well known that it is not possible to trisect a general angle with straightedge and compass, it is possible under different conditions. Dave Richeson defines a trisectrix—a curve that can be used to trisect an angle—based on ideas from 1928 when Henry Scudder described how to use a carpenter's square to trisect an angle. Richeson also describes a compass that could be used to draw the curve. Compelling images complement the article.

You may know that Richard Feynman gave the name "Morrie's law" to an trigonometric identity because he learned of the identity from his friend Morrie Jacobs. In this issue, Gaston Brouwer generalizes the double-angle formulas for trigonometric functions to arrive at identities in the spirit of Morrie's law. For example, Brouwer shows that $\sin(n\theta)$ can be written as the product of n sine functions in which the arguments involve phase shifts of θ and a possible reflection about the horizontal axis. He provides similar formulas for cosines and tangents.

David Treeby uses centers of mass to generate some combinatorial identities, including one that involves Fibonacci tilings. The proofs are visual. Bernhard Klaassen's article is also visual, providing many examples of spirals on way to define a spiral tiling. This answers a question raised by Grünbaum and Shephard in the late 1970's.

You probably know Kaprekar's operation on four-digit numbers that involves subtracting permutations of the digits, but you may not know it by its name. Manuel Ricardo Falcão Moreira examines this 60-year old problem using dihedral symmetry to show that 6174 is the unique fixed point of the map. S. Muralidharan also examines a well-known problem, but from a different perspective. He uses divide and conquer to solve the 15 puzzle.

Juan Luis Varona provides a proof without words that shows why $\ell^1(\mathbb{R})$ is a subset of $\ell^2(\mathbb{R})$. Using only properties that the Euclidean, spherical, and hyperbolic geometries have in common, Robert Foote provides a formula that unifies the Pythagorean theorem for these geometries. The final article in the issue is by Konstantinos Gaitana. He explains how an approach used by Euler to prove Fermat's little theorem can be adapted to prove a more recent result on primes.

The issue concludes with our regular departments. Tracy Bennett's crossword puzzle *Mathematics in Love?* offers a not-so-obvious nod to Valentine's Day. Mathematical challenges appear in the Problems, while the Reviews help keep you up to date with new developments.

The solutions to the 77th William Lowell Putnam Competition that was held in December 2016 will appear in the April 2017 issue. Because of the new production time table and the desire to include student solutions, it is not possible to provide the solutions in the February issue any longer.

Michael A. Jones, Editor

ARTICLES

Hammer and Feather: Some Calculus of Mass and Fall Time

C. W. GROETSCH
The Citadel, The Military College of South Carolina
Charleston, SC 29409 USA
charles.groetsch@citadel.edu

Does a heavy object fall faster than a lighter one? In 1971 NASA astronaut Dave Scott, Commander of the Apollo 15 lunar mission, used a hammer and a falcon feather on the surface of the moon to give a dramatic illustration that this is not so [**4**]. It seems Aristotle disagreed [**1**, p. 365]:

> A given weight moves a given distance in a given time; a weight which is as great and more moves the same distance in less time...

In the late sixteenth century a brash twenty-something mathematics professor at the University of Pisa, Galileo Galilei, had the temerity to dispute the great philosopher [**3**, p. 26]:

> Of those moving bodies which are of the same material, Aristotle said that the larger moves more swiftly. ...he says that a large piece of gold moves more swiftly than a small piece. ...But how ridiculous this view is, is clearer than daylight.

A proper mathematical investigation of the question became possible with Newton's development of dynamics. Newton's program shows that, if the *only* active force is constant gravity, then the time of fall from a given height is independent of the mass. Indeed, in this case, the gravitational force mg (where m is the mass and g is the gravitational acceleration) and the velocity v are related as

$$m\frac{dv}{dt} = mg,$$

and so the velocity is independent of the mass, as is the fall time. Commander Scott's lunar experiment can be taken as validation of the independence of mass and fall time and a vindication of Galileo. But was it a repudiation of Aristotle?

What if the object falls not in a vacuum, but through a resisting medium? After all, Aristotle formulated in his physical theory the notion of *horror vacui*—nature's rejection of a vacuum. So perhaps he couldn't conceive of dropping objects in a vacuum. Does a massive object fall from a given height in a resisting medium in less time than a less massive one of the same size and shape? The medieval scholar John Philiponus argued that this is so [**5**, p. 433]:

Math. Mag. **90** (2017) 3–7. doi:10.4169/math.mag.90.1.3. © Mathematical Association of America
MSC: Primary 34A00

> Clearly, then, it is the natural weight of bodies, one having a greater and another a lesser downward tendency, that causes differences in motion. For that which has a greater downward tendency divides the medium better. ... The same space will consequently be traversed by the heavier body in shorter time and the lighter body in a longer time...

Investigating the relationship between mass and fall time for the two basic resistance models introduced by Newton in his *Principia* [**2**, Book II, Sections I and II] is a useful vehicle for illustrating several techniques and results from calculus. The analysis involved can be used for enrichment topics in first year calculus and differential equations courses. Among these are separable and linear differential equations, implicit differentiation, hyperbolic trigonometry, properties of the parabola, and some fun with inequalities.

Quadratic resistance

A common model for air resistance posits that the resistive force is proportional to the square of the velocity. In this model the equation of motion is

$$m\,\frac{dv}{dt} = mg - kv^2, \quad v(0) = 0,$$

where k is a constant of proportionality, called the drag coefficient. On separating variables and integrating we find that the velocity v at time t satisfies

$$\int_0^v \frac{du}{1-(\sqrt{\frac{k}{mg}}u)^2} = gt,$$

that is,

$$\sqrt{\frac{m}{kg}}\tanh^{-1}\left(\sqrt{\frac{k}{mg}}v\right) = t.$$

One approach to our investigation of fall time is to solve for the velocity and then show that, for any fixed positive t, the velocity $v(t)$ is an *increasing* function of mass, leading to a fall time that decreases with the mass. We instead take the direct approach of analyzing the fall time itself.

If we denote by s the distance fallen from an arbitrary fixed height, say 1 unit, then,

$$s(t) = \int_0^t v(\tau)\,d\tau = \sqrt{\frac{mg}{k}}\int_0^t \tanh\left(\sqrt{\frac{kg}{m}}\tau\right)d\tau,$$

and hence the fall time $T(m)$ satisfies

$$\sqrt{\frac{k}{mg}} = \int_0^{T(m)} \tanh\left(\sqrt{\frac{kg}{m}}\tau\right)d\tau = \sqrt{\frac{m}{kg}}\ln\left(\cosh\left(\sqrt{\frac{kg}{m}}T(m)\right)\right).$$

In this quadratic model the fall time therefore has the explicit representation:

$$T(m) = \sqrt{\frac{m}{kg}}\cosh^{-1}\left(e^{\frac{k}{m}}\right).$$

We began this note with the question: Do heavier objects fall faster than lighter ones? In other words, is $T(m)$ a decreasing function of the mass m? Differentiating $T(m)$, we obtain

$$\sqrt{kg}\,\frac{d}{dm}T(m) = \frac{1}{2}m^{-1/2}\cosh^{-1}\left(e^{k/m}\right) + m^{1/2}\frac{1}{\sqrt{e^{2k/m}-1}}e^{k/m}\left(-\frac{k}{m^2}\right).$$

Therefore, $\frac{dT}{dm} < 0$ if and only if

$$\cosh^{-1}\left(e^{k/m}\right) < 2\frac{k}{m}\frac{1}{\sqrt{e^{2k/m}-1}}e^{k/m}, \quad \text{for all} \quad m > 0.$$

This seemingly formidable inequality brings to mind the comment attributed to Harald Bohr by G. H. Hardy [**6**, p. 81]:

> all analysts spend half their time hunting through the literature for inequalities which they want to use and cannot prove.

Fortunately, in this case a change of variables suggests itself and leads to a proof. Setting $w = e^{k/m}$, the inequality is equivalent to

$$\cosh^{-1}(w) < \frac{2w\ln w}{\sqrt{w^2-1}}, \quad \text{for all} \quad w > 1.$$

Now setting $w = \cosh u$, we find this is equivalent to

$$u\tanh u < 2\ln(\cosh u), \quad \text{for all} \quad u > 0.$$

Since the expressions on the left and right of the inequality sign agree at $u = 0$, this inequality is established if the derivatives of the left and right hand sides bear the same relationship, that is, if

$$\tanh u + u\frac{1}{\cosh^2 u} < 2\tanh u,$$

or equivalently,

$$u < \cosh u \sinh u, \quad \text{for all} \quad u > 0.$$

But this is equivalent to

$$2u < \sinh 2u, \quad \text{for all} \quad u > 0,$$

which is easily seen to be true. Score one for Aristotle. In the quadratic model the fall time is a decreasing function of mass.

Linear resistance

A simpler model in which resistance is taken to be proportional to velocity (according to Newton this is "more a mathematical hypothesis than a physical one" [**2**, p. 244]) offers a number of lessons and could provide enrichment in a first year calculus class. While one might expect this linear model to submit to simpler analysis than the quadratic model, it holds a surprise, as will be seen. In this model the velocity satisfies

$$m\frac{dv}{dt} = mg - kv, \quad v(0) = 0,$$

for a given positive drag coefficient k. This linear initial value problem has the unique solution:

$$v = \frac{gm}{k}\left(1 - e^{-\frac{k}{m}t}\right).$$

An additional integration yields the distance fallen in time t:

$$s(t) = \frac{gm}{k}t + \frac{gm^2}{k^2}\left(e^{-\frac{k}{m}t} - 1\right).$$

Therefore, the time $T(m)$ for the object to fall from a unit height satisfies

$$1 = \frac{gm^2}{k^2}\left[\frac{kT(m)}{m} + \left(e^{-\frac{kT(m)}{m}} - 1\right)\right]. \tag{1}$$

Our goal, as before, is to discover if $T(m)$ is a decreasing function of m by examining its derivative $T'(m)$. But, unlike in the quadratic model, we now lack an explicit representation of the fall time. Is all lost? Certainly not! Alert calculus students will note that an implicit relationship suggests use of a familiar technique: implicit differentiation. Before plunging into the differentiation process, it helps to simplify notation. A glance at (1) suggests setting

$$U(m) = \frac{kT(m)}{m} \quad \text{and} \quad a(m) = \frac{k^2}{gm^2}.$$

The characteristic relationship then takes the cleaner form:

$$a(m) + (1 - e^{-U(m)}) = U(m). \tag{2}$$

That is, $U(m)$ is a *fixed point* of the function $f(x) = a(m) + (1 - e^{-x})$. A quick sketch (or an easy analytic proof using the intermediate value theorem) convinces that f has a unique positive fixed point u and that a given positive number r is greater than u if and only if $r > f(r)$.

Since $T(m) = mU(m)$, we see that $T'(m) < 0$ if and only if

$$U(m) + mU'(m) < 0.$$

To tidy the typography, we suppress the explicit dependence of a and U on m in the notation below. By (2) we see that

$$U' = a' + e^{-U}U',$$

that is,

$$U'(1 - e^{-U}) = a'.$$

From (2) we have $1 - e^{-U} = U - a$. Substituting this above and multiplying by m, we find

$$mU'(U - a) = ma' = -2a,$$

and so,

$$(U + mU')(U - a) = U^2 - aU - 2a.$$

But $U - a$ is positive (by (2)), and therefore $T'(m) = U + mU' < 0$ if and only if

$$U^2 - aU - 2a < 0. \tag{3}$$

This inequality ties the monotonicity of the fall time function $T(m)$ to a simple algebraic feature of quadratics. Since $U(m)$ is positive for all positive m, the inequality (3) holds if and only if U is less than the positive root of the quadratic $y = x^2 - ax - 2a$. Denote this positive root by r. Then $T(m)$ is monotone decreasing if and only if

$$U < \frac{a + \sqrt{a^2 + 8a}}{2} =: r. \tag{4}$$

However, as noted above, the fixed point U of $f(x) = a + (1 - e^{-x})$ is less than r if and only if $f(r) < r$, that is if and only if

$$a + (1 - e^{-r}) < r. \tag{5}$$

The equality in (4) gives $a = r^2/(2 + r)$, therefore (5) is equivalent to

$$\frac{r^2}{2 + r} + (1 - e^{-r}) < r,$$

that is,

$$2 - r < (2 + r)e^{-r}.$$

But it is easy to see that this inequality holds for all positive r as both expressions on either side of the inequality vanish at $r = 0$, while the derivatives of the respective sides of the inequality satisfy

$$-1 < (-1 - r)e^{-r},$$

since $e^r > 1 + r$. Therefore, $T'(m) < 0$, for each positive m, the time of fall decreases as the mass increases.

REFERENCES

1. Aristotle, *On the Heavens*, Book I. Trans. J. L. Stocks. Works of Aristotle, Vol. I, Great Books of the Western World, Encyclopedia Britannica, Chicago, 1952.
2. F. Cajori (Ed.), *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World*. Trans. into English by Andrew Mott in 1729, the translations revised, and supplied with an historical and explanatory appendix by Florian Cajori, Univ. of California Press, Berkeley, 1934.
3. I.E. Drabkin, S. Drake, *On Motion and On Mechanics*, Comprising *De Motu* (ca. 1590), trans. with Introduction and Notes by I.E. Drabkin, and *Le Meccaniche* (ca. 1600), trans. with Introduction and Notes by Stillman Drake, Univ. of Wisconsin Press, Madison, 1960.
4. http://www.youtube.com/watch?v=03SPBXALJZI
5. M. Clagett, *The Science of Mechanics in the Middle Ages*. Univ. of Wisconsin Press, Madison, 1959.
6. R. Roy, *Sources in the Development of Mathematics: Infinite Series and Products from the Fifteenth to the Twenty-first Century*. Cambridge Univ. Press, Cambridge, 2011.

Summary. Techniques from calculus and elementary differential equations are used to explore the relationship between mass and fall time in the two models of resistance treated by Newton in his *Principia*. The material can be used to enrich various undergraduate classes and to acquaint, and perhaps interest, students in an aspect of the fascinating story of the mathematization of nature.

CHUCK GROETSCH (MR Author ID: 77275) hails from the Irish Channel neighborhood in New Orleans. He received his Ph.D. from LSU and was long affiliated with the University of Cincinnati where he served at times as head of the Department of Mathematical Sciences and dean of the College of Arts and Sciences. Currently he is Distinguished Professor of Mathematical Science at The Citadel where he was the founding dean of the School of Science and Mathematics. He enjoys blending mathematics, science, and history in his writing and teaching (the former, he believes, is a subset of the latter).

A Trisectrix From a Carpenter's Square

DAVID RICHESON
Dickinson College
Carlisle, PA 17013
richesod@dickinson.edu

In 1837, Pierre Wantzel proved that it is impossible to trisect an arbitrary angle using only a compass and straightedge [**13**]. However, it *is* possible to trisect an angle if we are allowed to add additional items to our toolkit.

We can trisect an angle if we have a marked straightedge [**7**, pp. 185–87], a Mira (a vertical mirror used to teach transformational geometry) [**4**], a tomahawk-shaped drawing tool [**11**], origami paper [**5**], or a clock [**9**]. We can also trisect an angle if we are able to use curves other than straight lines and circles: a hyperbola [**15**, pp. 22–23], a parabola [**3**, pp. 206–08], a quadratrix [**7**, pp. 81–86], an Archimedean spiral [**2**, p. 126], a conchoid [**15**, pp. 20–22], a trisectrix of Maclaurin [**10**], a limaçon [**15**, pp. 23–25], and so on; such a curve is called a *trisectrix*. In many cases we can use specially designed compasses to draw these or other trisectrices. For instance, Descartes designed such a compass [**1**, pp. 237–39].

The literature on different construction tools and techniques, new compasses, and their relationships to angle trisection and the other problems of antiquity is vast. A reader interested in learning more may begin with [**1, 6, 7, 8, 15**].

In this note we describe a trisection technique discovered by Henry Scudder in 1928 that uses a carpenter's square [**12**]. Then we use the ideas behind this construction to produce a new trisectrix, and we describe a compass that can draw the curve.

Angle Trisection Using a Carpenter's Square. A carpenter's square—a common drawing tool found at every home improvement store—consists of two straightedges joined in a right angle. To carry out Scudder's construction, we need a mark on one leg such that the distance from the corner is twice the width of the other leg. For instance, we will assume that the longer leg is one inch wide and that there is a mark two inches from the corner on the shorter leg.

Let's say we wish to trisect the angle $\angle AOB$ in Figure 1. First, we draw a line l parallel to and one inch away from AO; this can be accomplished using a compass and straightedge, but a simpler method is to use the long leg of the carpenter's square as a double-edged straightedge. We now perform the step that is impossible using Euclidean tools: Place the carpenter's square so that the inside edge passes through O, the two inch mark lies on the line BO, and the corner sits on the line l (at the point C, say). Then the inside edge of the carpenter's square and the line CO trisect the angle. This procedure works for any angle up to 270°, although the larger the angle, the narrower the short leg of the carpenter's square must be.

In fact, we do not need a carpenter's square to carry out this construction. All we need is a T-shaped device (shown on the right in Figure 1) in which the top of the T is two inches long. It is not difficult to see that this technique trisects the angle: The right triangles COF, COE, and DOE in Figure 1 are congruent.

A New Compass. We now use the carpenter's square as inspiration to create a compass to draw a trisectrix (see Figure 2). The device has a straightedge that is one inch

Math. Mag. **90** (2017) 8–11. doi:10.4169/math.mag.90.1.8.
MSC: Primary 01A20, Secondary 51M15

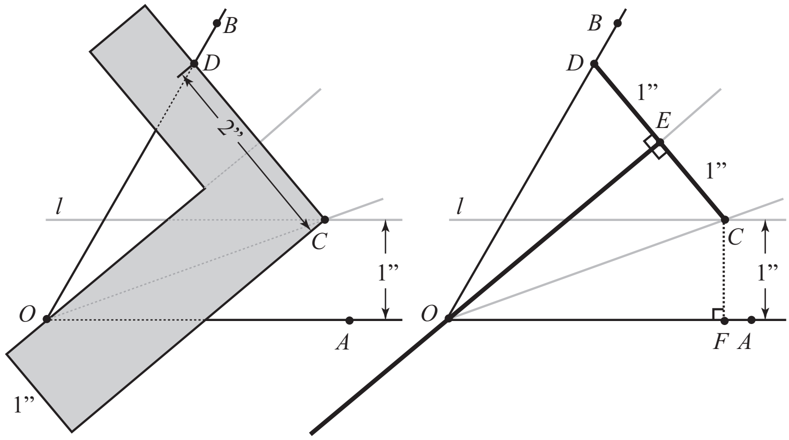


Figure 1 A carpenter's square or a T-shaped tool can be used to trisect an angle.

wide and a T-shaped tool with pencils at both ends of the two-inch top of the T. The long leg of the T passes through a device at one corner of the straightedge that can rotate and that allows the T to slide back and forth. One pencil draws a line along the straightedge. The other pencil draws the curve we call the *carpenter's square curve*.

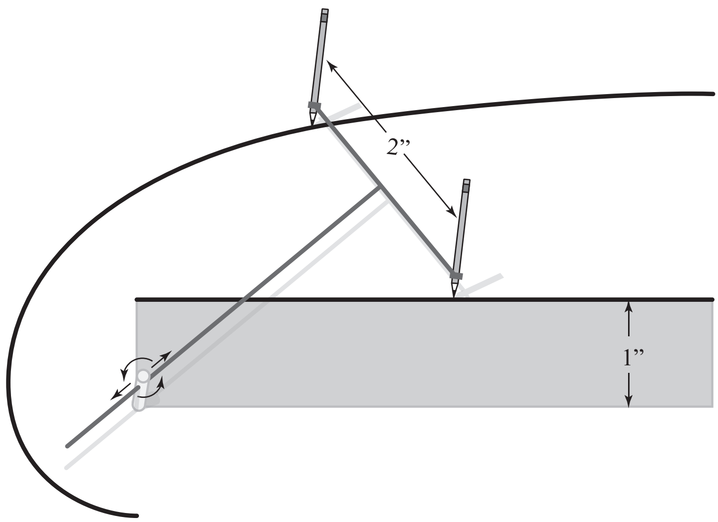


Figure 2 A compass to draw the carpenter's square curve.

We use the compass as follows. Suppose we would like to trisect $\angle AOB$ in Figure 3. Place the bottom of the straightedge along OA with the corner (to which the T is attached) at O. Use the compass to draw the straight line l and the carpenter's square curve. Say that BO intersects the curve at D. Open an ordinary compass two inches. (One way to do this is to draw a line perpendicular to OA at O that intersects the carpenter's square curve at F. As we will see in the next section, OF is two inches.) Use the compass to draw a circle with center D. It will intersect l at two points. Label the right-most point (viewed from the perspective of Figure 3) C. Then OC trisects the angle. Use an ordinary compass and straightedge to bisect $\angle COD$ to obtain the other trisecting ray.

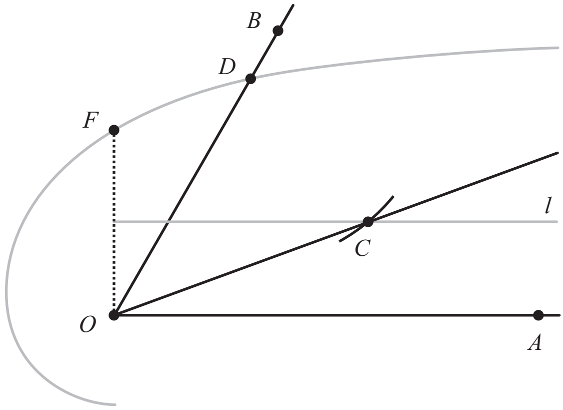


Figure 3 We can use the carpenter's square curve to trisect an angle.

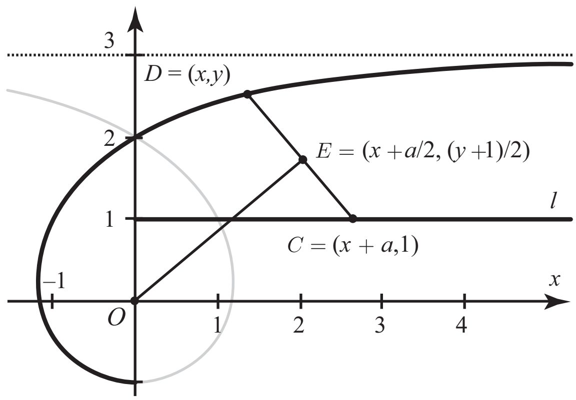


Figure 4 The carpenter's square curve.

The Carpenter's Square Curve. What is this carpenter's square curve? Does it have a closed form? Is it algebraic or transcendental? (In [**14**], Yates used a carpenter's square in a different way to generate a different curve—a cissiod. Yates gives an algebraic expression for his curve and shows that it can be used to compute cube roots.)

First, we introduce x- and y-axes. Let O be the origin and OA be the positive x-axis (see Figure 4). Let $D = (x, y)$ and $C = (x + a, 1)$. Because $|CD| = 2$, $a^2 + (1 - y)^2 = 4$, and hence, $a = \sqrt{(3 - y)(y + 1)}$. (Notice that $a \geq 0$ throughout the construction.) Also, E is the midpoint of CD, so $E = (x + a/2, (y + 1)/2)$. Because CD and EO are perpendicular,

$$\frac{(y+1)/2}{x + a/2} = -\frac{a}{1-y}.$$

Substituting our expression for a and simplifying, we obtain

$$x^2 = \frac{(y-2)^2(y+1)}{3-y}.$$

This algebraic curve has a self-intersection at (0, 2), and $y = 3$ is a horizontal asymptote. As we see in Figure 4, our compass traces one branch of the curve, namely

$$x = (y-2)\sqrt{\frac{y+1}{3-y}}.$$

To see an interactive applet of this trisection visit *ggbtu.be/mJpaNPATB*.

REFERENCES

1. H. J. M. Bos, *Redefining Geometrical Exactness: Descartes' Transformation of the Early Modern Concept of Construction*. Springer-Verlag, New York, 2001.
2. C. B. Boyer, U. C. Merzbach, *A History of Mathematics.* John Wiley& Sons, New York, 1991.
3. R. Descartes, *The Geometry of René Descarte: Translated from French and Latin by David Eugene Smith and Marcia L. Latham*. Dover Publications, New York, 1954.
4. J. W. Emert, K. I. Meeks, R. B. Nelsen, Reflections on a Mira, *Amer. Math. Monthly* **101** no. 6 (1994) 544–549.
5. K. Fusimi, Trisection of angle by Abe, *Saiensu Supplement* (1980).
6. T. L. Heath, *A History of Greek Mathematics. Vol. I: From Thales to Euclid.* Carendon Press, Oxford, 1921.
7. W. R. Knorr, *The Ancient Tradition of Geometric Problems.* Dover Publications, New York, 1993.
8. G. E Martin, *Geormetric Constructions.* Springer-Verlag, New York, 1998.
9. L. Moser, The watch as angle trisector, *Scripta Math.* **13** (1947) 57.
10. H. C. Reyerson, Anyone can trisect an angle, *Math. Teacher* **70** (1977) 319-321.
11. B. S. Sackman, The tomahawk, *Math. Teacher* **49** (1956) 280–281.
12. H. T. Scudder, Discussions: How to trisect an angle with a carpenter's square, *Amer. Math. Monthly* **35** no. 5 (1928) 250–251.
13. P. L. Wantzel, Recherches sur les moyens de reconnaître si un Problème de Géométrie peut se résoudre avec la règle et le compas, *Journal de Mathématiques Pures et Appliquées* **2** no 1 (1837) 366–372.
14. R. C. Yates, The angle ruler, the marked ruler and the carpenter's square, *National Math. Mag.* **15** no. 2 (1940) 61–73.
15. ———, *The Trisection Problem*. The Franklin Press, Baton Rouge, LA, 1942.

Summary. In 1928, Henry Scudder described how to use a carpenter's square to trisect an angle. We use the ideas behind Scudder's technique to define a trisectrix—a curve that can be used to trisect an angle. We also describe a compass that could be used to draw the curve.

DAVID RICHESON (MR Author ID: 642588) is a professor of mathematics at Dickinson College and is the editor of *Math Horizons*. He is the author of *Euler's Gem: The Polyhedron Formula and the Birth of Topology* (Princeton University Press, 2008), which received the MAA's 2010 Euler Book Prize.

A Generalization of the Angle Doubling Formulas for Trigonometric Functions

GASTON BROUWER
Middle Georgia State University
Macon, GA 31206
gaston.brouwer@mga.edu

A colleague of mine at Middle Georgia State University, David Vogel, brought the following identity to my attention:

$$\cos(20^\circ)\cos(40^\circ)\cos(80^\circ) = \frac{1}{8}.$$

I was intrigued by this identity since none of the factors in the product are rational, yet the product is. I discussed the identity with another colleague, Jeremy Aikin, and he posed a couple of interesting questions, one of which was whether there were other products like this, possibly containing more than three factors. Or, more formally, is it possible to find integers $0 < i_1 < i_2 < \ldots < i_n < 90$ such that $\prod_{k=1}^{n} \cos(i_k^\circ) \in \mathbb{Q}$ for values $n \geq 3$? I will call these products "rational products of integer cosines." I set out to find other such products, and my first approach was to use a computer to find them. To my surprise, there were many: the computer found 60 rational products of integer cosines consisting of five factors, for example. Upon further investigation, I learned that the original equation is also known as "Morrie's law." The name is due to the physicist Richard Feynman, who picked that name because he learned of the identity during his childhood from a boy named Morrie Jacobs [**3**]. In [**1**], the authors showed that Morrie's law is a special case of the Morrie-type formula:

$$2^k \prod_{j=0}^{k-1} \cos(2^j x) = \frac{\sin(2^k x)}{\sin x}$$

with $x = 20^\circ$ and $k = 3$. In [**4**] and [**5**], the authors give a very elegant geometric proof of Morrie's law and another Morrie-type formula, respectively. Unfortunately, Morrie-type formulas only account for a small fraction of all rational products of integer cosines, and the methods used in the geometric proof would be hard to generalize for other rational products of integer cosines. However, while I was studying the products more closely, I seemed to have stumbled across an intriguing pattern for $\sin(n\theta)$:

$$\sin(3\theta^\circ) = 4\sin(\theta^\circ)\sin(60^\circ - \theta^\circ)\sin(60^\circ + \theta^\circ),$$
$$\sin(4\theta^\circ) = 8\sin(\theta^\circ)\sin(45^\circ - \theta^\circ)\sin(45^\circ + \theta^\circ)\sin(90^\circ - \theta^\circ), \text{ and}$$
$$\sin(5\theta^\circ) = 16\sin(\theta^\circ)\sin(36^\circ - \theta^\circ)\sin(36^\circ + \theta^\circ)\sin(72^\circ - \theta^\circ)\sin(72^\circ + \theta^\circ).$$

I soon realized that this pattern for the sine function could be seen as a generalization of the angle doubling formula:

$$\sin(2\theta^\circ) = 2\sin(\theta^\circ)\sin(90^\circ - \theta^\circ).$$

Math. Mag. **90** (2017) 12–18. doi:10.4169/math.mag.90.1.12.
MSC: Primary 97G60

In this paper, I will use an analytical approach to prove the above multiple angle formulas and other similar formulas.

Results

For the remainder of this paper angles are assumed to be in radians unless specifically stated otherwise.

Theorem 1. *For n an odd integer,*

$$\sin(n\theta) = 2^{n-1}\sin(\theta)\prod_{m=1}^{\frac{n-1}{2}}\left[\sin\left(\frac{m}{n}\pi+\theta\right)\sin\left(\frac{m}{n}\pi-\theta\right)\right].$$

Proof. From De Moivre's theorem, it follows that

$$\cos(n\theta) + i\sin(n\theta) = (\cos\theta + i\sin\theta)^n = \sum_{k=0}^{n}\binom{n}{k}\cos^{n-k}\theta(i\sin\theta)^k. \tag{1}$$

By comparing the imaginary parts on the left and right side of the equation, we obtain for odd n:

$$\sin(n\theta) = \sum_{j=0}^{\frac{n-1}{2}}(-1)^j\binom{n}{2j+1}\cos^{n-(2j+1)}\theta\sin^{2j+1}\theta. \tag{2}$$

By using the fundamental identity $\sin^2\theta + \cos^2\theta = 1$ and by factoring out a $\sin\theta$, we can rewrite (2) as

$$\sin(n\theta) = \sin\theta\sum_{j=0}^{\frac{n-1}{2}}(-1)^j\binom{n}{2j+1}(1-\sin^2\theta)^{\frac{n-1}{2}-j}\sin^{2j}\theta. \tag{3}$$

We can see that, for n odd, $\sin(n\theta)$ can be written as $\sin\theta$ times a polynomial in $\sin^2(\theta)$. To find out what this polynomial looks like, we will use the binomial theorem to expand the powers of $(1-\sin^2\theta)$. After simplifying, we obtain

$$\sin(n\theta) = \sin\theta\sum_{j=0}^{\frac{n-1}{2}}\sum_{m=0}^{\frac{n-1}{2}-j}(-1)^{\frac{n-1}{2}-m}\binom{n}{2j+1}\binom{\frac{n-1}{2}-j}{m}\sin^{n-1-2m}\theta. \tag{4}$$

Hence,

$$\sin(n\theta) = \sin\theta\, P_{\frac{n-1}{2}}(\sin^2\theta) \tag{5}$$

where $P_{\frac{n-1}{2}}$ is a polynomial given by

$$P_{\frac{n-1}{2}}(\sin^2\theta) = \sum_{j=0}^{\frac{n-1}{2}}\sum_{m=0}^{\frac{n-1}{2}-j}(-1)^{\frac{n-1}{2}-m}\binom{n}{2j+1}\binom{\frac{n-1}{2}-j}{m}\sin^{n-1-2m}\theta. \tag{6}$$

By considering the term with $m = 0$ in (6), we obtain that the degree of $P_{\frac{n-1}{2}}$ is $\frac{n-1}{2}$ and the leading coefficient is

$$(-1)^{\frac{n-1}{2}} \sum_{j=0}^{\frac{n-1}{2}} \binom{n}{2j+1} = (-1)^{\frac{n-1}{2}} 2^{n-1}. \tag{7}$$

Since $\sin(m \cdot \pi) = 0$ for any integer m, it follows from (5) that $\sin^2\left(\frac{m}{n} \cdot \pi\right)$ is a root of $P_{\frac{n-1}{2}}$ for any nonzero integer m that is not a multiple of n. In particular, we have $\frac{n-1}{2}$ distinct roots for $m = 1, \ldots, \frac{n-1}{2}$. Hence, we can rewrite (5) as

$$\sin(n\theta) = (-1)^{\frac{n-1}{2}} 2^{n-1} \sin\theta \prod_{m=1}^{\frac{n-1}{2}} \left[\sin^2\theta - \sin^2\left(\frac{m}{n} \cdot \pi\right)\right]. \tag{8}$$

In order to obtain the desired result, we need to rewrite (8). By using the sum and difference identities for the sine function, we see that for every u and every x, the following holds:

$$\begin{aligned}\sin(u+x)\sin(u-x) &= (\sin u \cos x + \sin x \cos u)(\sin u \cos x - \sin x \cos u)\\ &= \sin^2 u \cos^2 x - \sin^2 x \cos^2 u\\ &= \sin^2 u(1-\sin^2 x) - \sin^2 x(1-\sin^2 u)\\ &= \sin^2 u - \sin^2 x. \end{aligned}\tag{9}$$

Hence, we can rewrite (8) as

$$\sin(n\theta) = (-1)^{\frac{n-1}{2}} 2^{n-1} \sin\theta \prod_{m=1}^{\frac{n-1}{2}} \left[\sin\left(\theta - \frac{m}{n} \cdot \pi\right) \sin\left(\theta + \frac{m}{n} \cdot \pi\right)\right]. \tag{10}$$

By moving $(-1)^{\frac{n-1}{2}}$ into the product and using the fact that the sine function is odd, we obtain the desired result. ■

Example. Theorem 1 can be used to prove Morrie's law. With $n = 3$ and $\theta = 10°$, we have

$$\begin{aligned}\sin(3 \cdot 10°) &= 2^{3-1} \sin(10°)\sin(60° - 10°)\sin(60° + 10°)\\ \frac{1}{2} &= 4\cos(80°)\cos(40°)\cos(20°).\end{aligned}$$

Theorem 2. *For n an even integer,*

$$\sin(n\theta) = 2^{n-1} \prod_{m=1}^{\frac{n}{2}} \left[\sin\left(\frac{m-1}{n}\pi + \theta\right) \sin\left(\frac{m}{n}\pi - \theta\right)\right].$$

Proof. By letting $n = 2k$, we obtain the equivalent statement

$$\sin(2k\theta) = 2^{2k-1} \prod_{m=1}^{k} \sin\left(\frac{m-1}{k}\frac{\pi}{2} + \theta\right) \prod_{m=1}^{k} \sin\left(\frac{m}{k}\frac{\pi}{2} - \theta\right) \text{ for all } k \in \mathbb{N}. \tag{11}$$

We will prove (11) by induction. The statement holds for $k = 1$ by the standard angle doubling formula for the sine function so that $\sin 2\theta = 2\sin\theta\sin(\frac{\pi}{2} - \theta)$. Now assume that (11) holds for all $k \leq n$. We will show that (11) holds for $k = n+1$. By the angle doubling formula for the sine function, we have that

$$\sin[2(n+1)\theta] = 2\sin[(n+1)\theta]\sin\left[\frac{\pi}{2} - (n+1)\theta\right]. \tag{12}$$

There are two cases to be considered.

Case 1: $n+1$ *is even.*

In this case, we can use the induction hypothesis with $k = \frac{n+1}{2}$ to obtain

$$\sin[(n+1)\theta] = 2^n \prod_{m=1}^{\frac{n+1}{2}} \sin\left(\frac{2m-2}{n+1}\frac{\pi}{2} + \theta\right) \prod_{m=1}^{\frac{n+1}{2}} \sin\left(\frac{2m}{n+1}\frac{\pi}{2} - \theta\right), \tag{13}$$

and, after some simplifying,

$$\sin\left[\frac{\pi}{2} - (n+1)\theta\right] = \sin\left[2\frac{n+1}{2}\left(\frac{\pi/2}{n+1} - \theta\right)\right] =$$

$$2^n \prod_{m=1}^{\frac{n+1}{2}} \sin\left(\frac{2m-1}{n+1}\frac{\pi}{2} - \theta\right) \prod_{m=1}^{\frac{n+1}{2}} \sin\left(\frac{2m-1}{n+1}\frac{\pi}{2} + \theta\right). \tag{14}$$

We can observe that the factors of (13) contain the even multiples of $\frac{\pi/2}{n+1}$, and the factors of (14) contain the odd multiples of $\frac{\pi/2}{n+1}$. Hence, by multiplying the expressions in (13) and (14), we get

$$\sin[2(n+1)\theta] = 2\sin[(n+1)\theta]\sin\left[\frac{\pi}{2} - (n+1)\theta\right] =$$

$$2 \cdot 2^n \cdot 2^n \prod_{m=1}^{n+1} \sin\left(\frac{m-1}{n+1}\frac{\pi}{2} + \theta\right) \prod_{m=1}^{n+1} \sin\left(\frac{m}{n+1}\frac{\pi}{2} - \theta\right). \tag{15}$$

This completes the induction step in the case that $n+1$ is even.

Case 2: $n+1$ *is odd.*

In this case, we can use Theorem 1 to see that $\sin[(n+1)\theta]$ is equal to

$$2^{n+1-1}\sin\theta \prod_{m=1}^{\frac{n+1-1}{2}} \sin\left(\frac{m}{n+1}\pi + \theta\right) \prod_{m=1}^{\frac{n+1-1}{2}} \sin\left(\frac{m}{n+1}\pi - \theta\right) =$$

$$2^n \sin\theta \prod_{m=1}^{\frac{n}{2}} \sin\left(\frac{2m}{n+1}\frac{\pi}{2} + \theta\right) \prod_{m=1}^{\frac{n}{2}} \sin\left(\frac{2m}{n+1}\frac{\pi}{2} - \theta\right), \tag{16}$$

and after some simplifying, that $\sin\left[\frac{\pi}{2} - (n+1)\theta\right]$ is equal to

$$2^n \sin\left(\frac{\pi/2}{n+1} - \theta\right) \prod_{m=1}^{\frac{n}{2}} \sin\left(\frac{2m+1}{n+1}\frac{\pi}{2} - \theta\right) \prod_{m=1}^{\frac{n}{2}} \sin\left(\frac{2m-1}{n+1}\frac{\pi}{2} + \theta\right). \tag{17}$$

We can complete the induction step by using a similar argument as was used in the first case. ■

Example. With $n = 6$ and $\theta = 5^\circ$, we obtain

$$\sin(5^\circ)\sin(25^\circ)\sin(35^\circ)\sin(55^\circ)\sin(65^\circ)\sin(85^\circ) = \frac{1}{64}. \tag{18}$$

Note that all the angles in (18) are odd multiples of $\pi/36$ radians. As is mentioned in [**5**], if r/q is a reduced fraction, the values of $\cos\left(\frac{r\pi}{q}\right)$ can be expressed in terms of ordinary arithmetical operations and finite root extractions on real rational numbers if and only if $q = 2^k p_1 p_2 \cdots p_s$, where k is a nonnegative integer, and $p_1 p_2 \cdots p_s$ is a finite collection (maybe empty) of distinct Fermat primes. This is a consequence of a result by Gauss and Wantzel [**2, 6**]. Even though 3 is a Fermat prime, since $36 = 2^2 \cdot 3^2$, none of the factors in (18) have a closed-form expression in terms of ordinary arithmetical operations and finite root extractions on real rational numbers.

Theorem 3. *For n an odd integer,*

$$\cos(n\theta) = 2^{n-1}\cos(\theta)\prod_{m=1}^{\frac{n-1}{2}}\left[\cos\left(\frac{m}{n}\pi + \theta\right)\cos\left(\frac{m}{n}\pi - \theta\right)\right].$$

Proof. Use Theorem 1 with $\cos x = \sin\left(\frac{\pi}{2} - x\right)$. ∎

Theorem 4. *For n an even integer,*

$$\cos(n\theta) = 2^{n-1}\prod_{m=1}^{\frac{n}{2}}\left[\cos\left(\frac{n+1-2m}{n}\frac{\pi}{2} + \theta\right)\cos\left(\frac{n+1-2m}{n}\frac{\pi}{2} - \theta\right)\right].$$

Proof. Use Theorem 2 with $\cos x = \sin\left(\frac{\pi}{2} - x\right)$. ∎

Remark. Note that this theorem gives an alternative double angle formula for the cosine, because

$$\cos(2\theta) = 2\cos\left(\frac{\pi}{4} - \theta\right)\cos\left(\frac{\pi}{4} + \theta\right).$$

By combining the results from Theorems 1 and 3, we obtain the following result.

Theorem 5. *For n an odd integer,*

$$\tan(n\theta) = \tan(\theta)\prod_{m=1}^{\frac{n-1}{2}}\left[\tan\left(\frac{m}{n}\pi + \theta\right)\tan\left(\frac{m}{n}\pi - \theta\right)\right].$$

Example. With $n = 3$ and $\theta = 20^\circ$, we obtain

$$\tan(60^\circ) = \tan(20^\circ)\tan(40^\circ)\tan(80^\circ).$$

By multiplying both sides of the above equation with $\tan(60^\circ)$, we obtain

$$\tan(20^\circ)\tan(40^\circ)\tan(60^\circ)\tan(80^\circ) = 3.$$

Unfortunately, the formula for $\tan(n\theta)$ when n is even will not look as simple because of the differences in Theorems 2 and 4. The following recurrence relation for $\tan(n\theta)$ can be derived from the sum and difference formula for tangent:

$$\tan(n\theta) = \frac{\tan[(n-1)\theta] + \tan\theta}{1 - \tan[(n-1)\theta]\tan\theta}. \tag{19}$$

Theorem 6. *For n an even integer,*

$$\tan(n\theta) = \frac{\tan(\theta)\prod_{m=1}^{\frac{n-2}{2}}\tan\left(\frac{m\pi}{n-1}+\theta\right)\prod_{m=1}^{\frac{n-2}{2}}\tan\left(\frac{m\pi}{n-1}-\theta\right)+\tan\theta}{1-\tan^2(\theta)\prod_{m=1}^{\frac{n-2}{2}}\tan\left(\frac{m\pi}{n-1}+\theta\right)\prod_{m=1}^{\frac{n-2}{2}}\tan\left(\frac{m\pi}{n-1}-\theta\right)}.$$

Applications

The multiple angle formulas in Theorems 3 and 4 can be used to generate many rational products of integer cosines. The trick is to choose n and θ (in degrees) in such a way that $\cos(n\theta)$ is rational and n divides 90° evenly. For example, if $n = 5$ and $\theta = 12^\circ$ in Theorem 4, we eventually get

$$\cos(12^\circ)\cos(24^\circ)\cos(48^\circ)\cos(84^\circ) = \frac{1}{16}.$$

The multiple angle formulas in Theorems 1 and 2 can also be used to evaluate certain products of sines as is shown in the following corollary.

Corollary. For $n \geq 3$,

$$\prod_{m=1}^{\left\lfloor\frac{n-1}{2}\right\rfloor}\sin\left(m\cdot\frac{\pi}{n}\right) = \frac{\sqrt{n}}{2^{\frac{n-1}{2}}}. \tag{20}$$

Proof. If n is even, then $\left\lfloor\frac{n-1}{2}\right\rfloor = \frac{n}{2} - 1$, and we can use Theorem 2 to obtain the following equation:

$$\frac{\sin(n\theta)}{\sin\theta} = 2^{n-1}\prod_{m=1}^{\frac{n}{2}-1}\sin\left(\frac{m}{n}\pi+\theta\right)\prod_{m=1}^{\frac{n}{2}-1}\sin\left(\frac{m}{n}\pi-\theta\right)\sin\left(\frac{\pi}{2}-\theta\right). \tag{21}$$

Take the limit as θ approaches zero on both sides of the equation to obtain

$$n = 2^{n-1}\prod_{m=1}^{\frac{n}{2}-1}\sin^2\left(\frac{m}{n}\pi\right). \tag{22}$$

Divide both sides of (22) by 2^{n-1}, and take the square root on both sides to obtain the result when n is even. If n is odd, then $\left\lfloor\frac{n-1}{2}\right\rfloor = \frac{n-1}{2}$, and we can use Theorem 1 to obtain the following equation:

$$\frac{\sin(n\theta)}{\sin\theta} = 2^{n-1}\prod_{m=1}^{\frac{n-1}{2}}\sin\left(\frac{m}{n}\pi+\theta\right)\prod_{m=1}^{\frac{n-1}{2}}\sin\left(\frac{m}{n}\pi-\theta\right). \tag{23}$$

We can complete the proof by the exact same limit argument as before. ∎

Example. With $n = 180$, we get $\prod_{m=1}^{89} \sin(m^\circ) = \prod_{m=1}^{89} \cos(m^\circ) = \frac{3\sqrt{10}}{2^{89}}$.

REFERENCES

1. W. A. Beyer, J.D. Louck, D. Zeilberger, A generalization of a curiosity that Feynman remembered all his life, *Math. Mag.* **69**, (1996), 43–44.
2. C. F. Gauss, *Disquititiones Arithmeticae*. Springer-Verlag, New York, 1986.
3. J. Gleick, *Genius: The Life and Science of Richard Feynman.* First ed. Pantheon Books, New York, 1992.
4. S. G. Moreno, E. M. García-Caballero, A geometric proof of Morrie's law, *Amer. Math. Monthly* **122**, (2016), 168.
5. S. G. Moreno, E. M. García-Caballero, A geometric proof of a Morrie-type formula, *Math. Mag.* **90**, (2016), 214–215.
6. M. L. Wantzel, Recherches sur les moyens de reconnaître si un problème de Géométrie peut se résoudre avec la règle et le compas, *J. Math. Pures Appl.* **2** , (1837), 366–372.

Summary. The angle doubling formula $\sin 2\theta = 2\sin\theta\cos\theta$ for the sine function is well known. By replacing the cosine in this formula with $\sin(\pi/2 - \theta)$, we see that $\sin 2\theta$ can be written as the product of two sine functions where the second sine function is obtained from the basic sine function by only using a phase shift of the angle θ and a reflection about the horizontal axis. In this paper, we will show that, for any natural number n, $\sin n\theta$ can be written as the product of n sine functions involving only phase shifts of the angle θ and a possible reflection about the horizontal axis. Similar formulas will be derived for the cosine and tangent functions.

GASTON A. BROUWER (MR Author ID: 703378) received his M.S. in applied mathematics from the Technical University in Delft, the Netherlands. He earned his Ph.D. in applied mathematics from the University of Alabama at Birmingham under the direction of Lex Oversteegen. He is currently an associate professor of mathematics at Middle Georgia State University in Macon, Georgia. His favorite courses to teach are Calculus and Computing for the Mathematical Sciences. In his spare time, he enjoys playing computer games and spending time with his family.

Are You Looking for the Putnam Competition?

Due to a change in the production schedule for the MAGAZINE, problems and solutions for the 2016 William Lowell Putnam Mathematical Competition will appear in the April issue.

Mark Krusemeyer, Gerald Alexanderson, and Leonard Klosinski

A Moment's Thought: Centers of Mass and Combinatorial Identities

DAVID TREEBY
Monash University
Clayton VIC 3800, Australia
david.treeby@monash.edu

Regular readers of the MAGAZINE will be aware that many mathematical identities can be demonstrated by way of a simple diagram without accompanying text. These are called *proofs without words*, and various books are devoted to presenting the best of these, most notably [**1**], [**7**], and [**8**]. For instance, Figure 1 provides a well-known demonstration that

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

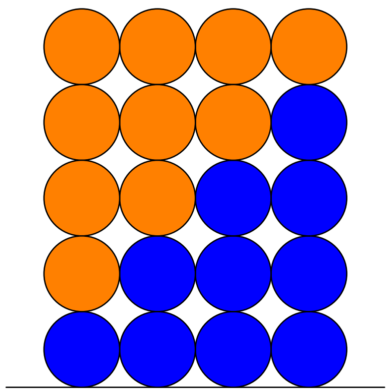


Figure 1

Readers might be surprised to learn that the above diagram can be supplanted by a simpler one, if one assumes a little knowledge of physics. Consider a set of points with coordinates $\{(x_j, y_j)\}_{j=1}^n$. At each of these points we place a weight of mass $\{m_j\}_{j=1}^n$, respectively. The *center of mass* of this configuration has coordinates $(\overline{x}, \overline{y})$ where

$$\overline{x} = \frac{\sum_{j=1}^n m_j x_j}{\sum_{j=1}^n m_j} \text{ and } \overline{y} = \frac{\sum_{j=1}^n y_j m_j}{\sum_{j=1}^n m_j}. \tag{2}$$

We will primarily be concerned with the x-coordinate of the center of mass, which we call the *x-mean*. Although the demonstrations in this article are of a physical nature, all units have been omitted to preserve the clarity of our arguments.

As our first example, we place one unit mass at each of the points 1 through to n on the number line, as shown in Figure 2 below. Since the configuration of masses

Math. Mag. **90** (2017) 19–25. doi:10.4169/math.mag.90.1.19.
MSC: Primary 70C02, Secondary 11A99.

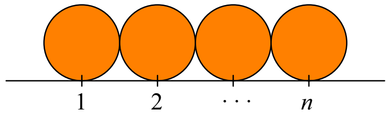


Figure 2 Unit masses arranged in a line

symmetric, the x-mean will be the midpoint $\frac{n+1}{2}$ of the interval from 1 to n. The x-mean can also be found using Equation (2) so that

$$\frac{1 \times 1 + 2 \times 1 + \cdots + n \times 1}{n} = \frac{n+1}{2},$$

giving Equation (1). So we have obtained a physical demonstration of a well-known result by finding the center of mass of a collection of evenly spaced unit masses. We don't provide this demonstration because we regard it as an improvement of the proof implied by Figure 1. However, it provides a simple example of a fruitful approach to deducing various formulae, some familiar and others less-so. This approach is not new; Archimedes, most famously, determined the volume of various figures using similar arguments. This approach is fully expounded in *The Method of Archimedes*, which can be found in [**3**], a collation of his extant works.

Sums of squares

There are well-known proofs without words for the sum of squares of consecutive natural numbers. For instance, a rather elaborate series of pictures in [**5**] provides a proof of the familiar identity,

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}. \tag{3}$$

A more concise proof in [**9**] has just four pictures. Can we provide a one-picture proof if we permit a little physics? A teaching note by Nick Lord in [**6**] provides a clue to one such approach. His argument proceeds as follows.

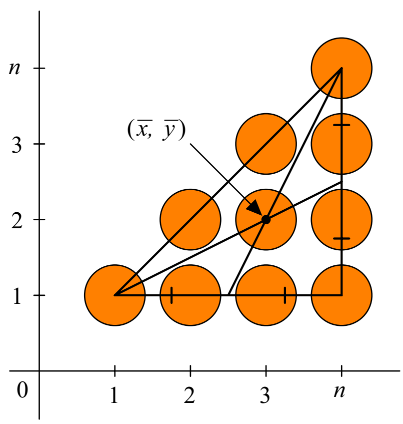


Figure 3 Unit masses arranged in a triangular array

Arrange unit masses in a triangular array as shown above in Figure 3, and let $(\overline{x}, \overline{y})$ be the center of mass of this configuration. Now consider the right-angled triangle indicated on the same figure. Since each vertical and horizontal line of masses has a center of mass on a median of this triangle, point $(\overline{x}, \overline{y})$ corresponds with intersection of its medians. That is, at its center of mass

$$(\overline{x}, \overline{y}) = \left(\frac{2n+1}{3}, \frac{n+2}{3}\right).$$

On the other hand, for each $1 \leq j \leq n$ there are j unit masses located where $x = j$. Therefore the x-mean can also be found by using Equation (2) giving

$$\frac{1 \times 1 + 2 \times 2 + \cdots + n \times n}{1 + 2 + \cdots + n} = \frac{2n+1}{3},$$

from which we deduce Equation (3). Perhaps the proof's only flaw is that it might assume a little *too much* physics, and this makes the explanation subsequent to the diagram crucial. Moreover, Lord's argument depends crucially on two facts. First, to locate $(\overline{x}, \overline{y})$, one must assume that the center of mass of a triangle is located at the intersection of its medians. This is not hard to prove, and we can recommend [**2**] for an entertaining proof that avoids the use of calculus. Second, the proof assumes that if a configuration is divided into parts, and if the center of mass of each of these parts is on the same line, then the center of mass of the entire configuration lies on the same line. This is a believable result, but it does require further thought and substantiation.

We have found a new configuration of masses that obviates the need for explanation and which assumes a little less. We refer the reader to Figure 4 below. Before reading the subsequent explanation you might want to see for yourself how this configuration provides a proof without words of Equation (3).

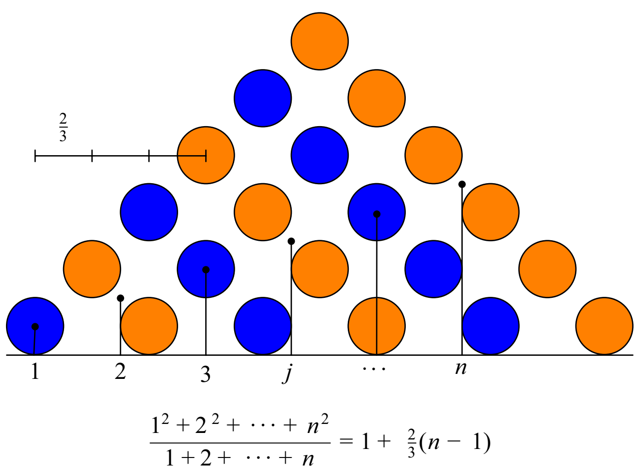


Figure 4

The words. In case the diagram is insufficient, here's the explanation. The configuration is symmetric, so the x-mean coincides with the x-coordinate of the topmost weight,

$$1 + \frac{2}{3}(n-1) = \frac{2n+1}{3}.$$

On the other hand, for each $1 \leq j \leq n$, the jth diagonal comprises j unit masses whose center of mass is located at position $x = j$. Therefore, the x-mean can be found using Equation (2), giving the equation indicated in the figure.

Further generalizations

We have seen how to physically demonstrate well-known formulae by finding the center of mass of symmetric configurations. How might this idea be extended? A couple of obvious things come to mind. Take any set of weights whose masses are $\{m_j\}_{j=0}^n$ where $m_j = m_{n-j}$ for $0 \le j \le n$. These are arranged so that weight j is located at position $x = j$ on the number line as shown in Figure 5 below. Since $m_j = m_{n-j}$ for

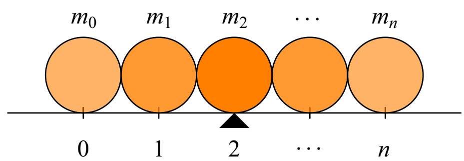


Figure 5 A symmetric sequences of masses

all $0 \le j \le n$, the x-mean of this configuration is located at the midpoint $\frac{n}{2}$ of the interval from 0 to n. On the other hand, we can find the x-mean by using Equation (2) so that

$$\frac{\sum_{j=0}^n j m_j}{\sum_{j=0}^n m_j} = \frac{n}{2}. \tag{4}$$

Example 1. We can apply Equation (4) to the symmetric sequence of binomial coefficients, $\{\binom{n}{j}\}_{j=0}^n$. This gives

$$\frac{\sum_{j=0}^n j\binom{n}{j}}{\sum_{j=0}^n \binom{n}{j}} = \frac{n}{2}$$

yielding

$$\sum_{j=0}^n j\binom{n}{j} = n2^{n-1}.$$

This result is ordinarily obtained by either a combinatorial argument or by differentiating the generating function for the binomial coefficients $(1+x)^n$. Here the physics renders the result a triviality. This technique obviously generalizes to any symmetric and evenly spaced sequence of binomial coefficients, $\{\binom{mn}{mk}\}_{k=0}^n$, the sums of which are considered in [**4**].

More generally, we define a *symmetric sequence* to be any finite sequence $\{a_j\}_{j=0}^n$ such that $a_j = a_{n-j}$ for $0 \le j \le n$. Now consider any two symmetric sequences $\{m_j\}_{j=0}^n$ and $\{a_j\}_{j=0}^{n-1}$. The first of these defines a set of masses while the second defines the distances separating these masses. First place a weight of mass m_0 at the origin. Subsequent to this, we place mass m_j on the x-axis so that its distance from the previous mass is a_{j-1}, as shown in Figure 6. For $1 \le j \le n$, mass j is located at position

$$x_j = \sum_{k=0}^{j-1} a_k. \tag{5}$$

In fact, the formula holds for $j = 0$ so long as we adopt the convention that $\sum_{j=0}^{-1} a_k = 0$.

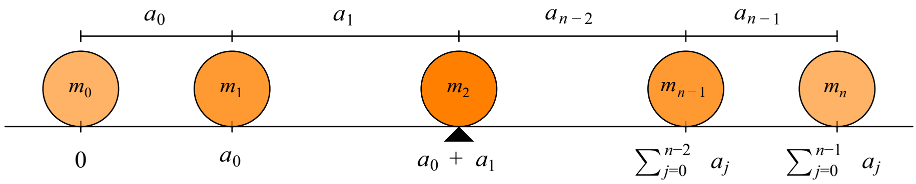


Figure 6 A symmetric configuration of symmetric masses

Since the configuration is symmetric, the x-mean of the configuration is located at its midpoint

$$\overline{x} = \frac{1}{2}\sum_{j=0}^{n-1} a_j.$$

The x-mean can also be found using Equation (2). With the added help of Equation (5) we obtain

$$\frac{\sum_{j=0}^{n} m_j x_j}{\sum_{j=0}^{n} m_j} = \frac{\sum_{j=0}^{n}\sum_{k=0}^{j-1} m_j a_k}{\sum_{j=0}^{n} m_j} = \frac{1}{2}\sum_{j=0}^{n-1} a_j.$$

We obtain the following proposition.

Proposition 1. *If* $\{m_j\}_{j=0}^{n}$ *and* $\{a_j\}_{j=0}^{n-1}$ *are symmetric sequences then*

$$\sum_{j=0}^{n}\sum_{k=0}^{j-1} m_j a_k = \frac{1}{2}\sum_{j=0}^{n} m_j \cdot \sum_{j=0}^{n-1} a_j. \tag{6}$$

Example 2. We can apply Equation (6) to the symmetric sequences of binomial coefficients $\{\binom{n}{k}\}_{k=0}^{n}$ and $\{\binom{n-1}{k}\}_{k=0}^{n-1}$ to obtain another combinatorial identity,

$$\sum_{j=0}^{n}\sum_{k=0}^{j-1}\binom{n}{j}\binom{n-1}{k} = \frac{1}{2}\sum_{j=0}^{n}\binom{n}{j}\cdot\sum_{j=0}^{n-1}\binom{n-1}{j} = \frac{1}{2}\cdot 2^n \cdot 2^{n-1} = 2^{2n-2}.$$

A formula-finding recipe

Note that we have actually developed a recipe for generating formulae from figures with reflective symmetry. We take any such figure and then

1. locate the x-mean along the line of symmetry then,
2. locate the x-mean by dividing the figure into parts then,
3. equate the two results.

To further demonstrate this approach we now consider the center of mass of a rectangular subset of the plane $X \subseteq \mathbb{R}^2$ with uniform density and total area A. If we decompose X into disjoint parts $X_1, \ldots, X_n$, then the center of mass can be found by first finding the center of mass $(\overline{x}_j, \overline{y}_j)$ and area A_j of each part and then computing the weighted averages

$$\overline{x} = \frac{\sum_{j=1}^{n} A_j\overline{x}_j}{A} \quad \text{and} \quad \overline{y} = \frac{\sum_{j=1}^{n} A_j\overline{y}_j}{A}. \tag{7}$$

An application to Fibonacci tilings

We now show how this technique can be used to generate formulae involving Fibonacci numbers. First comes a little revision. The *Fibonacci numbers* are defined by the recurrence relation $F_n = F_{n-1} + F_{n-2}$ where $F_0 = 0$ and $F_1 = 1$. We remind the reader of four basic results:

$$\sum_{j=1}^{n} F_j^2 = F_n F_{n+1}, \qquad (8)$$

$$\sum_{j=1}^{n} F_{2j} = F_{2n+1} - 1, \qquad (10)$$

$$\sum_{j=1}^{n} F_j = F_{n+2} - 1, \qquad (9)$$

$$\sum_{j=1}^{n} F_{2j-1} = F_{2n}. \qquad (11)$$

Each of these can be discovered in Figure 8 below. Specifically, comparing the whole area to the sum of its parts gives (8). To obtain (9) we compare the perimeter of the outer rectangle to the sum of its parts. Comparing the length of the left edge to the sum of its parts gives (10). Finally, comparing the length of the lower edge to the sum of its parts yields Equation (11).

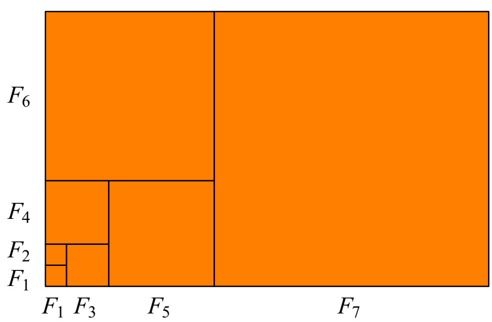


Figure 7 A familiar Fibonacci tiling

We now show how these identities, combined with a center of mass argument, can be used to generate further formulae. Take, for example, a set of rectangles whose lengths are consecutive Fibonacci numbers,

$$F_1 \times 1,\ F_2 \times 1, \ldots, F_n \times 1.$$

Arrange these rectangles along the x-axis as shown in Figure 8. By Equation (9), these combine to make a rectangle of total length

$$F_1 + F_2 + \cdots + F_n = F_{n+2} - 1.$$

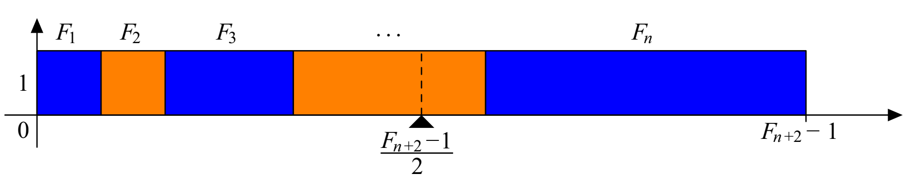


Figure 8

Being symmetric, the x-mean of the combined rectangle is located at its midpoint:

$$\bar{x} = \frac{F_{n+2} - 1}{2}. \qquad (12)$$

The x-mean can also be found as a weighted average of its parts. To this end, we first note that the x-mean of the jth rectangle can be found with the help of Equation (9). This gives

$$\bar{x}_j = \frac{1}{2}\left(\sum_{k=0}^{j-1} F_k + \sum_{k=0}^{j} F_k\right) = \frac{1}{2}\left(F_{j+1} - 1 + F_{j+2} - 2\right) = \frac{1}{2}\left(F_{j+3} - 2\right). \quad (13)$$

We now use Equation (7) to find the x-mean of the combined rectangle. With a further application of Equations (9) and (13) we obtain

$$\bar{x} = \frac{\sum_{j=1}^{n} \bar{x}_j A_j}{A} = \frac{\sum_{j=1}^{n} \frac{1}{2}(F_{j+3} - 2)F_j}{F_{n+2} - 1} = \frac{1 - F_{n+2} + \frac{1}{2}\sum_{j=1}^{n} F_{j+3}F_j}{F_{n+2} - 1}. \quad (14)$$

Finally, equating (12) and (14) yields

$$\sum_{j=1}^{n} F_{j+3}F_j = F_{n+2}^2 - 1. \quad (15)$$

Similar results can be found with sets of rectangles whose lengths are given by $\{F_{2j}\}_{j=1}^{n}$ and $\{F_{2j-1}\}_{j=1}^{n}$. In this case, Equations (10) and (11) are required. In fact, the approach can be profitably applied to more complicated Fibonacci tilings. These are explored in further depth by the author in [**10**].

Acknowledgment With thanks to the support of my supervisors, Burkard Polster, Heiko Dietrich, and Marty Ross.

REFERENCES

1. C. Alsina, R. B. Nelsen, *Math Made Visual: Creating Images for Understanding Mathematics*. MAA Press, Washington, DC, 2006.
2. T. M. Apostol, M. A. Mnatsakanian, Finding centroids the easy way, *New Horizons* **8** (2000) 7–12.
3. Archimedes, *The Works of Archimedes.* Ed. T. L. Heath. Cambridge Library Collection, Cambridge Univ. Press, Cambridge, 2009.
4. A. T. Benjamin, B. Chen, K. Kindred, Sums of evenly spaced binomial coefficients, *Math. Mag.* **83** no. 5 (2010) 370–373.
5. G. Goldoni, A visual proof for the sum of the first n squares and for the sum of the first n factorials of order two, *Math. Intelligencer* **24** no. 4 (2009) 67–70.
6. N. Lord, Evaluating $\sum_1^n r^2$ and $\sum_1^n r(r+1)$ using moments, *Math. Gazette* **98** no. 542 (2014) 346–347.
7. R. B. Nelsen, *Proofs Without Words I: Exercises in Visual Thinking*. MAA Press, Washington, DC, 1993.
8. ———, *Proofs Without Words II: More Exercises in Visual Thinking*. MAA Press, Washington, DC, 2001.
9. M. K. Siu, Proof without words: Sum of squares, *Math. Mag.* **57** no. 2 (1984).
10. D. Treeby, Hidden formulae in Fibonacci tilings, *Fibonacci Quart.* **54** no. 1 (2016) 23–30.

Summary. We provide proofs of well-known formulae using physical arguments. Specifically, we locate the center of mass of a configuration of masses two different ways, and then equate the results. Most notably, we show how this idea leads to a new proof without words for the sum of squares of consecutive natural numbers. We also demonstrate how the method can be profitably applied to certain combinatorial identities, and Fibonacci summations.

DAVID TREEBY (MR Author ID: 972844) has recently begun his Ph.D. in Mathematics at Monash University, Australia. He is concurrently teaching high school Mathematics at Presbyterian Ladies' College and toddler Mathematics to his two young sons.

How to Define a Spiral Tiling?

BERNHARD KLAASSEN
Fraunhofer Institute SCAI
53754 Sankt Augustin, Germany
bernhard.klaassen@scai.fraunhofer.de

Dedicated to Branko Grünbaum on the occasion of his 87th birthday.

What is the main difference between the two arrangements of polygons in Figure ? One answer could be: "The left one has mirror symmetry, the right one doesn't." The answer is correct but ignores another quite obvious effect: The right figure has a spiral structure, the left one doesn't.

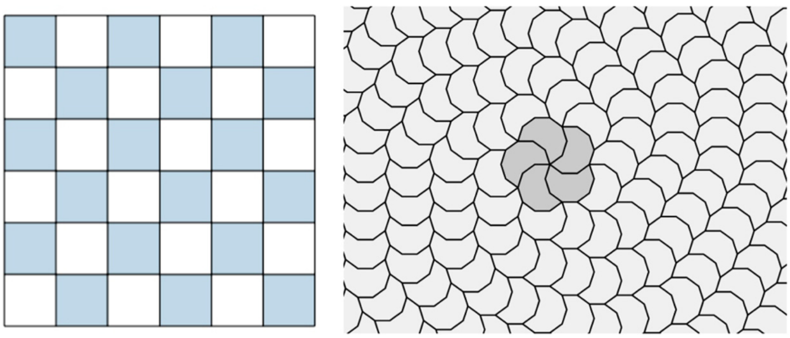


Figure 1 How can these two arrangements be distinguished? (Source: [13]).

It may be easy to recognize the spiral structure, but how does one define it mathematically? For the reader it might be surprising that no such definition has been proposed. Later we will see that the task is more complex than one may expect at first glance. In the tilings "bible" [**6**] the authors admit that "to some extent, at least, the spiral effect is psychological." Nevertheless they declare it as an open problem to define this effect mathematically.

In this paper we investigate plane tilings built in spiral form and will try to fill this gap. However, before going on to answer the above question, we have to say what a tiling of the plane is in general. Therefore, we cite again the "grandmasters" of tilings, Grünbaum and Shephard [**5**]:

> "A **tiling** of the plane is a family of sets—called **tiles**—that cover the plane without gaps or overlaps. ("Without overlaps" means that the intersection of any two of the sets has measure (area) zero.)"

Since it is no secret that an arbitrary subset of the plane can easily turn out to be something really strange, one should restrict oneself to those sets (i.e., tiles) that are topologically equivalent to a closed disk. (If the reader is familiar with the famous drawings of M. C. Escher, there are lots of plane tilings to be found.) To keep it even simpler, in all our examples the tiles will be polygons. Many (but not all) of the polygonal plane tilings have the property that all tiles are mutually congruent, as in Figure . Such tilings are called **monohedral**. However, polygonal tilings are not always simple as some of the following figures will demonstrate.

Math. Mag. **90** (2017) 26–38. `doi:10.4169/math.mag.90.1.26.` © Mathematical Association of America
MSC: Primary 52C20

On the other hand, in left tiling (of Figure) nobody will see a spiral pattern. We will keep such a simple chessboard tiling in mind as a test case, since any definition for spiral tilings should be checked to see if it excludes such cases.

Definitions

There are only few definitions of spiral tilings in the literature, all of which are specialized to certain areas of interest. To our knowledge, the earliest one is given by Grünbaum and Shephard [**6**] (in the exercise section of Chapter 9.5) and it is restricted to monohedral plane tilings. In [**2**], a definition for "similarity tilings" with growing size of tiles is given. Here we give an alternative approach.

In Figure (right) you can find long sequences of tiles starting at one of the dark gray "center tiles" and spinning around in spiral-like manner. You can count five of these sequences that will be called *spiral arms* throughout this paper. We will see that the following definition is not directly applicable for one-armed spirals. So, for the start, the number of arms should be larger than one. *Singular points*, i.e., points where an infinite number of tiles are clustered, will be considered later, as well as one-armed spirals. Observe that we do not restrict this investigation to monohedral tilings only.

Definition L (spiral-like). A partition of a plane tiling into more than one separate classes (called *arms* here) is defined as a *spiral-like partition* or *L-partition* under the following conditions. (The plane is identified with the complex plane $\mathbb{C}$ where the origin is represented by a selected point of the tiling.)

L1: For each arm A (as a union of tiles from one class) there exists a curve $\theta : \mathbb{R}_0^+ \to A \subset \mathbb{C}$ with $\theta(t) = r(t)\exp(i\varphi(t))$ called a *thread*, where both r and φ are continuous and unbounded and φ is monotone. Curve θ does not meet or cross itself or any thread from another arm of the tiling.

L2: For each tile T in A the intersection of the interior of T with the image of θ is non-empty and connected.

Remark: A plane tiling (without singular points) with an L-partition shall be called an *L-tiling* or a *spiral-like tiling*.

The properties of the threads $\theta(t)$ from L1 and L2 may be summarized in one sentence: they must be simple unbounded separate curves meeting each tile exactly once while spinning infinitely often around a point. They can be regarded as threads of infinitely long pearl necklaces running through all tiles of each spiral arm. See Figure 2 for a simple tiling with two threads shown as lines within the spiral arms. We see that two arms can be coupled in a curved segment. Where two arms meet in such a way, there is no "natural" separation of them. By choosing the starting points for the threads, a partition into separated arms is being made (shown by the coloring). One should keep in mind that the choice of the partition cannot be unique here.

Observe that the checker board from Figure (partitioned into "white fields" and "gray fields") would also fulfill definition L. (The reader might try to find a suitable thread before looking at Figure 3). Hence, though it seems to be necessary for a spiral tiling that the above definition is satisfied, it is not enough. What is the property of Figure 2 to be easily recognized as a spiral in contrast to the above checkerboard tiling?

A closer inspection of Figure 2 shows that all triangles within an arm meet at the longer edges while the boundaries of the arms mainly consist of the shorter ones. In the curved segment only, where two arms are coupled, the long edge is needed in the boundary of arms. We can reformulate this observation: All neighboring tiles within an arm are connected in a "typical" way that can be distinguished from the situation

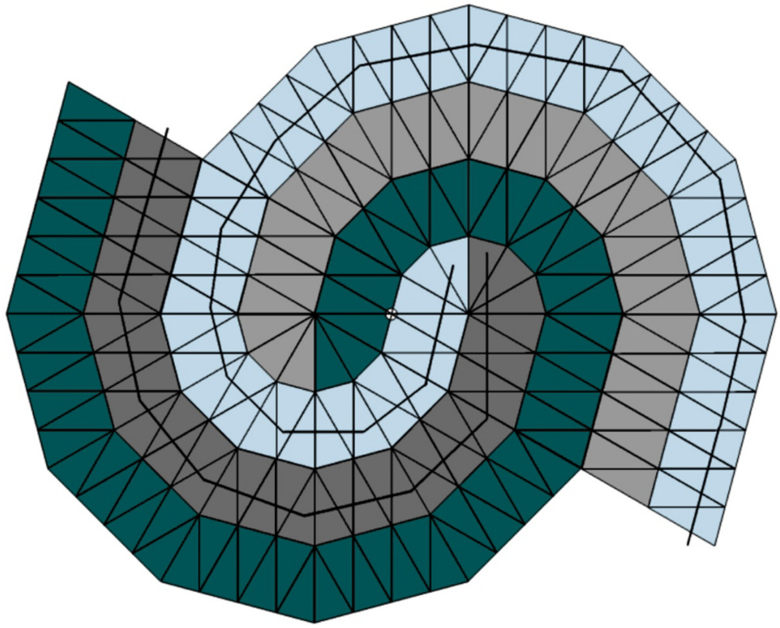


Figure 2 A four-arms tiling from [**1**] with two threads displayed.

when two tiles meet at the boundary between two arms. The only exception may occur at the beginning of an arm.

Definition S (for multi-armed spirals). A partition of a plane tiling is defined as a *spiral partition* or *S-partition* under the following conditions.

S1: It must be an L-partition (see definition L).

S2: If any two tiles $T_1, T_2 \in A$ are direct neighbors and can be mapped (as a pair) by rotation and translation onto another pair of tiles, these must also be direct neighbors within an arm. This rule can be ignored if the image pair contains the beginning of an arm, i.e., contains $\theta(0)$. ($T_1, T_2 \in A$ are called *direct neighbors* if $T_1 \cap T_2$ is cut by the thread of A or contains more than a finite number of points.)

Remark: A plane tiling (without singular points) which allows an S-partition shall be called an *S-tiling* or a *spiral tiling*.

It should be clear that the term L-tiling should only be used for those tilings which satisfy definition L but *not* S. Condition S2 was introduced mainly to distinguish an obvious spiral partition from unwanted cases as shown in the chessboard tiling. So, let us take a look at Figure 3 to see if it satisfies S2 or not.

We take the tiles (2 and 3), which are direct neighbors (in the meaning of S2), and map them to the pair (2 and 9). Geometrically the mapping is possible, but the image pair is not connected by the thread (dashed line) and the intersection is one point only. Hence, S2 is not satisfied, which means that the chessboard partition is only spiral-like according to definition L. The latter is illustrated by the thread, which is able to connect all colored fields if suitably prolongated. An analogous second thread can be drawn for the white fields, so we have two arms here.

We have seen above that Figure 2 shows a tiling with a partition satisfying definition S, so we can call it a spiral tiling with four arms. In the next section we will give a series of examples for known tilings with spiral character that allow partitions to fulfill all conditions of definition S. Moreover, many examples for threads will be given, all

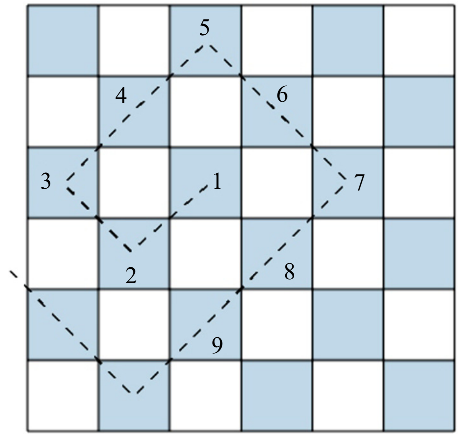


Figure 3 The checker tiling does not satisfy condition S2.

of which are represented by polygonal lines. For reasons of elementary geometry it can be seen that the condition of monotonicity (of φ) only needs to be checked at the corner points of such lines simplifying the analysis.

Examples

We start our tour through typical examples with the historically first occurrence of a spiral tiling, the famous Voderberg tiling with two arms (see Figure 4). Observe that all these complicated tiles have the same shape. To highlight the structure of this tiling, neighboring tiles were colored in different color shades.

Since the full tiling has 2-fold rotational symmetry, it is sufficient to draw function θ from definition L just for one arm. The origin of the complex plane, which is in this case also the center of symmetry, is marked by a dot. S2 can be checked easily, since all direct neighbors are joined at the longer edges while adjacent pairs from different arms are sharing a short edge only. (The reason for the quite sophisticated form of the tiles is not to create a spiral form. The Voderberg tiles have other surprising properties not to be discussed here in full detail.)

The monotonicity of φ and the other properties from the definition are obvious in this case, as, by construction, the tiles have a canonical order within the arm.

The following example shown in Figure 5 (again from [**1**]) shows that the threads cannot always run completely within the interior of the arms. Where the arm's boundary meets itself, there is only one single point through which the thread must lead. This was also the case in Figure 3 but here we have a spiral satisfying all conditions of definition S. Observe that such a partition also could be applied to the tiling of Figure 2 if we identify each rhomb with two triangles. However, the investigation of this tiling can be carried on a step further. One should keep in mind that our definitions always need a tiling *plus* a partitioning. So, the question is here: How will other partitions deal with definition L or S? Take a look at Figure 6.

It is quite obvious that here all conditions for definition L are satisfied. But S2 is violated, since the adjacent pairs with both dark and light color have the same shape as several pairs within the arms. So, this partition is only spiral-like but not an S-partition. As Figures 5 and 6 demonstrate, it cannot be expected that the partitioning of a tiling into spiral arms could always be done in a unique way. Our definition can only check *if* a given partition is of type L or S, or neither nor. But it cannot directly help to find a suitable structure within the tiling.

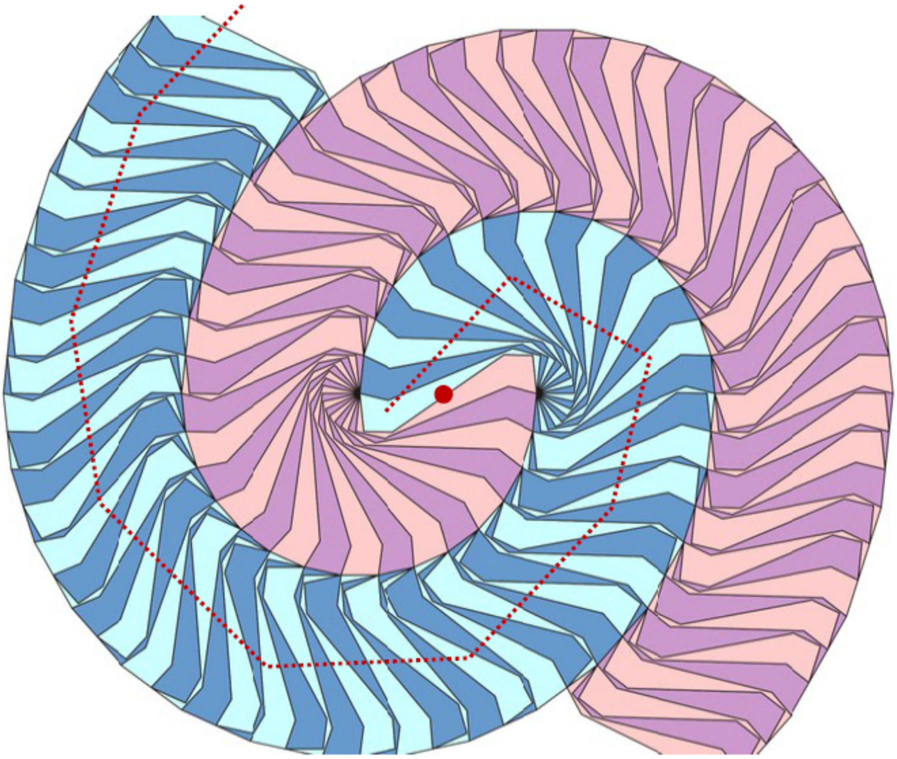


Figure 4 Voderberg's spiral [**16**] with thread from definition S (dotted line).

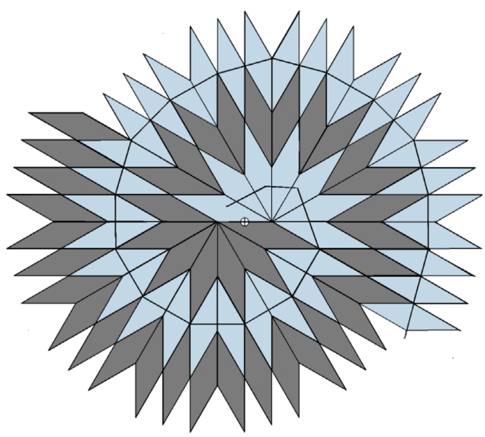


Figure 5 A spiral from [**1**].

Figure 6 leads to the observation that function r—the radius in polar coordinates of $\theta(t)$—cannot always be monotone, due to the star-shape of this tiling. This motivates the decision to demand only φ to be monotone in definition L and S.

Since the first definition by Grünbaum and Shephard was restricted to monohedral tilings, it is important that definition S is suitable for other types also. We will test it in Figure 7 with a ten-armed pentagon-rhombus tiling from [**9**], which has a connection back to the German renaissance painter Albrecht Dürer (who, incidentally, wrote the first geometry book in the German language in 1525).

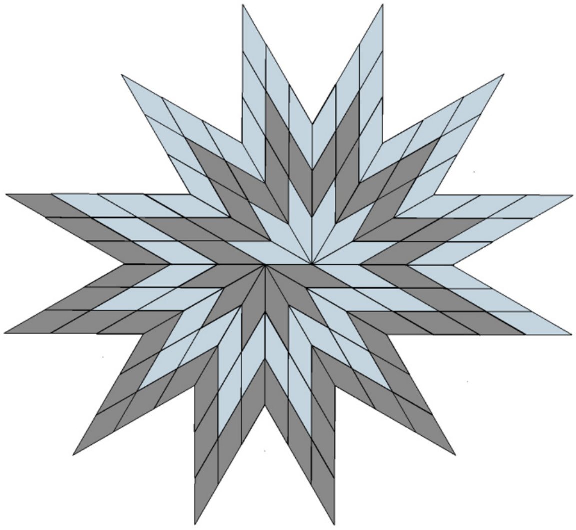


Figure 6 Same tiling as in Figure 5 but with different partitioning.

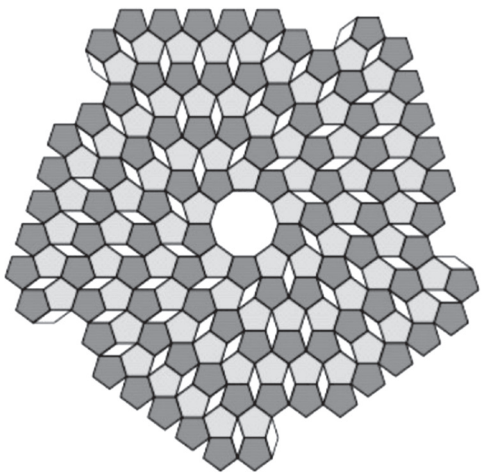


Figure 7 A tiling consisting of rhombuses, pentagons, and one decagon [**9**].

First, one has to decide about the partitioning. (The coloring in Figure 7 is from the original drawing and therefore cannot represent the partition needed for definition S.) The pentagons are obviously grouped into ten parts. The rhombi are not clearly partitioned and each rhombus has common edges with both dark and light gray pentagons, which creates problems for condition S2. In some sense the rhombi—as they are positioned—disturb condition S2. The other conditions can easily be fulfilled but, due to the rhombi, this tiling remains an L-tiling. However, one should not draw the conclusion that tilings which are not monohedral must have problems with S2. Imagine that all rhombi were cut along the short diagonal into two triangles and the decagon (like a tart) into ten triangles. In such a tiling the partitioning would be straightforward and also condition S2 would easily be fulfilled.

The following examples are again monohedral.

In [**11**] we constructed spiral tilings consisting of convex pentagons with an arbitrary number of arms and for each possible rotational symmetry. They all satisfy definition S as Figure 8 illustrates for two different partitions with seven arms.

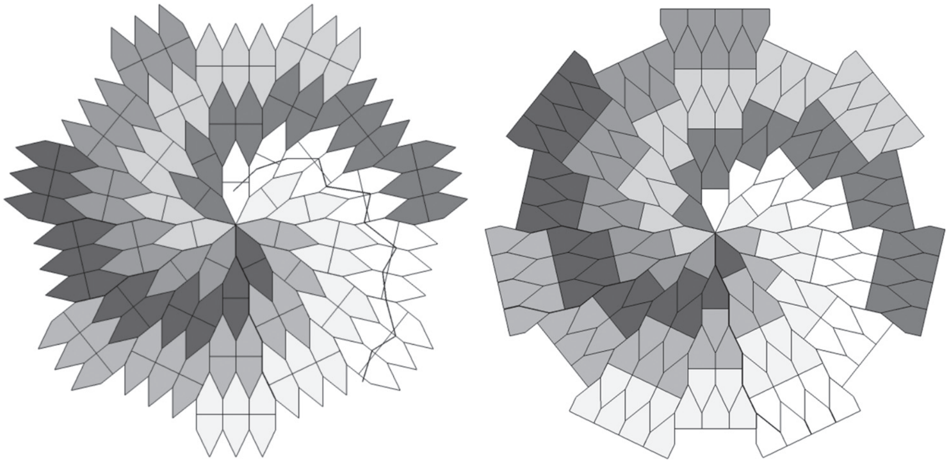


Figure 8 A tiling with 7-fold symmetry consisting of pentagons [**11**].

All other spirals with an odd number of arms larger than 5 (e.g., in [**13**]) had been constructed with nonconvex tiles as can be seen in the two spirals in Figure 9. Both are monohedral but have no overall symmetry, so we have to draw all threads. Here it is not so easy to count the number of arms. It is easier to count the number of beginning threads near the center point. In both cases we have five threads.

Again (as in Figure 2) we have some freedom to choose the partition in the case of coupled arms in curved segments near the center. With the help of the chosen threads it should be clear how the arms are separated. In all cases a partition was found in a way that threads with a monotone φ could be constructed. Condition S2 can easily be checked since the tile pairs within the arms are put together in a very typical manner.

In discussions about spirals one could ask "Why define it? You will know it when you see it." This is not always the case, as we can see in the following examples. Since the creations of Voderberg or Penrose or M. C. Escher, tilings are not always directly comprehensible. From [**1**] and [**6**]—originally from [**12**]—we know that rather strange S- and Z-shaped tiles (see Figure 10) are able to tile the plane in a spiral structure which is not easily recognized without a suitable partition.

In Figure 10 (left) the coloring was chosen to illustrate the construction of the tiling and not the spirals. The threads help the viewer to make sure that there are four

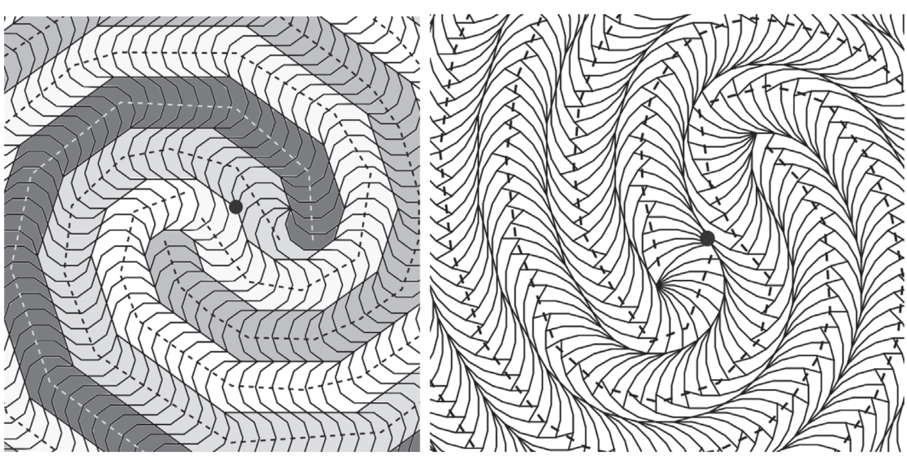


Figure 9 Two spirals from [**13**].

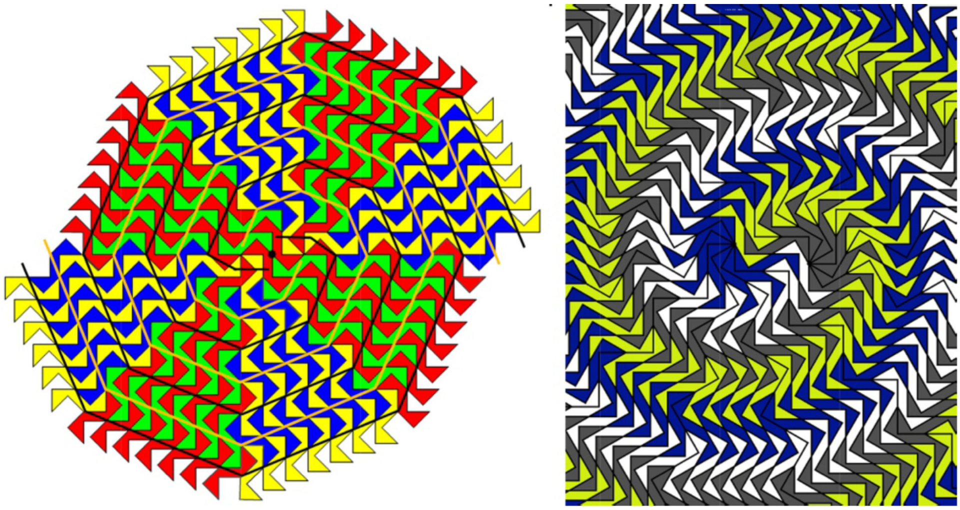


Figure 10 A tiling built of S- and Z-shaped tiles (taken from [**1**] and [**15**]).

connected lines of tiles from the origin outwards. The right half of Figure 10 shows a variant of the same tiling with a colored partitioning that highlights the spiral effect.

So, definition L and S might not only help to decide if a certain pattern should be called a spiral tiling, but also serve as analytical tools making the structure of a tiling more visible. In Figure 10 (and also in Figure 5) it is neither necessary nor possible for the image of θ to run completely in the interior of the arm. Some tiles might be touched by two or three threads—but this is also not forbidden. In all cases it is clear which is the arm's thread and which is a neighboring one, since neighbor threads can only touch (but not cross) the arm. However S2 can be checked easily: Each arm consists of either S- or Z-tiles, which can form direct neighbors. And those pairs, which are sharing an edge of the arm's boundary, are mixed S-Z-pairs, so there is no chance for a conflict with S2.

It is an interesting question whether there are tilings with partitions, which at first glance seem to have spiral character but do not fulfill definition S. Such an example exists and it is a famous one: the Hirschhorn tiling [**7**] (Figure 11, left). Outside of the inner circle of 18 tiles there are the next 18 tiles forming a ring in C18 symmetry.

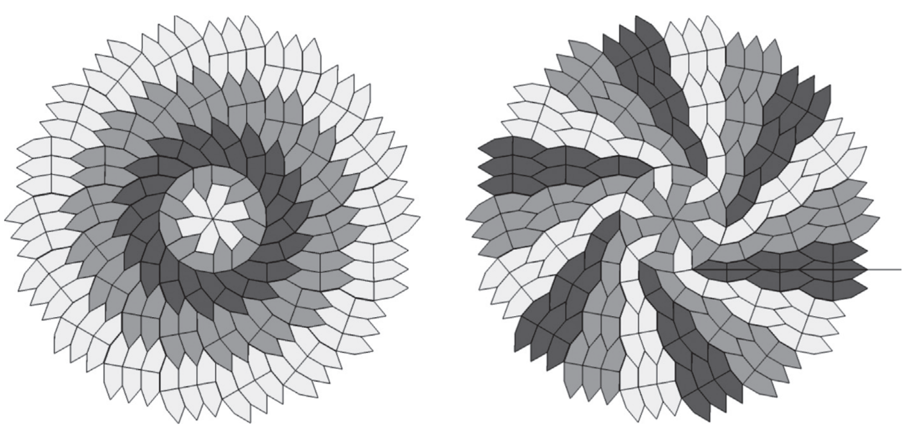


Figure 11 The Hischhorn tiling (spiral tiling or not?).

Each of these tiles can be regarded as the root of a tree-like structure with a growing size of rows. Each tree can be chosen as one part to get a partition into 18 classes (see Figure 11, right) and seems at first glance to bend itself around the center like a spiral.

But on closer inspection one realizes that this structure will always follow a straight line shown in Figure 11 (right). Such an “arm” will never turn around in a spiral-like manner as demanded in definition L or S. With this partition we have a case where S2 is satisfied but not S1, hence neither L nor S is fulfilled. One could also imagine a structure with an unlimited number of arms, since each of the tree-like parts in Figure 11 can be split into a series of subarms, one of which starting at each “row” of the trees. Anyway, all those arms will never bend around the center.

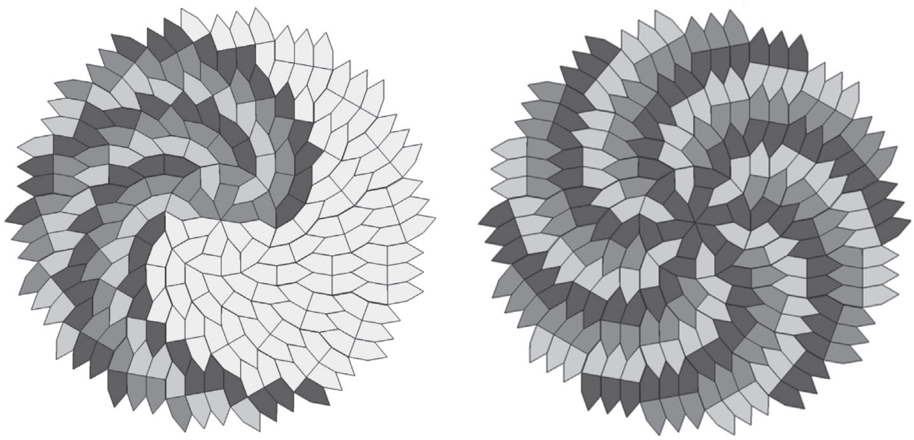


Figure 12 Same tiling as in Figure 11 but with different partitions.

However, we learned from Figures 5 and 6 that in some cases several partitions should be tested, and we should inspect Figure 12. Here we can see on the left side that there is a partitioning with all properties to fulfill definition L, but S2 cannot be satisfied here. We left one half unpartitioned to show that it is more or less a hidden structure within the tiling, which has not much to do with the tiling’s original construction, being better illustrated by Figure 11 (right).

Nevertheless, sometimes a shift in direction can help: Following a suggestion of Branko Grünbaum [**4**] we can show a nice spiral partitioning in Figure 12 (right). It satisfies all conditions of definition S. So, in some sense the Hischhorn tiling represents an optical illusion, since the first impression is an anticlockwise spiral effect, while the “true” spiral is running in clockwise direction.

We can summarize the example section by stating that all multi-armed spiral tilings from [**1**], [**3**], [**6**], [**11**], and [**13**] (altogether more than 30 cases) can be partitioned to fulfill the conditions of definition S. Figure 7 is the only L-tiling (not satisfying S) found in the literature, together with the spirals from regular tilings mentioned in the exercise section of Chapter 9.5 in [**6**]. The Hirschhorn tiling is a special case: The canonical partition into 18 tree-like arms does not form spirals. The structure allowing an S-partition could be revealed but is not to be seen at first glance.

The reader might try to apply the definitions to those spirals from the given sources that could not be shown here. (At the tilingsearch.org website [**15**] there is a very nice collection to be found in excellent graphical quality.) At this site a very special case is tiling 129/F19. Although this tiling has in some sense a spiral character, it is not possible to find suitable threads to satisfy definition L. So, this tiling (with the given partition) is neither spiral nor spiral-like.

The one-armed case

After this series of examples the reader may have a better chance to see that definition S has its benefits. However, in the case of one single arm, we have to formulate a modified version.

Definition O (for one-armed spirals). A tiling of the plane without singular points is called a *spiral tiling with one arm* under the following conditions (The plane is identified with the complex plane $\mathbb{C}$ where the origin is represented by a selected point of the tiling.)

O1: There exists a curve $b : \mathbb{R}_0^+ \to \mathbb{C}$ with $b(t) = r(t)\exp(i\varphi(t))$ called *spiral boundary*, where both r and φ are continuous and unbounded. Curve b does not meet or cross itself and runs completely on boundaries of tiles

O2: If T_1, T_2 are direct neighbors and can be mapped (as a pair) by rotation and translation onto another pair, these tiles must also be direct neighbors. This rule can be ignored if the image pair lies at the beginning of the boundary (i.e., contains $b(0)$). (*Direct neighbors* means here that $T_1 \cap T_2$ contains more than a finite number of points but not from the spiral boundary.)

In Figure 13 the marked edges are shown forming the spiral boundary. In all known one-armed spirals (e.g., in [**6**] or [**15**]) we can find such a structure. It is interesting that the same tiling could also be interpreted as a two-armed spiral, shown on the right half of Figure 13. With this partition it satisfies definition S.

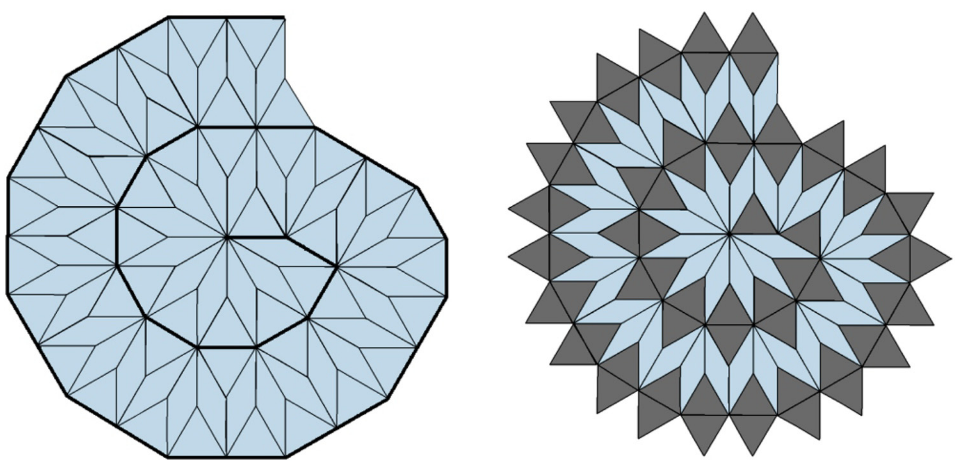


Figure 13 One-armed spiral with marked spiral boundary (left side).

The singular point case

Up to this point, the nonspecialist reader might think that in all tilings the size of a tile should be limited. This was the case in our examples but is not a rule in general. For the following tilings (known as "locally infinite") there will be no upper or lower limit for the size of tiles, and arbitrarily small tiles will be clustered at *singular points*. In this context it is required to say more precisely what is meant by "without gaps" in the cited definition of the term *tiling* (see first section). It is possible that the singular point is not part of any tile. So, with P_s being the set of singular points and T the union of all tiles, we can say that "without gaps" means here that the entire plane should be $T \cup P_s$ which is $\overline{T}$, since singular points are limit points by definition.

For those tilings with exactly one singular point (some examples are given in [**6**], Chapter 5.1 or 10.1 under the keyword *similarity tilings*) it is possible to modify our definition slightly in the following way.

Definition P (for tilings with one singular point). A tiling of the Euclidean plane with exactly one singular point together with a partition is called a *spiral tiling* under the following conditions. (The plane is identified with $\mathbb{C}$ where the origin is represented by the singular point.)

P1: The partition fulfills L1 and L2 but with $\theta : \mathbb{R} \to A \subset \mathbb{C}$ and with φ being unbounded in both directions.

P2: If any two tiles $T_1, T_2 \in A$ are direct neighbors and can be mapped (as a pair) by rotation, translation and scaling onto another pair, these tiles must also be direct neighbors within an arm. (*Direct neighbors* means here that $T_1 \cap T_2$ is cut by the thread of A.)

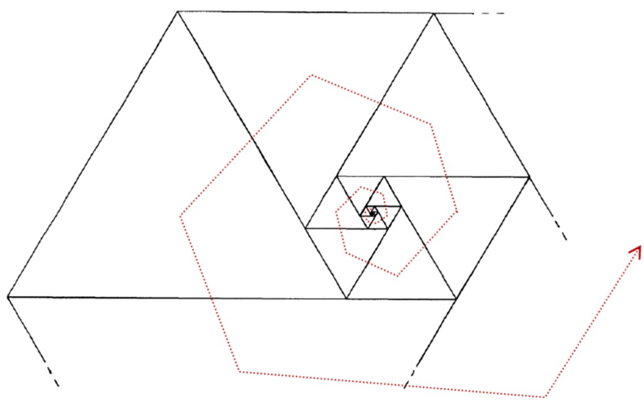


Figure 14 A spiral tiling with equilateral triangles of different size [**10**].

An example from [**10**] illustrates a locally infinite tiling with one singular point. (See Figure 14). The dotted line is the image of θ. There are several tilings of this kind satisfying definition P ([**1**] or [**2**]). Also multi-armed spirals exist with one singular point (see also [**14**] for many more examples to which definition P can be applied). In the right half of Figure 15 we can see a three-armed spiral within a rectangular structure. At first glance this seems to be surprising, but the construction is quite straightforward (all three arms are scale symmetric to each other). Here again it can be demonstrated that the partitioning into spiral arms is in many cases not unique. The left side of Figure 15 shows exactly the same tiling with one thread, illustrating that the tiling can be treated as both three- and one-armed spiral.

Also spirals with more than one singular point are possible, but then our definition becomes even longer and less readable. So, we restrict ourselves to the case with one singular point only, which is also the standard case in the literature.

Discussion

The definition for spiral tilings given in the first sections of this paper is useful for multi-armed partitions of locally finite tilings. Here we distinguish between *spiral-like* (definition L) and *spiral* tilings (definition S). For the latter we can say that it is

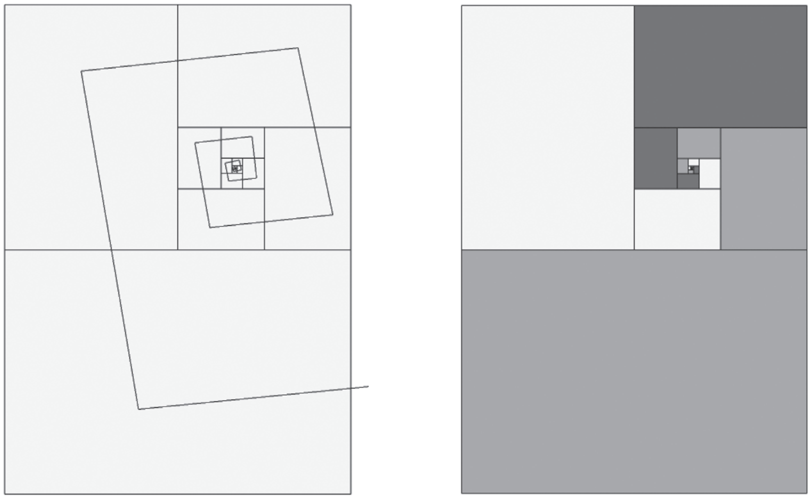


Figure 15 A logarithmic spiral tiling with one arm (left) and three arms (right).

not only the partition which introduces the spiral form. The tiling itself must have the spiral structure to admit a spiral partitioning. (We conjecture that an algorithm could be found based on definition S, which automatically reconstructs the structure of a spiral tiling without a given partition. The resulting partition might not be unique in all cases but should always be spiral.) There exist some unwanted cases mainly from regular tilings, which all can be classified to be only spiral-like. Here it remains as an open question whether every monohedral (or every periodic) tiling admits an L-partition.

A modified version (definition O) deals with one-armed spirals and a special variant (definition P) is intended for locally infinite tilings with one singular point and an arbitrary number of arms. Hence, for each known case of tilings with spiral character, an appropriate definition from this paper has been shown to be applicable.

At first glance one might get the impression that for the definitions a lot of properties have to be checked, but it could be demonstrated that it is sufficient just to draw a polygonal line for each spiral arm and to inspect the typical neighboring tiles within the arms and at boundaries. Then it can easily be determined whether a tiling and its partitioning fulfills the conditions or not.

The strength of the definitions is that not only can one check whether a tiling forms a spiral, but that they provide a tool to highlight the structure of complicated tilings.

Acknowledgment I am indebted to Branko Grünbaum for fruitful discussions as well as to the anonymous reviewers and the editor of the MAGAZINE and to Dina Hess and Sebastian Quack, who helped a lot to improve this paper. I would also like to thank Steven Dutch (Univ. of Wisconsin-Green Bay), Craig Kaplan (Univ. of Waterloo, Cda.) and, last but not least, Brian A. Wichmann [**15**] who kindly gave permission to use spiral drawings from their websites (see references below).

REFERENCES

1. S. Dutch, Radial and Spiral Tilings, https://www.uwgb.edu/dutchs/symmetry/rad-spir.htm, Dec. 15th 2015.
2. D. Eppstein, Spiral Tilings, http://www.ics.uci.edu/~eppstein/junkyard/spiraltile/, Dec. 15th 2015.
3. P. Gailiunas, Spiral Tilings, http://vismath4.tripod.com/gal/, Dec. 15th 2015.
4. B. Grünbaum, private communication, June 2016.
5. B. Grünbaum, G. C. Shephard, Tilings by regular polygons, *Math. Mag.* **50** no. 5 (1977) 227–247.
6. ———, *Tilings and Patterns*. W.H. Freeman and Co., New York, 1987.

7. M. D. Hirschhorn, D. C. Hunt, Equilateral convex pentagons which tile the plane, *J. Comb. Theory Ser. A* **39** (1985) 1–18.
8. https://de.wikipedia.org/wiki/Parkettierung.
9. C. Kaplan, The Trouble with Five, https://plus.maths.org/content/trouble-five, Dec. 15th 2015.
10. B. Klaassen, Infinite perfekte Dreieckszerlegungen auch für gleichseitige Dreiecke, *Elem. Math.* **50** (1995) 116–121.
11. ———, Rotationally symmetric tilings with convex pentagons and hexagons, *Elem. Math.* **71** (2016) 137–144.
12. E. A. Lord, C. B. Wilson, *The Mathematical Description of Shape and Form*. Ellis Horwood, Chichester, 1984.
13. D. L. Stock, B. A. Wichmann, Odd spiral tilings, *Math. Mag.* **73** no. 5 (2000) 339–346.
14. T. Sushida, A. Hizume, Y. Yamagishi, Triangular spiral tilings, *J. Phys. A: Math. Theor.* **45** (2012) 235203.
15. B. A. Wichmann, http://www.tilingsearch.org/tree/spiral.htm.
16. Wikipedia contributors, *Wikipedia, The Free Encyclopedia*, https://commons.wikimedia.org/wiki/File:Voderberg-1.png (drawing by T. E. Drozinski, 2008).

Summary. A precise mathematical definition is given for spiral plane tilings. It is not restricted to monohedral tilings and is tested on a series of examples from the literature. Unwanted cases from regular tilings can be excluded. In case of one single arm a modified definition can be applied. Also the special case of locally infinite tilings with one singular point can be treated with any number of spiral arms. The question whether such a definition in mathematical terms could be given was posed by Grünbaum and Shephard in the late 1970s.

BERNHARD KLAASSEN (MR Author ID: 365627) studied mathematics and computer science in Munich and Bonn and received a Ph.D. from the University of Duisburg-Essen. Currently he is a senior scientist at the Fraunhofer Institute SCAI in Sankt Augustin, Germany.

PINEMI PUZZLE

7		7	5		6		6		
		11				6		8	5
	10			11	7		6		
4	10			11			6	6	
	6		11						9
6		10		7	7		9		
			6		8	9		12	9
	9	6	5	8					
		5		9			12		
	5				7	8		6	

How to play. Place one jamb (|), two jambs (||) and three jambs (|||) in empty cells, where numbers indicate how many of jambs in the surrounding cells (including diagonally adjacent cells), and each row (column) has 10 jambs. There cannot be a jamb in any cell that contains a number.

The solution is on page 57.

—contributed by Lai Van Du Thinh, Munigeni Studio Game,
Nha Trang, Khanh Hoa, Vietnam; fibona2cis@gmail.com

Dihedral Symmetry in Kaprekar's Problem

MANUEL R. F. MOREIRA
Faculdade de Ciências da Universidade do Porto
Porto, Portugal
mfmorei@fc.up.pt

Take an integer u greater than zero, less than ten thousand, and not a multiple of 1111. If necessary complete with leading zeros to obtain a four digit integer. Apply the following operation to the integer u: Let N be the integer obtained from u by placing its four digits in descending order and n the integer with the same four digits in ascending order. Then consider the difference $N - n$. Apply the operation again taking this difference as the new starting number. For example, if $u = 1960$ then $N = 9610$, $n = 0169$ and the difference is 9441. Take this difference as the new number to which apply the operation and get 7992, since $9441 - 1449 = 7992$. Iterate the process and get successively 7173, 6354, 3087, 8352, 6174. At this stage something different happens. Applying the operation to 6174 gives 6174 since $7641 - 1467 = 6174$. This number 6174 is a fixed point of the operation.

We knew of course that we would end up with a fixed point or we would get caught up in a loop, since that is necessarily the fate for a set of the type $\{u, f(u), f^2(u), f^3(u), \ldots\}$, where f is a map of a finite set X into itself and u is an element of X. But the interesting thing about the particular set X and map f that we are considering is that no matter what element u we start with and apply the described process we will end up with the fixed point 6174. Moreover, also independently of the initial integer u in X, $f^7(u) = 6174$. The number of our example $u = 1960$ is as tough as there is. No number resists more than seven operations before surrendering to the fixed point 6174. Now we easily can produce a computer program that will show experimentally these assertions. But we choose another option. In the following sections we use symmetries of the operation, and keep the computations to a minimum, to get the announced results and a little bit more. Before we start, we should acknowledge that the math problem we are dealing with was discovered more than sixty years ago by the Indian mathematician Kaprekar. And for that reason 6174 is called *Kaprekar's constant*.

Preliminaries

Let X be the set of natural integers less than 10000 without the nine integers represented by four equal digits. An element of X can be written as $1000a + 100b + 10c + d$, where a, b, c, and d are integers between zero and nine, not all equal. Given such an element u let $e \geq f \geq g \geq h$ be the digits a, b, c, and d in descending order of magnitude. Define a map $f: X \longrightarrow X$ by $f(u) = 1000e + 100f + 10g + h - 1000h - 100g - 10f - e$. By regrouping we get $f(u) = 999(e - h) + 90(f - g)$. These integers are in X since they are less than 10000 and not a multiple of 1111. Because $f(u)$ is always a multiple of 9 and the only multiple of 1111 less than 10,000 that is also a multiple of 9 is 9999, then the map f is injective. If we consider the set Y whose elements are the integers that can be put in the form of $999m + 90n$, where $0 \leq m \leq 9$, $0 \leq n \leq 9$, and m and n not both zero, we have that $f(X) \subset Y \subset X$. In particular f maps Y into Y. We divide the set Y into two disjoint sets: (a) the set

Math. Mag. **90** (2017) 39–47. doi:10.4169/math.mag.90.1.39.
MSC: Primary 11B83, Secondary 00A08.

L of elements of the form $999m$, $1 \le m \le 9$, or of the form $90n$, $1 \le n \le 9$. Written in the usual way, these are the elements $1000(m-1)+900+90+(10-m)$ or $100(n-1)+10(10-n)$; (b) the set Q of elements of the form $999m+90n$, $1 \le m \le 9$, $1 \le n \le 9$. Written in the usual way, we get the integers $1000m + 100(n-1) + 10(9-n) + (10-m)$.

Instead of using the integers in Y we can use the pairs (m, n) associated with the elements of Y as described, since the correspondence between elements of Y and pairs (m, n) is clearly a bijection. The elements in L correspond to pairs $(m, 0)$, $1 \le m \le 9$ or pairs $(0, n)$, $1 \le n \le 9$, and the elements in Q correspond to pairs (m, n), $1 \le m \le 9$, $1 \le n \le 9$. Better yet for our purposes, we consider pairs (x, y) obtained from the previous pairs (m, n) by the relations $x = m - 5$, $y = n - 5$. We will be considering the set Y_1 of pairs (x, y) where x and y are integers and $-5 \le x \le 4$, $-5 \le y \le 4$, $(x, y) \neq (-5, -5)$. Similar to what we did with Y, we consider two disjoint subsets of Y_1 defined by

(a) $M = \{(x, -5); -4 \le x \le 4 \text{ or } (-5, y); -4 \le y \le 4\}$ and

(b) $S = \{(x, y); -4 \le x \le 4; -4 \le y \le 4\}$.

The sets Y_1 and Y are identified via the bijection $\psi : Y_1 \longrightarrow Y$ given by

$$\begin{cases} \psi(x, -5) = 1000(x+4) + 900 + 90 + (5-x) \\ \psi(-5, y) = 100(y+4) + 10(5-y) \\ \psi(x, y) = 1000(x+5) + 100(y+4) + 10(4-y) + (5-x) \quad \text{if } x, y \neq -5. \end{cases}$$

ψ maps M into L and S into Q. And we use the letter g to represent the map of Y_1 into Y_1 induced by the previously defined map f of Y into Y and the bijection ψ, that is, $g = \psi^{-1} \circ f \circ \psi$.

Symmetries of the map g

To exploit the symmetry of the map g, consider the following two permutations of the set Y_1:

(i) σ defined by

$$\begin{cases} \sigma(x, y) = (x, -y) & \text{if } (x, y) \in S \\ \sigma(x, -5) = (x, -5) & \text{if } -3 \le x \le 4 \\ \sigma(-5, y) = (-5, 1-y) & \text{if } -3 \le y \le 4 \\ \sigma(-4, -5) = (-4, -5) & \\ \sigma(-5, -4) = (-5, -4). & \end{cases}$$

(ii) τ that sends (x, y) to (y, x).

Theorem 1. *We have $g\sigma(x, y) = g(x, y)$ and $g\tau(x, y) = g(x, y)$ for every (x, y) in the set Y_1.*

Proof. Note that $\sigma(M) = M$, $\sigma(S) = S$, $\tau(M) = M$, and $\tau(S) = S$, so we separate the cases where $(x, y) \in S$ and $(x, y) \in M$. First we suppose that $(x, y) \in S$. That $g(x, -y) = g(x, y)$ is a consequence of the fact that the numbers $1000(x+5) + 100(y+4) + 10(-y+4) + (5-x)$ and $1000(x+5) + 100(-y+4) + 10(y+4) + (5-x)$ of $Q \subset Y$ have the same image under f, since the two numbers are formed by the same digits. We remark that we also have $g(-x, y) = g(x, y)$, by a similar argument.

Next we show that $g(y, x) = g(x, y)$ for (x, y) in S where $x \neq y$, for otherwise the equality is obvious. Suppose that $x > y \ge 0$, changing, if necessary, x in $-x$, y in $-y$, and interchanging x and y. Then we have the following inequalities:

$$x + 5 > y + 4 \geq -y + 4 \geq -x + 5,$$

$$x + 4 \geq y + 5 \geq -y + 5 > -x + 4.$$

To prove that $g(x, y) = g(y, x)$ is the same as proving that the numbers

$$1000(x + 5) + 100(y + 4) + 10(-y + 4) + (-x + 5)$$

and

$$1000(y + 5) + 100(x + 4) + 10(-x + 4) + (-y + 5)$$

have the same image under f. Taking into account the above inequalities, both numbers are mapped to $999(2x) + 90(2y)$ under f. This shows that the theorem holds for (x, y) in S.

Suppose now that $(x, y) \in M$. To prove that $g(-5, y) = g(-5, 1 - y)$, for $-3 \leq y \leq 4$, we look at the numbers $\psi(-5, y) = 100(y + 4) + 10(5 - y)$ and $\psi(-5, 1 - y) = 100(5 - y) + 10(y + 4)$. These two numbers have the same image under f, proving our assertion. Finally we show that $g(x, y) = g(y, x)$ for $(x, y) \in M$. This amounts to prove that $g(x, -5) = g(-5, x)$ for $-4 \leq x \leq 4$. Now $\psi(x, -5) = 1000(x + 4) + 900 + 90 + (5 - x)$ and $\psi(-5, x) = 100(x + 4) + 10(5 - x)$. If $x + 4 > 5 - x$, then f maps the first number to $(9 - (5 - x))999 + (9 - (x + 4))90$ and the second to $(x + 4)999 + (5 - x)90$ which are the same. The case where $x + 4 < 5 - x$ is dealt in a similar fashion. ∎

The permutations σ and τ of Y_1 generate a subgroup isomorphic to the dihedral group D_4 of eight elements, the group of symmetries of the square. For, if we call $\alpha = \sigma\tau$, then α and τ generate the same group as σ and τ, α has order 4, τ has order 2, and $\tau\alpha = \alpha^{-1}\tau$. This is the usual representation of D_4 by generators and relations, expressing D_4 as a semidirect product of a cyclic group of order 4 and a cyclic group of order 2. As an immediate consequence of this observation and the previous theorem we have

Theorem 2. *For any permutation λ of Y_1 in D_4, $g\lambda(x, y) = g(x, y)$.*

If we consider the orbits of the action of D_4 on the subset S, that is, the subsets of the form $D_4(x, y) = \{\lambda(x, y);\ \lambda \in D_4\}$, the previous theorem says that the map g is constant on each orbit. Now, obviously, the orbit of an element (x, y) of S under this action is constituted by all elements obtained from (x, y) by changing the sign of x, or of y, by interchanging x and y or by any combination of these procedures. For example, the orbit of (3, 1) has the following eight elements

$$\{(3, 1), (1, 3), (-1, 3), (-3, 1), (-3, -1), (-1, -3), (1, -3), (3, -1)\}.$$

There are six orbits of eight elements: the orbits of (x, y) where $4 \geq x > y > 0$. There are eight orbits of four elements: the orbits of $(x, 0)$, and the orbits of (x, x), where $1 \leq x \leq 4$. Finally $(0, 0)$ is alone in its own orbit.

In M there is an orbit of two elements, the orbit of $(-4, -5)$, which pairs with $(-5, -4)$. The other four orbits have four elements. These are the orbits of $(x, -5)$, where $1 \leq x \leq 4$, that bring together $(x, -5)$, $(1 - x, -5)$, $(-5, x)$, and $(-5, 1 - x)$.

Now we look for a formula for $g(x, y)$, where (x, y) is in S. We have seen during the proof of Theorem 1 that if $x > y \geq 0$ then f maps the number $1000(x + 5)$

$+\,100(y+4)+10(-y+4)+(-x+5)$ to $999(2x)+90(2y)$. It follows immediately that $g(x, y) = (2x-5, 2y-5)$ for $x > y \geq 0$. Using the symmetric properties of g we get

$$g(x, y) = (2|x|-5, 2|y|-5); \quad \text{if} \quad |x| > |y|,$$

$$g(x, y) = (2|y|-5, 2|x|-5); \quad \text{if} \quad |x| < |y|.$$

Suppose that $1 \leq x \leq 4$. We consider $g(x, x)$ by looking at the number $1000(x+5)+100(x+4)+10(-x+4)+(-x+5)$. Since

$$x+5 > x+4 > -x+5 > -x+4$$

f maps the number to $999(2x+1)+90(2x-1)$. It follows that $g(x, x) = (2x-4, 2x-6)$. Considering the constancy of g in the orbit of (x, x) we get

$$g(x, x) = g(x, -x) = (2|x|-4, 2|x|-6); \quad \text{if} \quad x \neq 0.$$

Finally by an easy computation we get

$$g(0, 0) = (-4, -4).$$

Let us turn our attention to the elements of the set M beginning with the pairs $(x, -5)$ where $-4 \leq x \leq 4$. To find $g(x, -5)$, we consider the number $1000(x+4)+900$
$+\,90+(-x+5)$. If $x \leq 0$, then $9 \geq 9 \geq -x+5 > x+4$ and f maps the number to $999(-x+5)+90(x+4)$. If $x \geq 1$, then $9 \geq 9 > x+4 > -x+5$ so the number goes to $999(x+4)+90(-x+5)$. Consequently

$$g(x, -5) = (-x, x-1) \quad \text{if} \quad x \leq 0,$$

$$g(x, -5) = (x-1, -x) \quad \text{if} \quad x > 0.$$

We remark that $g(1-x, -5) = g(x, -5)$ for $x \neq -4$, as expected. For the pairs $(-5, y)$, where $-4 \leq y \leq 4$, we know that $g(-5, y) = g(y, -5)$, and

$$g(-5, y) = (-y, y-1) \quad \text{if} \quad y \leq 0,$$

$$g(-5, y) = (y-1, -y) \quad \text{if} \quad y > 0.$$

At this point we can prove the following.

Theorem 3. *The map $f: X \longrightarrow X$ has one and only one fixed point. This fixed point is the number* 6174.

Proof. Clearly we can replace the set X by the smaller set Y since any fixed point, should it exist, must be in Y. We prove that the map g on Y_1 has one and only one fixed point. The map g does not fix any point of $M \subset Y_1$ since it is apparent from the formulas for g on M that the only element that maintains the coordinate equal to -5, after applying g, is $(-4, -5)$ but $g(-4, -5) = (4, -5)$. In S, g does not fix any point of the form (x, x) or $(x, -x)$ as the formulas for g clearly show. Of the remaining points of S the result after applying g is a point whose both coordinates are odd. Now there is only one orbit where the elements have odd coordinates. It is the orbit of $(3, 1)$. At this point we can say that there is at most a fixed point for g. If $g(3, 1)$ is an element of the same orbit then that image will be our fixed point. If $g(3, 1)$ is not in the orbit of

$(3, 1)$ we will have no fixed point. Now our formulas for g on S give $g(3, 1) = (1, -3)$, clearly an element in the orbit of $(3, 1)$. So $(1, -3)$ is the only fixed point for g. Since $\psi(1, -3) = 6174$, we see that 6174 is the only fixed point of the map $f : X \longrightarrow X$. ■

As usual denote by $g^2 = g \circ g$. In the set $S \subset Y_1$ consider the elements (x, y) where $|x| \geq 1$ and $|y| \geq 1$. These elements constitute a subset of S made up of four square-like disjoint sets. In this subset consider the permutation

$$\mu(x, y) = \left(\frac{x}{|x|}(5 - |y|), \quad \frac{y}{|y|}(5 - |x|) \right).$$

These are reflections on the sides of the square of equation $|x| + |y| = 5$. The next theorem presents another symmetry of the map g.

Theorem 4. $g^2\mu(x, y) = g^2(x, y)$, *where* $1 \leq |x| \leq 4$ *and* $1 \leq |y| \leq 4$.

Proof. If (x_1, y_1) belongs to the orbit of (x, y), then $\mu(x_1, y_1)$ belongs to the orbit of $\mu(x, y)$. So we can restrict the proof to the case where $1 \leq y \leq x$. Then $\mu(x, y) = (5 - y, 5 - x)$. If $x > y$, then $5 - y > 5 - x$ and hence $g(5 - y, 5 - x) = (5 - 2y, 5 - 2x)$ which is in the same orbit as $g(x, y) = (2x - 5, 2y - 5)$. So in this case $g^2\mu(x, y) = g^2(x.y)$. If $y = x$, then $g\mu(x, x) = g(5 - x, 5 - x) = (6 - 2x, 4 - 2x)$ which is in the same orbit as $g(x, x) = (2x - 4, 2x - 6)$. The conclusion follows as in the previous case. ■

Working in the quotient space

We consider now a set Z which is a quotient space of Y_1. The elements of Z are: (i) The fifteen orbits of the action of D_4 on S and (ii) the five orbits of the action of D_4 on M.

Z is a set with twenty elements. We represent the elements of Z by $[x, y]$, where (x, y) is any element in the equivalence class. Since the map g is constant on each of these equivalence class, it induces a map φ of Z into Z. From the formulas for g in the section Preliminaries, we see that the elements of Y_1 in the image of g are of one of these forms: (a) (x, y) where both x and y are odd; (b) (x, y) where both x and y are even and $x - y = 2$; (c) (x, y) where $x + y = -1$, and (d) (-4,-4).

From this we conclude that the six equivalence classes

$$[3, 0], [4, 0], [4, 1], [0, 0], [2, 2], [-4, -5]$$

have no elements in the image of g. Hence these elements are not in the image of φ. So we pick these equivalence classes and apply the map φ iteratively. We get the following chains for the first three of these equivalence classes:

$$[3, 0] \longrightarrow [1, -5] \longrightarrow [1, 0] \longrightarrow [-3, -5] \longrightarrow [4, 3] \longrightarrow [3, 1]$$

$$[4, 0] \longrightarrow [3, -5] \longrightarrow [3, 2] \longrightarrow [1, 1] \longrightarrow [4, 2] \longrightarrow [3, 1].$$

$$[4, 1] \longrightarrow [3, 3] \longrightarrow [2, 0] \longrightarrow [-1, -5] \longrightarrow [2, 1] \longrightarrow [3, 1]$$

We get three chains of six elements all ending with $[3, 1]$ which is the fixed point for the map φ. $[3, 1]$ is also the only common element in these three chains. We used therefore sixteen of the twenty elements of Z. Starting with $[-4, -5]$ we get the sequence

$$[-4, -5] \longrightarrow [-3, -5] \longrightarrow [4, 3] \longrightarrow [3, 1].$$

This, except for the first element, is part of the first of the six elements chains. Starting with [0, 0] we have

$$[0, 0] \longrightarrow [4, 4] \longrightarrow [4, 2] \longrightarrow [3, 1].$$

This has two new classes and then blends into the second of the six elements chains. Finally, starting with [2, 2],

$$[2, 2] \longrightarrow [2, 0] \longrightarrow [-1, -5] \longrightarrow [2, 1] \longrightarrow [3, 1].$$

This, apart from the first element, coincides with the last four elements of the third of the six elements chains. We already knew that g and hence φ has one and only one fixed point. But now we can state the following theorem.

Theorem 5. *The map φ is such that φ^5 is constant.*

Define depth of an equivalence class $[x, y]$ in Z, $d_\varphi([x, y])$, as the smallest number k such that $\varphi^k([x, y]) = [3, 1]$. We saw that for all $[x, y]$ in Z, $0 \leq d_\varphi([x, y]) \leq 5$. Now since $g(4, 3) = (3, 1)$, $g(4, 2) = (3, -1)$, and $g(2, 1) = (-1, -3)$, the only elements in Y_1 that are mapped by g into $(1, -3)$, the fixed point of g, are those in the equivalence class [3, 1]. Hence, if we define the depth of an element (x, y) of Y_1 as the smallest integer k such that $g^k(x, y) = (1, -3)$, we have

$$\begin{cases} d(x, y) = d_\varphi([x, y]) + 1 & \text{for } (x, y) \neq (1, -3) \\ 0 & \text{for } (x, y) = (1, -3). \end{cases}$$

Here is a table with the depth of the elements of Y_1.

Depth of elements of Y_1	
depth	*Subset of Y_1*
6	$[3, 0] \cup [4, 0] \cup [4, 1]$
5	$[1, -5] \cup [3, -5] \cup [3, 3] \cup [2, 2]$
4	$[1, 0] \cup [3, 2] \cup [2, 0] \cup [-4, -5] \cup [0, 0]$
3	$[-3, -5] \cup [1, 1] \cup [-1, -5] \cup [4, 4]$
2	$[4, 3] \cup [4, 2] \cup [2, 1]$
1	$[3, 1] - \{(1, -3)\}$
0	$\{(1, -3)\}$

Coming back to the map f of X into X

The restriction of the map f to the set Y has the same properties of the map g of Y_1 into Y_1, since $g = \psi^{-1} \circ f \circ \psi$. In particular f^6 is constant in Y and we can find the depth of an element u of Y by looking at the depth of the element $\psi^{-1}(u)$ of Y_1. If we consider an element u of X then, as we pointed out before, $f(u)$ belongs to Y so we can find out the depth of $f(u)$. And, unless $u = 6174$, the fixed point of f, we have the depth of u, $d(u)$, equal to $d(f(u)) + 1$. So all the elements of X have depth less or equal to seven. Hence we can state the following theorem.

Theorem 6. *The map $f\colon X \longrightarrow X$ is such that f^7 is constant.*

But are there elements of depth equal to seven? Take for example 51. $f(u)$ corresponds to the pair $(m, n) = (5, 1)$ and hence to the pair $(0, -4)$ of Y_1. If we look at our table we see that all elements in the class [4, 0] have depth equal to six. Hence the element 51 of X has depth equal to seven. More generally, take any element of

X less than 100, $u = 10m + n$, where m and n are between 0 and 9 and are not both zero. Then the depth of u is equal to $d(m-5, n-5)$ plus one. For example, there are sixteen elements of X of depth seven and less than 100. They are

85 58 52 25 95 59 51 15 96 69 94 49 61 16 41 14.

The first four correspond to the four elements in the equivalence class [3, 0], the next four to the four elements in the equivalence class [4, 0], and the last eight to the eight elements in the equivalence class of [4, 1], the elements of depth equal to six of Y_1 (see table in the previous section). Continuing we can see that there are also sixteen numbers between 1 and 99 of depth six, since the equivalence classes in the row of depth five in the table of the previous section total sixteen elements. And there are, between 1 and 99, 19 elements of depth five, 16 elements of depth four, 24 elements of depth three, 6 of depth two, and 2 of depth one. One has to take care not to be misled by the last two rows of the table that seems to point out to seven elements of depth two and one element of depth one. 26 and 62 are the two elements of depth one, between 1 and 99.

If we take any element $u = 1000a + 100b + 10c + d$ of X we may start by transforming it to $v = 1000e + 100f + 10g + h$, where $e \geq f \geq g \geq h$ are the digits a, b, c, and d in descending order. Then we consider $w = 10(e-h) + (f-g)$. Of course $f(u) = f(v) = f(w)$. So, $d(u) = d(e-h-5, f-g-5) + 1$, unless $u = 6174$, in which case, $d(u) = d(1, -3) = 0$. Suppose we want a list of all the numbers u in X such that $f(u) = 6174$, that is, a list of all the numbers in x of depth equal to one. From the table in the previous section, we see that we should look for the elements $u = 1000e + 100f + 10g + h$ in X such that $e - h = 6$, $f - g = 2$, where $e \geq f \geq g \geq h$, and their permutations. This gives us 6200, 6310, 6420, 6530, 6640, 7311, 7421, 7531, 7641, 7751, 8422, 8532, 8642, 8752, 8862, 9533, 9643, 9753, 9863, 9973, and their permutations. We get 384 numbers, but we have to ignore 6174, so in fact we have only 383 numbers of depth equal to one in X.

A formula for the depth

It would be nice to have a formula for the depth of an element of Y_1, and hence of X, instead of a table. Theorem 4, by revealing a new symmetry, albeit for the map g^2 and only in a subset of Y_1, provided sufficient new relations in the depth function as to make us believe that there was such a formula and that the formula could be a quadratic expression in x and y, at least for $|x|$ and $|y|$ between 1 and 4, invariant under the permutation μ. Note that $g^2\mu(x, y) = g^2(x, y)$ is not equivalent to $d(x, y) = d\mu(x, y)$, but the only exceptions are those between elements in [3, 1] and [4, 2]. After some calculations we got the formula in the next theorem.

Theorem 7. *Let* $\delta(x, y) = (x - y - 1)^2 + (5 - x)y - 2$. *Then for* $4 \geq x \geq y \geq 1$ *we have*

$$d(x, y) = \begin{cases} \delta(x, y) & \text{for } x = y \text{ or } (x, y) = (3, 1) \\ 2\delta(x, y) & \text{for } x \neq y \text{ and } (x, y) \neq (3, 1). \end{cases}$$

Note that $\delta(5 - y, 5 - x) = \delta(x, y)$. Note also that, except for the special case of (3, 1), $d(x, y) = \delta(x, y)$ if the orbit of (x, y) has four elements and $d(x, y) = 2\delta(x, y)$ if the orbit of (x, y) has eight elements. Taking into account the symmetry properties of the map g we get formulas for $d(x, y)$, where $|x|$ and $|y|$ are between 1 and 4. We have just to take care with $d(1, -3) = 0$ and not 1, since it is the fixed point.

The formula in the theorem takes care of the majority of cases. For the remaining values of $d(x, y)$ we can get formulas for the depth, not as nice as the formula in the last theorem. We have

$$d(x,0) = d(0,x) = \begin{cases} (5-|x|)|x| & x \quad \text{odd} \\ (5-|x|)|x| + 2(-1)^{\frac{x}{2}} & x \neq 0 \quad \text{even} \\ 4 & x = 0. \end{cases}$$

To obtain $d(x, -5)$, where x is equal to -3, -1, 1, or 3 we have to remember that $g(x, 0)$, for x equal to 1, 2, 3, or 4 is equal to $(2x - 5, -5)$, hence

$$\begin{aligned} d(-5,x) &= d(x,-5) \\ &= d(\frac{x+5}{2}, 0) - 1 \\ &= \begin{cases} \dfrac{(5-x)(5+x)}{4} - 1 & \text{for } x = -3 \quad \text{or } x = 1 \\ \dfrac{(5-x)(5+x)}{4} + 2(-1)^{\frac{x+5}{4}} - 1 & \text{for } x = -1 \quad \text{or } x = 3. \end{cases} \end{aligned}$$

The transformation $y = 1 - x$ gives us the formulas for $d(x, -5)$ for x equal to 4, 2, 0, or -2.

$$d(-5,x) = d(x,-5) = d(1-x,-5)$$

for x equal to 4, 2, 0, or -2. It only remains that $d(-4, -5) = 4$.

As an application let us find quickly the depth of 2014. We have $m = 4 - 0 = 4$ and $n = 2 - 1 = 1$. Hence $x = m - 5 = -1$ and $y = 1 - 5 = -4$. The depth of $(-1, -4)$ is equal to the depth of $(4, 1)$. By theorem 7, we get $d(4, 1) = 2\delta(4, 1) = 6$. Hence 2014 has the maximum depth equal to seven.

Remark. The contour lines of the quadratic function in two variables $\delta(x, y) = (x - y - 1)^2 + (5 - x)y - 2$, that we used to compute the depth of the majority of the elements of Y_1, are hyperbolas or the two intersecting lines that are asymptotes to all the other contour lines. Since the axis of symmetry of all these hyperbolas make 45 degree angles with the coordinate axis, the slopes m_1 and m_2 of the asymptotes must satisfy $m_1 m_2 = 1$. The interesting part is that $m_1 = 1 + \Phi = \Phi^2$ and $m_2 = 1 + \Phi' = \Phi'^2$, where Φ is the golden ratio and Φ' is the conjugate root of Φ in the equation $u = 1 + \frac{1}{u}$.

And beyond...

Instead of four digit integers we can consider a similar problem for integers composed of a fixed number of digits. The cases of two and three digits integers are too simple. What about the problem concerning five digit integers? A similar approach, keeping the notation, would lead us to a set Y whose elements are the integers that can be put in the form $9999m + 990n$, where m and n have the same range of variation as in the case of four digit integers. That is right. The case of five digit integers is also bidimensional as the previous case. But some symmetry is lost. The new map g in the set Y_1 is the following:

$$\begin{cases} g(x, y) = (|x| - 1, |x| + |y| - 4) & \text{if } |y| < |x| \neq 5 \\ g(x, y) = (|y|, |x| + |y| - 6) & \text{if } |x| < |y| \neq 5 \\ g(x, y) = (|x|, 2|x| - 5) & \text{if } |x| = |y| \neq 0 \\ g(0, 0) = (0, -4) & \\ g(x, -5) = (-x, x - 1) & \text{if } x \leq 0 \\ g(x, -5) = (x - 1, -x) & \text{if } x > 0 \\ g(-5, y) = (4, -x) & \text{if } y \leq 0 \\ g(-5, y) = (4, x - 1) & \text{if } y > 0. \end{cases}$$

The lost symmetry is because $g(x, y) \neq g(y, x)$. Although, as can easily be checked by the previous formulas $g(y, x) - g(x, y) = (1, -2)$ for $0 \leq y < x$. The formulas also show that there are no fixed points. And starting with any integer we will end up with one of three loops or cycles {53955, 59994}, {74943, 62964, 71973, 83952}, {63954, 61974, 82962, 75933}.

What about the problem for six digit integers? Well, the complexity increases. That is now a three-dimensional case, that is, we have to deal with a set Y whose members are the integers that can be put in the form $99999m + 9990n + 900t$. But some of the lost symmetry comes back, and that is a hopeful sign.

We leave the rest for the interested reader.

REFERENCES

1. D. R. Kaprekar, An interesting property of the number 6174, *Scr. Math.* **15** (1955) 244–245.
2. D. R. Kaprekar, On Kaprekar numbers, *J. Recreat. Math.* **13** (1980–81) 81–82.
3. D. Deutsch, B. Goldman, Kaprekar's constant, *Math. Teacher* **98** (2004) 234–242.

Summary. In a problem involving an operation on four digit integers, introduced more than sixty years ago, a dihedral type of symmetry is used to prove the existence of a unique fixed point and to assess the speed of convergence towards this fixed point. Additional symmetry points the way to a formula to compute how far an integer is from the fixed point in terms of the operation.

MANUEL R. F. MOREIRA (MR Author ID: 235163) is a professor of Mathematics at the Faculty of Sciences of the University of Porto. He has a first degree in Mathematics from the University of Porto and a Ph.D. in Mathematics from M.I.T. He is currently Chair of the Math Department.

The Fifteen Puzzle—A New Approach

S. MURALIDHARAN
Decision Sciences and Algorithms Lab
Tata Consultancy Services, Chennai, INDIA
drsmuralidharan@gmail.com

You probably have seen the 15-puzzle before. This puzzle, also called the Gem Puzzle, Boss Puzzle, Game of Fifteen, Mystic Square, and other names, consists of sliding square tiles numbered 1 to 15 that occupy all but one cell of a 4×4 box (Figure 1). The basic goal is to use the blank cell to slide the square tiles so that the final arrangement is the natural order of the numbers 1 to 15. Do you know how to solve it? Perhaps a trial and error approach might work. Starting from an arbitrary initial arrangement, is it possible to arrive at the natural order? Have you seen an algorithm that you could use for any initial position? This puzzle has a rich history. Sam Loyd claimed that he was the inventor of the puzzle but it may have been invented by Noyes Chapman, a postmaster in Canastota, New York.

Only half of the possible 15! initial arrangements can be restored to the natural order but that was not evident when this puzzle swept the United States and other parts of the world in the 1880's. The major challenge was to achieve the goal starting from a near perfect initial array in which the last row contained 13, 15, 14 in that order—the only difference being the permutation of the last two adjacent tiles. This challenge gripped the people of all ages and walks of life. Sam Loyd even announced a prize of 1000 US dollars for the solver of this challenge.

1	**2**	**3**	**4**
5	**6**	**7**	**8**
9	**10**	**11**	**12**
13	**14**	**15**	

Figure 1 The 15-puzzle.

It can be proved that of the 15! possible initial arrangements, half of them can be restored to the natural order and the other half to the order in which the first three rows are in natural order and the last row is 13-15-14 (see [**1, 2, 5, 6, 7, 10**]). In particular, the 13-15-14 arrangement cannot be restored to the natural order.

Almost all the proofs use the theory of the alternating group A_{15} but do not exhibit a set of moves to arrive at one of the above two possible arrangements from a given initial arrangement.

There are also several algorithms for solving the puzzle. Even for the general $N \times N$ board, one can find a solution using heuristics, but finding the optimal number of moves in which the general puzzle can be solved is known to be an NP-hard problem (*e.g.*, [**8, 9**]).

In this paper we give a simpler, elementary (nongroup-theoretic) proof that exactly half of the 15! arrangements can be restored to the natural order. We also describe

Math. Mag. **90** (2017) 48–57. doi:10.4169/math.mag.90.1.48.
MSC: Primary 91A60, Secondary 05A05.

an algorithm to restore any arrangement into one of the two arrangements mentioned above. The key strategy we use is "divide and conquer." We reduce the problem to the 2×4 board and use the results proved for this board to derive the main theorem. Our approach generalizes to the $N \times N$ board as well.

2×4 Puzzle We first consider the 2×4 puzzle in which we arrange the numbers $1, 2, \ldots, 7$ in some order and study which of those arrangements can be restored to the natural order. With any arrangement, we associate a sequence formed by following the arrow in Figure 2. For example, for the arrangement shown in Figure 2, the sequence

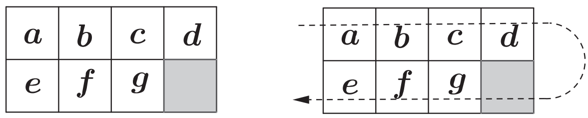


Figure 2 2×4 board and listing order.

associated is (a, b, c, d, g, f, e). Note that we ignore the blank space in forming the sequence.

In any sequence we say that an *inversion* occurs each time a larger number precedes a smaller number. For example, in the sequence $(1, 2, 3, 4, 7, 6, 5)$, the number of inversions is 3 and in $(7, 6, 5, 4, 3, 2, 1)$, the number of inversions is 21.

An arrangement is said to have *even parity* if the corresponding sequence has an even number of inversions and *odd parity* if the sequence has odd number of inversions. Inversions hold the key to determine which arrangements can be restored to the natural order. We prove the following theorem.

Theorem 1. *An arrangement of the numbers in the* 2×4 *board can be restored to the natural order if and only if the arrangement has odd parity.*

We begin with a lemma.

Lemma 1. *If two adjacent numbers in a sequence are interchanged, the resulting sequence has the opposite parity of the original sequence.*

Proof. Suppose that in the sequence $s_1 = (\ldots, a, b, \ldots)$ we interchange a, b to get the sequence $s_2 = (\ldots, b, a, \ldots)$. If $a < b$, then s_2 has one more inversion caused by the larger number b preceding the smaller a, since all other inversions of the sequence s_1 are preserved in s_2. Thus s_2 has opposite parity to the parity of s_1. Also, if $a > b$, the inversion in s_1 caused by the larger a preceding the smaller b is removed in s_2 by the interchange. Thus s_2 has one less inversion than s_1. Thus again, s_2 has opposite parity to the parity of s_1. ∎

As a corollary, we make the following observation. The sequences

$$s_1 = (\ldots, a, x_1, x_2, \ldots, x_r, \ldots),$$

$$s_2 = (\ldots, x_1, x_2, \ldots, x_r, a, \ldots)$$

have the same parity if r is even and have opposite parity if r is odd.

This is clear from the lemma, since we can obtain s_2 by interchanging a with $x_1, x_2, \ldots, x_r$, in that order. Since each interchange reverses the parity, there are r reversals. We first prove that the legal moves do not change the parity.

Lemma 2. *Any legal move does not change the parity of the arrangement.*

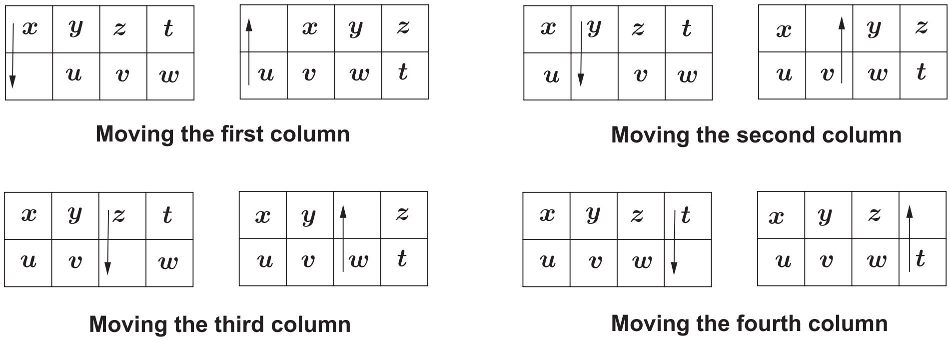


Figure 3 Vertical moves.

Proof. Any horizontal shift does not change the sequence and hence the parity of the sequence is not affected.

Consider the vertical moves (Figure 3).

When we move the x in the first column, the sequence (x, y, z, t, w, v, u) becomes (y, z, t, w, v, u, x) and hence x moves to the final position after 6 interchanges. Hence by the observation above, the resulting sequence has the same parity as the original sequence.

When we move u in the first column, second row to the first column, first row, the sequence changes from (x, y, z, t, w, v, u) to (u, x, y, z, t, w, v) and again u is interchanged six times. Thus in this case also the parity of the sequence does not change.

Similarly, when we move a symbol in the second column, the number of interchanges is 4 and hence the parity is maintained.

When we move a symbol in the third column, the number of interchanges required is 2 and finally, moving a symbol in the last column does not change the sequence at all. Thus in all cases the parity of the sequence is maintained. This completes the proof of the lemma. ■

We now introduce some basic moves that play a key role in the proof of Theorem 1.

$S(\pm n)$: In this move, we slide the symbols either clockwise or anticlockwise so that each symbol moves by n cells. The moves $S(+1)$ and $S(-1)$ are shown in Figure 4.

R_1: This move allows for three adjacent symbols to cycle in a row under certain conditions (as shown in Figure 5). The proof that R_1 can be obtained by a sequence of legal moves is shown in Figure 6.

E_1, E_2: These moves interchange two adjacent columns, as shown in Figure 7. E_1 can be obtained as follows. (See Figure 8.) Shift e, move b down, apply $S(-2)$, push d down, and move a, c anticlockwise. The legal moves constituting E_2 are shown in Figure 9.

T: This move shifts elements in a triangular cycle among two adjacent rows. For example, in Figure 10 we shift a from the top row to the bottom row and shift d from the bottom row to the top row. Figure 11 shows how this can be accomplished using legal moves.

R_2: This moves three elements in the bottom row cyclically, as shown in Figure 12. The legal moves leading to R_2 are given in Figure 13.

Now we prove Theorem 1. Start with an arbitrary arrangement of $1, 2, \ldots, 7$ in the 2×4 board, as given in Figure 14. By using the moves T and R_1, we see that any symbol in the first row can be interchanged with the first symbol in the second row, with no other symbols switching rows (although their order within the row may change). Again, using E_1 and E_2, we can bring any symbol in the second row to the

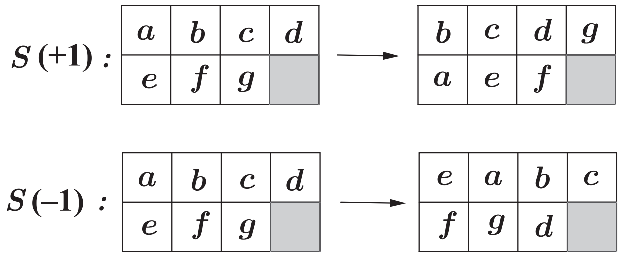


Figure 4 Slide move $S(\pm 1)$.

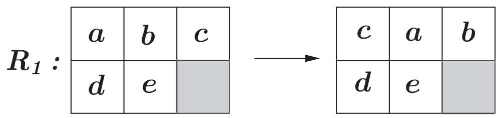


Figure 5 Move R_1.

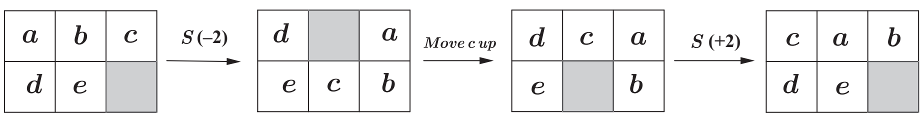


Figure 6 Legal moves for R_1.

first cell in the second row (again without impacting the set of symbols in the individual rows) and hence it follows that we can exchange any pair of symbols between the first and second rows. Thus we can bring the symbols 1, 2, 3, 4 to the first row and 5, 6, 7 to the second row. Now using E_1, E_2, we can assume that 4 is in the first row, fourth column. Using R_2, we can assume that the second row is either 5, 6, 7 in that order or 5, 7, 6. Thus we have reached one of the two arrangements A_1, A_2 in Figure 15 where (x, y, z) is 1, 2, 3 in some order. Suppose that the initial arrangement had odd parity, and we reached the A_1 arrangement. The parity of the sequence $(x, y, z, 4, 7, 6, 5)$ must also be odd and since $(4, 7, 6, 5)$ has three inversions, it follows that (x, y, z) must have an even number of inversions. Hence (x, y, z) must be one of $(1, 2, 3)$, $(3, 1, 2)$, or $(2, 3, 1)$. Now, using R_1, we can move this arrangement to the natural order.

Now suppose that we reach the arrangement A_2. Since $(4, 6, 7, 5)$ has two inversions, the sequence (x, y, z) must have an odd number of inversions. Thus it must be one of $(1, 3, 2)$, $(2, 1, 3)$, or $(3, 2, 1)$. The $(1, 3, 2)$ case is already solved (we use E_1 to interchange the second and third columns and we obtain the natural order). For $(2, 1, 3)$, first use E_1 to interchange second and third columns and then R_1 to obtain the natural order. For $(3, 2, 1)$, we use R_1 to change the first row to $(1, 3, 2)$ and use E_1

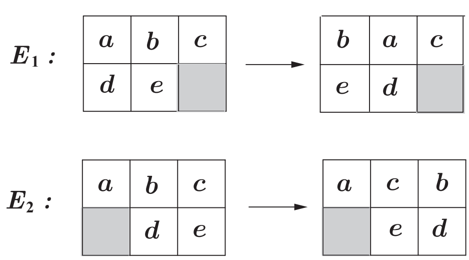


Figure 7 Moves E_1 and E_2.

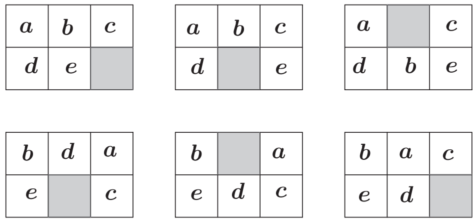


Figure 8 Legal moves for E_1.

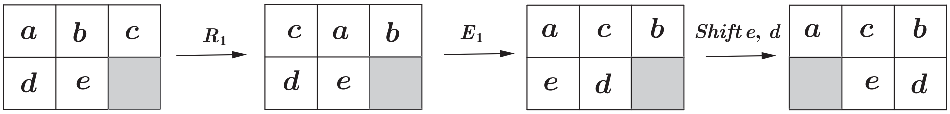


Figure 9 Legal moves for E_2.

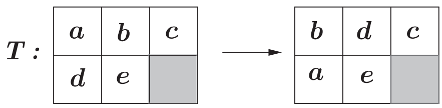


Figure 10 Move T.

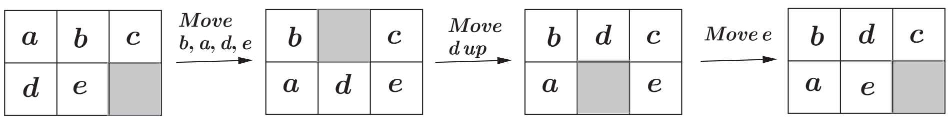


Figure 11 Legal moves for T.

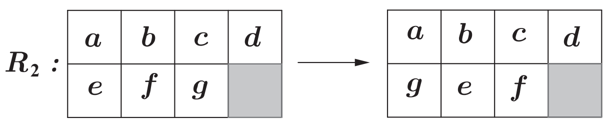


Figure 12 Move R_2.

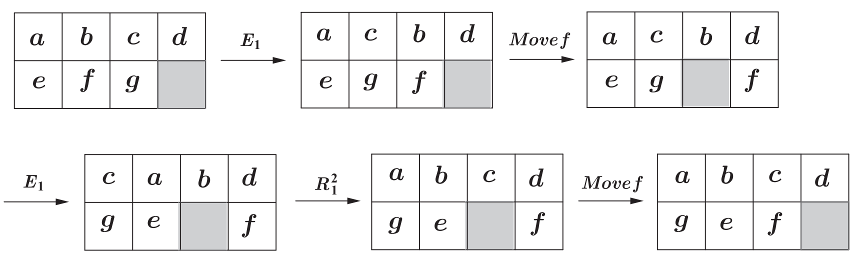


Figure 13 Legal moves for R_2.

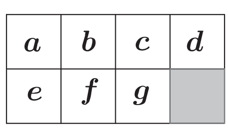


Figure 14 Starting arrangement.

x	y	z	4
5	6	7	

A_1

x	y	z	4
5	7	6	

A_2

Figure 15 Proof of Theorem 1.

to interchange the second and third columns. This results in the natural order. Hence if we start with an arrangement with odd parity, we can reach the natural order.

A similar argument shows that if we start with an arrangement with even parity we reach the arrangement in which the first row is 1, 2, 3, 4 in that order and the second row is 5, 7, 6 in that order. Thus we have proved the following theorem.

Theorem 2. *Starting with any arrangement on the* 2×4 *board, we reach the arrangements* A_{nat} *or* A_{rev} *(Figure 16) when the parity of the original arrangement is odd or even, respectively.*

1	2	3	4
5	6	7	

A_{nat}

1	2	3	4
5	7	6	

A_{rev}

Figure 16 Final positions for 2×4 board.

4×4 Puzzle As in the 2×4 case, we first choose a suitable listing order and associate a sequence with any arrangement in such a way that any legal move does not alter the parity of the associated sequence. For the 4×4 board, we choose the listing sequence by following the path shown in Figure 17. With this listing order, the

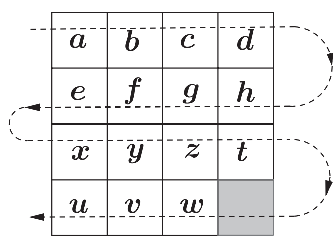


Figure 17 Listing order.

horizontal moves do not change the sequence. For the vertical moves, there are those that do not change the sequence. These are shown in Figure 18. The other vertical moves of an arbitrary element u will move u two, four, or six places in the sequence, as shown in Figures 19, 20, and 21, respectively. In all the cases, the symbols are shifted in the sequence by an even number of places and hence the parity of the sequence is not altered.

Now we are ready to prove the following theorem.

Theorem 3. *An arrangement of the numbers in the* 4×4 *board can be restored to the natural order if and only if the arrangement has odd parity.*

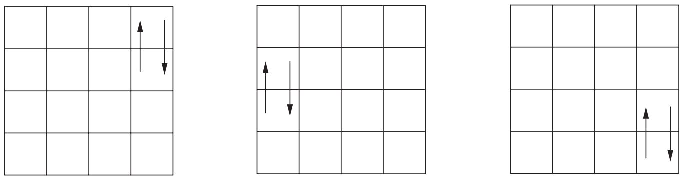


Figure 18 Moves that do not change the sequence.

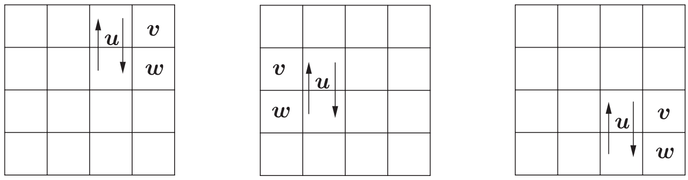


Figure 19 Moves that shift a symbol by 2 places.

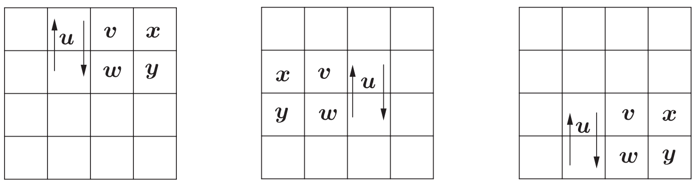


Figure 20 Moves that shift a symbol by 4 places.

The "only if" part is obvious—if the arrangement can be restored to the natural order using legal moves, then it necessarily has odd parity since the sequence associated with the natural order

$$(1, 2, 3, 4, 8, 7, 6, 5, 9, 10, 11, 12, 15, 14, 13)$$

has 9 inversions and parity is not changed by legal moves.

Now suppose that we start with an arrangement with odd parity. Let us call the 2×4 board consisting of the first and second rows of the board B_{top} and the 2×4 board consisting of the third and fourth rows B_{bot}. Using the move T repeatedly, we can move all numbers $1, 2, 3, \dots, 8$ to B_{top} and the numbers $9, 10, \dots, 15$ to the B_{bot}. Again through a combination of the moves E_1, E_2, R_1, R_2, we can assume that 4 is at the first row, fourth column, 8 is at the second row, fourth column, and 12 is at the third row, fourth column with blank occupying the last row, last column. Now since the parity of the board is odd, we have two cases to consider:

1. B_{top} has odd parity, B_{bot} has even parity, and
2. B_{top} has even parity, B_{bot} has odd parity.

1. B_{top} has odd parity, B_{bot} has even parity: Bring the blank cell to the second row,

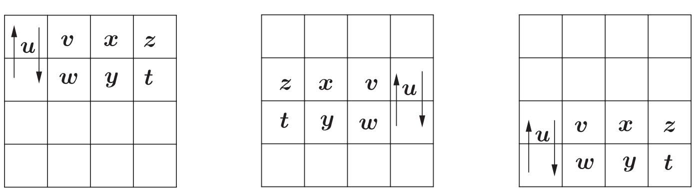


Figure 21 Moves that shift a symbol by 6 places.

fourth column by pushing 12 to the fourth row, fourth column and 8 to third row, fourth column. Since B_{top} has odd parity, the sequence $(a, b, c, 4, 8, x, y, z)$ has odd parity. Since 8 is involved in three inversions (all x, y, z are less than 8), the sequence $(a, b, c, 4, x, y, z)$ has even parity. Thus by Theorem 2, we can bring the board B_{top} to the arrangement A_2 (Figure 22). Now move 8 back to second row,

1	2	3	4
5	7	6	

A_2

9	10	11	12
13	15	14	

A'_2

Figure 22 B_{top} and B_{bot} arrangement.

fourth column, 12 to third row, fourth column, and consider the bottom board B_{bot}. Since this has even parity, again by Theorem 2, we can bring the board B_{bot} to the arrangement A'_2 (Figure 22). Now, applying the move E_1 in A'_2, we can change it to A''_2 (Figure 23). Stitching together A_2 and A''_2, and moving 12, we obtain the

9	11	10	12
13	14	15	

A''_2

1	2	3	4
5	7	6	8
9	11	10	
13	14	15	12

A_3

Figure 23 B_{bot} arrangement and full board.

arrangement in A_3. Now, applying the move E_1 to the second and third rows, we obtain the natural order. This completes proof for this case.

2. B_{top} has even parity, B_{bot} has odd parity

   Since B_{top} has even parity, the sequence $(a, b, c, 4, 8, x, y, z)$ has even parity. Since 8 is involved in three inversions (all x, y, z are less than 8), the sequence $(a, b, c, 4, x, y, z)$ has odd parity. Sliding 8 and 12 down, applying Theorem 2, and sliding 8 and 12 up show that we can bring the board B_{top} to the natural order. Also, since B_{bot} has odd parity, it can be brought to the natural order by Theorem 2. Now, stitching these together, we obtain the natural order on $1, 2, \ldots, 15$.

A similar argument proves the following result.

Theorem 4. *An arrangement of the numbers in the* 4×4 *board can be restored to the order* $13 - 15 - 14$ *if and only if the arrangement has even parity.*

The above proof also gives an algorithm for restoring the board to one of the two arrangements. The steps are as follows:

1. Move $1, 2, \ldots, 8$ to the top two rows, bring 4, 8, and 12 into their proper positions, slide 8 and 12 down, bring $1, 2, \ldots, 7$ into the natural order or $5 - 7 - 6$ order, slide 8 and 12 back up.
2. (a) top board is in natural order, bottom board is in natural order: the full board is in natural order

(b) top board is in natural order, bottom board is in $13 - 15 - 14$ order: the full board is in $13 - 15 - 14$ order

(c) top board is in $5 - 7 - 6$ order, bottom board is in natural order: Use E_1 in second and third rows and again in third and fourth rows. The board will be in $13 - 15 - 14$ order

(d) top board is in $5 - 7 - 6$ order, bottom board is in $13 - 15 - 14$ order: Use E_1 in second and third rows and again in third and fourth rows. The board will be in natural order.

The above algorithm may not yield the optimal number of moves required to restore the board. It is known that the 15-puzzle can be solved in maximum of 80 single tile moves [**11**].

In general, permutation puzzles (*e.g.,* [**3, 4, 6**]) are interesting and any discussion of them involves some amount of group theory. Another famous permutation puzzle, Rubik's cube, has been studied extensively and there are algorithms for solving the cube. The usual proofs in the literature for the 15-puzzle are all existence proofs—that is, using the theory of permutation groups, one shows that there exists a set of moves—without explicitly describing the moves—to restore any starting arrangement to one of the two arrangements mentioned in the beginning of this paper. An arrangement is first mapped to a permutation in the symmetric group S_{15} on 15 symbols (the blank cell is assumed to be at the 16th place). One shows that an arrangement can be restored to the natural order of the numbers 1 to 15 if and only if the mapped permutation is even. The proof uses the fact that the subgroup A_{15} of even permutations is generated by the 3-cycles [**1**]. For the 15-puzzle, there are algorithms [**6, 10**] but none of the published algorithms use a divide and conquer approach. They also appear to be specific to 4×4 board. The proof given in this paper provides a better intuition for solving the puzzle and also extends to the more general $N \times N$ board.

Generalization It is easy to see that the above proof can be modified to prove that for any board of size $m \times n$, two arrangements can be reached from one another if and only if they have the same parity.

Acknowledgment The author would like to thank the referees and the editor for their helpful suggestions towards improving this paper. Thanks to Sachin Lodha for reviewing the draft version of the paper and for his encouragement.

REFERENCES

1. A. F. Archer, A modern treatment of the 15 puzzle, *Amer. Math. Monthly* **106** (1999) 793–799.
2. B. Averbach, O. Chein, *Problem Solving through Recreational Mathematics*. Dover Publications, New York, 1980. 288–299.
3. T. A. Brown, A note on "Instant Insanity," *Math. Mag.* **41** (1968) 167–169.
4. A. P. Grecos, R. W. Gibberd, A diagrammatic solution to "Instant Insanity" problem, *Math. Mag.* **44** (1971) 119–124.
5. H. Liebeck, Some generalizations of the 14-15 puzzle, *Math. Mag.* **44** (1971) 185–189.
6. J. Mulholland, Permutation Puzzles: A Mathematical Perspective, http://www.sfu.ca/jtmulhol/math302/,2015.
7. Wm. Woolsey Johnson, W. E. Story, Notes on the "15" puzzle, *Amer. J. Math.* **2** (1879) 393–404.
8. D. Ratner, W. Manfred, Finding a Shortest Solution for the $N \times N$ Extension of the 15-PUZZLE Is Intractable. National Conference on Artificial Intelligence, 1986.
9. D. Ratner, W. Manfred, The $(n^2 - 1)$-puzzle and related relocation problems, *J. Symbolic Comput.* **10** no. 2 (1990) 111–137.
10. E. L. Spitznagel, Jr., A new look at the fifteen puzzle, *Math. Mag.* **40** (1967) 171–174.
11. http://oeis.org/A087725.

Summary. We give an elementary, nongroup theoretic proof that exactly half of the 15! arrangements of the fifteen puzzle can be restored to the natural order.

S. MURALIDHARAN (MR Author ID: 196417) heads Decision Sciences and Algorithms lab at Tata Consultancy Services in Chennai, India. His research interests include analytics and algorithms.

1 A	2 D	3 O	4 S			5 S	6 N	7 A	8 P			9 S	10 P	11 I	12 N
13 W	I	S	P		14 B	O	O	B	O	15 O		16 I	O	T	A
17 O	V	A	L		18 O	N	T	A	P	E		19 D	I	S	H
20 L	O	G	I	21 S	T	I	C	C	U	R	22 V	E			
23 S	T	E	N	O		24 C	H	I	P		25 O	V	26 I	27 N	28 E
			29 E	D	30 U				31 M	32 A	X	I	M	A	L
33 P	34 T	35 A		36 O	N	37 E	38 O	39 V	E	R		40 E	M	I	L
41 L	O	N	42 G	I	T	U	D	I	N	A	43 L	43 W	A	V	E
44 O	L	G	A		45 I	C	E	C	U	B	E		46 D	E	N
47 T	E	L	L	48 A	L	L				49 S	I	50 N			
51 S	T	E	I	N		52 I	53 D	54 E	55 M		56 C	O	57 S	58 E	59 C
			60 L	O	61 G	D	E	R	I	62 V	A	T	I	V	E
63 M	64 A	65 Z	E		66 N	E	B	U	L	A		67 N	E	E	D
68 A	R	E	A		69 C	A	R	P	E	T		70 O	G	R	E
71 V	E	N	N			72 N	A	T	S			73 W	E	T	S

SOLUTION TO PINEMI PUZZLE

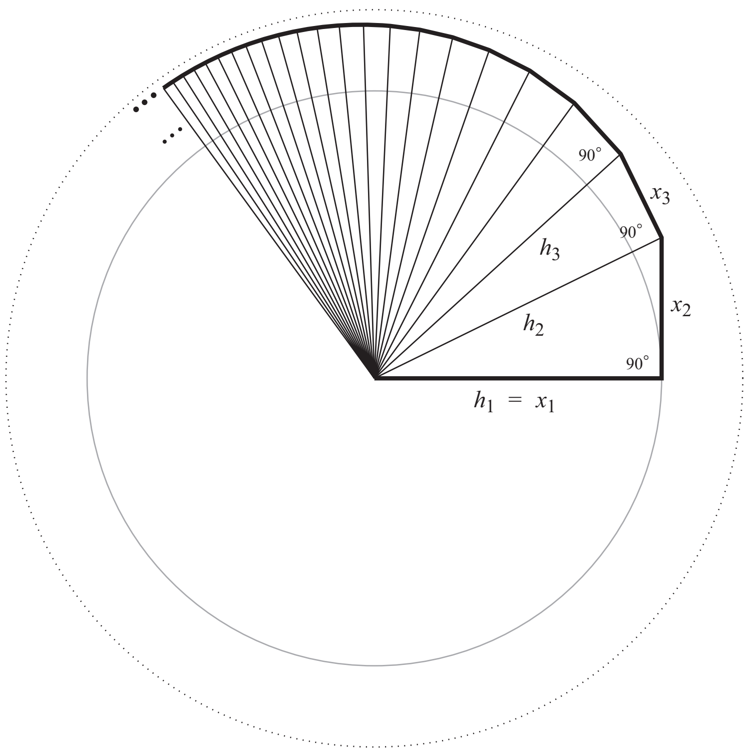
<table>
<tr><td>7</td><td>|||</td><td>7</td><td>5</td><td>|||</td><td>6</td><td>|</td><td>6</td><td>||</td><td>|</td></tr>
<tr><td>|</td><td>|||</td><td>11</td><td>|</td><td>|</td><td>|</td><td>6</td><td>|||</td><td>8</td><td>5</td></tr>
<tr><td>|||</td><td>10</td><td>||</td><td>||</td><td>11</td><td>7</td><td>|</td><td>6</td><td>|</td><td>|</td></tr>
<tr><td>4</td><td>10</td><td>|</td><td>|||</td><td>11</td><td>|||</td><td>|</td><td>6</td><td>6</td><td>||</td></tr>
<tr><td>|</td><td>6</td><td>|||</td><td>11</td><td>||</td><td>|</td><td>|</td><td>|</td><td>|</td><td>9</td></tr>
<tr><td>6</td><td>|</td><td>10</td><td>||</td><td>7</td><td>7</td><td>|</td><td>9</td><td>|||</td><td>|||</td></tr>
<tr><td>||</td><td>||</td><td>||</td><td>6</td><td>||</td><td>8</td><td>9</td><td>||</td><td>12</td><td>9</td></tr>
<tr><td>|</td><td>9</td><td>6</td><td>5</td><td>8</td><td>|||</td><td>||</td><td>|</td><td>||</td><td>|</td></tr>
<tr><td>|</td><td>|</td><td>5</td><td>|</td><td>9</td><td>||</td><td>|||</td><td>12</td><td>|</td><td>|</td></tr>
<tr><td>|</td><td>5</td><td>||</td><td>|</td><td>||</td><td>7</td><td>8</td><td>|||</td><td>6</td><td>|</td></tr>
</table>

Proof Without Words: $\ell^1(\mathbb{R})$ Is a Subset of $\ell^2(\mathbb{R})$

JUAN LUIS VARONA
Universidad de La Rioja
26006 Logroño, Spain
jvarona@unirioja.es

Let's take the sequence $\mathbf{x} = \{x_n\}_{n=1}^{\infty}$, with $x_n > 0$ for every n. We define a new sequence $\{h_n\}_{n=1}^{\infty}$ as follows.

$$h_1 = x_1, \quad h_n^2 = h_{n-1}^2 + x_n^2, \quad n \geq 2 \quad \Rightarrow \quad h_n = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}$$



Because Length(spiral) $= \|\mathbf{x}\|_{\ell^1}$ and Radius(asymptotic outer circle) $= \|\mathbf{x}\|_{\ell^2}$, then

$$\text{Length(spiral)} < \infty \quad \Rightarrow \quad \text{Radius(asymptotic outer circle)} < \infty, \text{ or}$$

$$\ell^1(\mathbb{R}) \subseteq \ell^2(\mathbb{R}).$$

Summary. We prove that $\ell^1(\mathbb{R})$ is a subset of $\ell^2(\mathbb{R})$ by means of a spiral.

JUAN LUIS VARONA (MR Author ID: 260232) is professor in the Department of Mathematics and Computation in the University of La Rioja, in Spain. He has published more that 50 research papers as well as an undergraduate book on number theory. He is editor in-chief of La Gaceta de la Real Sociedad Matemática Española.

Math. Mag. **90** (2017) 58–58. `doi:10.4169/math.mag.90.1.58`.
MSC: Primary 46B45, Secondary 40A05.

A Unified Pythagorean Theorem in Euclidean, Spherical, and Hyperbolic Geometries

ROBERT L. FOOTE
Wabash College
Crawfordsville, IN 47933
footer@wabash.edu

Consider a right triangle with legs x and y and hypotenuse z. If the triangle is on the unit sphere S^2 or in the hyperbolic plane H^2 with constant curvature -1, its sides satisfy

$$\cos z = \cos x \cos y \quad \text{or} \quad \cosh z = \cosh x \cosh y, \tag{1}$$

respectively. (On the sphere we assume the triangle is *proper*, that is, its sides have length less than half of the circumference.) These formulas for the Pythagorean theorem seem to have little to do with the familiar $z^2 = x^2 + y^2$ in Euclidean geometry. Where are the squares (regular quadrilaterals) and their areas? Why are the right hand side products and not sums? What is the meaning of the cosine of a distance?

Some authors (e.g., [**6**]) find an analogy by expanding the formulas in (1) into power series. The constant terms cancel, the first-order terms vanish, and the second-order terms agree with the Euclidean formula. The author finds this comparison unsatisfying, and to students not yet comfortable with power series it seems more like a parlor trick. Yes, for very small triangles in S^2 and H^2 the power series say that the Euclidean formula is approximately true, but their meaning is unclear for large triangles. No additional geometric insight is gained, and one wonders how to interpret the higher-order terms that are ignored.

Our main goal is the following theorem, which gives a common formula for the Pythagorean theorem in all three geometries. Throughout the paper let M denote $\mathbb{R}^2$, S^2, or H^2.

Theorem 1. (Unified Pythagorean Theorem) *A right triangle in M with legs x and y and hypotenuse z satisfies*

$$A(z) = A(x) + A(y) - \frac{K}{2\pi} A(x)A(y), \tag{2}$$

where $A(r)$ is the area of a circle of radius r and K is a constant.

Before getting into the proof, a few comments are in order. This theorem and its proof are in neutral geometry, the geometry that $\mathbb{R}^2$, S^2, and H^2 have in common. We tend to focus on the differences between these geometries, but they share quite a bit. For example, rotations are isometries, something needed in the proof.

The constant K turns out to be the Gaussian curvature of M. While the treatment of this is beyond the scope of this article, the reader will see K arise in the proof along with consequences its sign has on the geometry of circles in M. In S^2 we have $K = 1/R^2$, where R is the radius of the sphere. In $\mathbb{R}^2$, $K = 0$. In H^2, K can be any negative value and R defined by $K = -1/R^2$ is often called the pseudoradius. (For an expository treatment of Gaussian curvature, see [**11**]. For technical details, see [**14, 15**].)

Math. Mag. **90** (2017) 59–69. doi:10.4169/math.mag.90.1.59.
MSC: Primary 51M09.

Hopefully the reader finds (2) more geometric than (1). The somewhat mysterious term $KA(x)A(y)/2\pi$ is interpreted as an area following the proof of Theorem 1. Note that when $K = 0$, (2) reduces to the Euclidean version, although the squares on the sides of the triangle have been replaced by circles. When $K \neq 0$ and x and y are small, (2) is approximately Euclidean since the product $A(x)A(y)$ is small compared to the other terms. Of course in $\mathbb{R}^2$ the squares can be replaced by any shape since area for similar figures scales in proportion to the square of their linear dimensions. Not so in S^2 and H^2 where similarity implies congruence. The author believes that it remains an open problem to find an expression of the Pythagorean theorem in S^2 and H^2 with figures on the sides other than circles. (See [**8**] and [**9**, p. 208] for the use of semirectangles in the "unification" of a related theorem.)

There have been a number of efforts to find unifying formulas for the three geometries going back to Bolyai, who gave a unified formula for the law of sines (see [**2**], especially pp. 102, 114). Following the proof of Theorem 1 we give a unified version of (1) as a corollary. Additional consequences of the proof include some differential-geometric results and formulas for circumference, area, and curvature of circles. At the end we indicate how (2) generalizes to a unified law of cosines.

The first formula in (1) was well known to students of spherical trigonometry in the early 1800s [**3**, p. 164]. Lobachevski proved a version of the second formula in his development of hyperbolic geometry in the 1820s–30s, although he did not express it in terms of hyperbolic cosine [**3**, p. 174].

Our arguments are intrinsic in the sense used in [**9, 14**]. In particular, for S^2 we do not refer to its embedding in $\mathbb{R}^3$; for H^2 we do not make use of any model. For extrinsic proofs of (1) using vector algebra, see [**17**, pp. 50, 86]. For proofs of the second formula in (1) using models of H^2 in $\mathbb{R}^2$, see [**6, 13**].

Proof in $\mathbb{R}^2$

Our starting point is the following rotational "bicycle" proof of the Pythagorean theorem in $\mathbb{R}^2$ [**12**]. Given $\triangle XYZ$ with right angle at Z, rotate the triangle in a circle centered at X (Figure 1). The sides $\overline{XY}$ and $\overline{XZ}$ sweep out areas πz^2 and πy^2, respectively. The third side, $\overline{YZ}$, sweeps out the annulus between the circles. This segment can be thought of as moving like a bicycle of length x with its front wheel at Y and its rear wheel at Z.

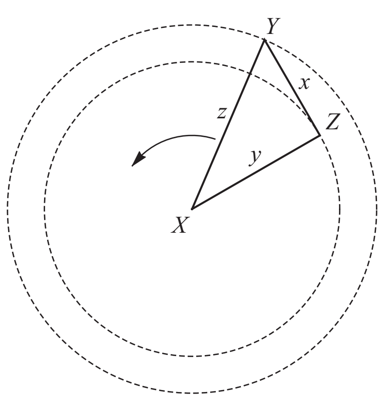


Figure 1 Rotational proof of the Pythagorean theorem.

A simple model of a bicycle is a moving segment of fixed length ℓ, where ℓ is the wheelbase (the distance between the points of contact of the wheels with the ground).

The segment (frame of the bicycle) moves in such a way that it is always tangent to the path of the rear wheel R (see Figures 2 and 3). An infinitesimal motion of the bicycle is determined by ds, the rolling distance of the rear wheel, and $d\theta$, the change of direction or turning angle, shown in Figure 2. (It's possible for the rear wheel to roll backwards, in which case $ds < 0$, but we will not need that in this paper.)

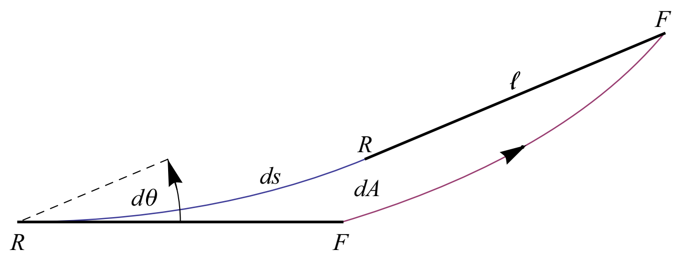


Figure 2 Infinitesimal bicycle motion.

Consider the area dA swept out by the bicycle frame in Figure 2. It is some linear combination of ds and $d\theta$: $dA = m\,ds + n\,d\theta$. In fact,

$$dA = \frac{\ell^2}{2}\,d\theta \tag{3}$$

because the forward motion of the bicycle (when $ds \neq 0$ and $d\theta = 0$) sweeps out no area (so $m = 0$), while the turning motion ($d\theta \neq 0$, $ds = 0$) sweeps out area at the rate of $n = \pi\ell^2/2\pi = \ell^2/2$ per radian turned.

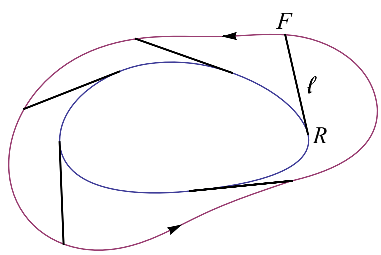


Figure 3 The annular area is $\pi\ell^2$ in $\mathbb{R}^2$.

If the rear wheel of the bicycle goes around a convex loop (Figure 3), its direction has turned by $\Delta\theta = 2\pi$, and so by (3) it sweeps out an area of $(\ell^2/2)\Delta\theta = \pi\ell^2$. Applying this to the segment $\overline{YZ}$ of the rotating triangle, we see that this side sweeps out an area of πx^2. The sides $\overline{XZ}$ and $\overline{YZ}$ combined sweep out the same area as the side $\overline{XY}$, that is, $\pi z^2 = \pi x^2 + \pi y^2$, and we obtain the Pythagorean theorem in $\mathbb{R}^2$.

These simple descriptions and consequences of a moving segment of fixed length sweeping out area (whether it moves like a bicycle or not) go back at least to 1894 in a paper [**10**] giving the history and theory of planimeters up to that time. (See [**4**] for more details and additional references.) The more recent notions of tangent sweeps and tangent clusters [**1**] incorporate and extend these ideas. Figure 4a shows the annular area of Figure 1 as a tangent sweep of infinitesimal triangles formed by the positions of the bicycle. Figure 4b is the corresponding tangent cluster in which the triangles have been rearranged into a disk of radius x, illustrating that the annulus and disk have

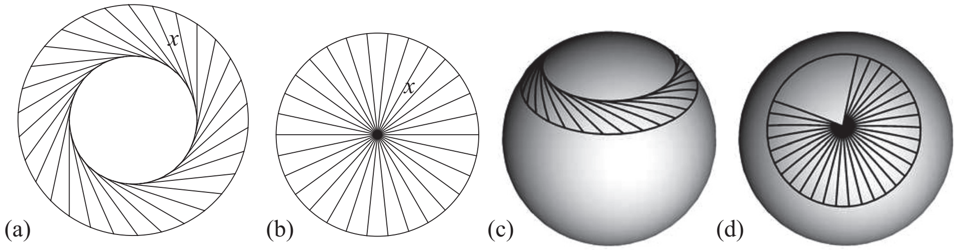


Figure 4 Tangent sweeps and tangent clusters in $\mathbb{R}^2$ and S^2.

the same area. The last two parts of Figure 4 show a tangent sweep and cluster on a sphere. They suggest that something different may happen there, as the tangent cluster does not form a full disk. They also suggest that we may have made an assumption that the tangent cluster *does* form a full disk in $\mathbb{R}^2$! This discrepancy is resolved in the next two sections.

Turning around a circle

Consider a circle of radius ρ in M. Let $A(\rho)$ and $C(\rho)$ denote its area and circumference, respectively. It is convenient to let $a(\rho) = A(\rho)/2\pi$ and $c(\rho) = C(\rho)/2\pi$, which we think of as the area and circumference per radian of a circular sector of radius ρ (Figure 5). Formulas for $A(\rho)$ and $C(\rho)$ are given in the last section in (13). We do not need them in the general proof; in fact, we derive them as *consequences* of the proof.

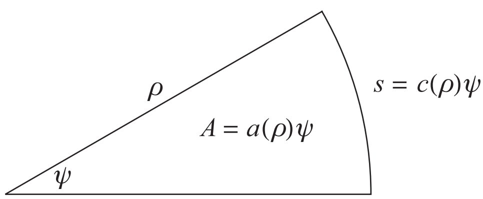


Figure 5 Definitions of $a(\rho)$ and $c(\rho)$.

The argument in the previous section works in S^2 and H^2, but one must be careful. The analog of (3) for the area dA in Figure 2 is

$$dA = a(\ell)\, d\theta, \tag{4}$$

however the turning angle $d\theta$ is more subtle.

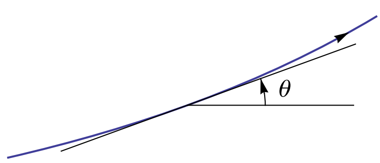


Figure 6 Geodesic curvature in $\mathbb{R}^2$: $\kappa = d\theta/ds$.

Turning angle is closely related to geodesic curvature. A common way to define these for a curve in $\mathbb{R}^2$ is to let the turning angle be the angle θ from some fixed direction to the tangent line (Figure 6), and then to let the geodesic curvature be $\kappa = d\theta/ds$, where s is arc length along the curve. Unfortunately this doesn't work in S^2 or H^2. In S^2 there is no notion of fixed direction. In H^2 there are fixed directions (given by the points at infinity), but then $d\theta/ds$ would depend on the direction used. We leave their precise definitions to the section on circle geometry. For the purpose of proving Theorem 1, it suffices to intuitively note some of their properties.

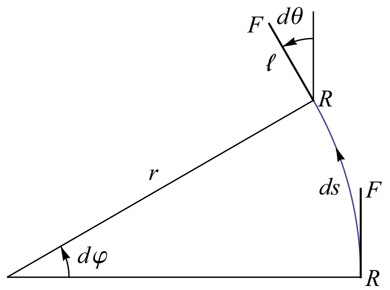


Figure 7 Bicycling around a circle: $d\theta = d\varphi$ only in $\mathbb{R}^2$.

The turning angle and geodesic curvature of a curve depend on the curve's orientation. We orient a circular arc in M by traversing it counterclockwise when viewed from its center, that is, $d\varphi > 0$ in Figures 7 and 8. Note that a circle in S^2 has two centers; specifying one of them determines its radius and orientation.

When bicycling around a circle (partially shown in Figure 7), we expect the total turning angle to be $\Delta\theta = \int d\theta = 2\pi$, however, this need not be the case on a curved surface. The equality we expect between the turning angle $d\theta$ and the central angle $d\varphi$ in Figure 7 is a feature of Euclidean geometry. To make this plausible, consider Figure 8. The first picture shows that we likely have $d\theta < d\varphi$ on S^2. The second shows a bicycle following a great circle (consequently going straight, turning neither left nor right), in which case $d\theta \equiv 0$. The true relationship between $d\theta$ and $d\varphi$ is given below in (5) and more fully in (9) as a consequence of the general proof of Theorem 1.

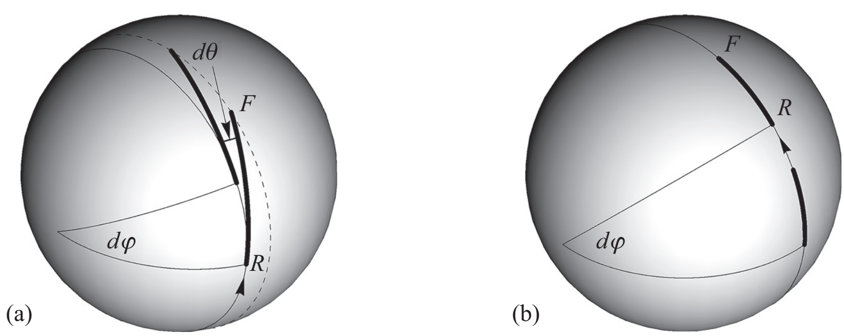


Figure 8 Bicycling around circles in S^2: (a) $d\theta < d\varphi$, (b) $d\theta \equiv 0$.

Any two circular arcs on M with the same length s and radius r are congruent since one can be mapped to the other with a composition of rotations and translations. As a result, they have the same constant curvature $\kappa(r)$ and the same turning angle $\Delta\theta = \kappa(r)s$. For a circle of radius r we have

$$d\theta = \kappa(r)\, ds = c(r)\kappa(r)\, d\varphi \quad \text{and} \quad \Delta\theta(r) = C(r)\kappa(r) = 2\pi c(r)\kappa(r), \tag{5}$$

where $d\theta$, ds, and $d\varphi$ are shown in Figure 7 and $\Delta\theta(r)$ is the total turning angle around the circle. These formulas partially explain the spherical tangent cluster in Figure 4d. We expect $d\theta < d\varphi$ on S^2, in which case $\Delta\theta(r)/2\pi = c(r)\kappa(r)$ represents the fraction of the disk taken up by the tangent cluster.

Proof of the general case and a corollary

We are now ready to prove Theorem 1 in all three geometries simultaneously. As the triangle in Figure 1 rotates around X, the sides $\overline{XY}$ and $\overline{XZ}$ sweep out areas $A(z)$ and $A(y)$, respectively. From (4) and (5), the area of the annulus swept out by $\overline{YZ}$ is $\Delta A = a(x)\Delta\theta(y) = A(x)\Delta\theta(y)/2\pi = A(x)c(y)\kappa(y)$.

The sides $\overline{XZ}$ and $\overline{YZ}$ together sweep out the same area as the side $\overline{XY}$, that is, $A(z) = A(y) + A(x)\Delta\theta(y)/2\pi = A(y) + A(x)c(y)\kappa(y)$, which is an asymmetric Pythagorean theorem. Rotating around Y instead of X, we get $A(z) = A(x) + A(y)c(x)\kappa(x)$. Dividing these by 2π yields

$$a(z) = a(y) + a(x)c(y)\kappa(y) \quad \text{and} \quad a(z) = a(x) + a(y)c(x)\kappa(x). \tag{6}$$

Setting the expressions in (6) equal to each other and separating the variables leads to $\big(1 - c(x)\kappa(x)\big)/a(x) = \big(1 - c(y)\kappa(y)\big)/a(y)$. Since x and y are independent, this quantity is constant, i.e., there is a constant K such that

$$1 - c(r)\kappa(r) = Ka(r) \tag{7}$$

for every $r > 0$ that is the radius of a circle.

Now use (7) with $r = x$ to eliminate $c(x)\kappa(x)$ in the second equation of (6) resulting in

$$a(z) = a(x) + a(y) - Ka(x)a(y).$$

Multiplying by 2π yields (2), and completes the proof.

Using (7) in the opposite way, that is, to eliminate $a(z)$, $a(x)$, and $a(y)$ in the second equation of (6), leads to the formula in the following corollary.

Corollary. *A right triangle in M with legs x and y and hypotenuse z satisfies*

$$c(z)\kappa(z) = c(x)\kappa(x)\, c(y)\kappa(y). \tag{8}$$

Given the formulas for $c(r)$ and $\kappa(r)$ in (13), this becomes the first formula in (1) when $K = 1$ and the second formula when $K = -1$. Ironically, when $K = 0$ (in $\mathbb{R}^2$), (8) becomes the true but useless equation $1 = 1{\cdot}1$, since $c(r) = r$ and $\kappa(r) = 1/r$ in that case. Thus (8) is a unified Pythagorean theorem only for the non-Euclidean geometries (cf. [**2**, p. 114]).

Two related observations come out of the proof. First, using (7) the expressions in (5) become

$$d\theta = \big(1 - Ka(r)\big)\, d\varphi \quad \text{and} \quad \Delta\theta(r) = 2\pi\big(1 - Ka(r)\big). \tag{9}$$

Thus, the relative sizes of $d\theta$ and $d\varphi$ in Figures 7 and 8 depend on the sign of K and the area of the circle followed by the rear wheel. This goes a bit farther than (5) in explaining the tangent clusters in Figure 4. When $K = 0$, the cluster exactly fills the disk (Figure 4b). When $K > 0$, the cluster falls short of filling the disk (Figure 4d). When $K < 0$, the cluster overlaps itself and exceeds the disk.

Second, from the proof we have

$$A(z) = A(y) + A(x)\frac{\Delta\theta(y)}{2\pi} = A(y) + A(x) - \frac{K}{2\pi}A(x)A(y).$$

The first equality shows how the area of the circle of radius z in Figure 1 breaks into the areas of the circle of radius y and the annulus. The factor of $\Delta\theta(y)/2\pi$ indicates that the area of the annulus falls short of, equals, or exceeds $A(x)$ accordingly as $K > 0$, $K = 0$, or $K < 0$. The second equality shows that $KA(x)A(y)/2\pi$ is the difference between the areas of the circle of radius x and the annulus—it is the area of the gap in Figure 4d between the tangent cluster and the full disk when $K > 0$. When $K < 0$, then $|K|A(x)A(y)/2\pi$ is the area of the overlap of the tangent cluster on itself.

Circle geometry

In this section we prove two fundamental relationships about the geometry of circles in M. These are used in the last section to make some additional conclusions from Theorem 1, including formulas for $A(r)$, $C(r)$, and $\kappa(r)$. Both use the fact that small regions in S^2 and H^2 are approximately Euclidean. Intuitively, a creature confined to a small region would not be able to determine which surface it is on empirically.

Proposition 1. *If r is the radius of a circle in M, then $A'(r) = C(r)$.*

Proposition 2. *The geodesic curvature and total turning angle of a circle of radius r in M are $\kappa(r) = C'(r)/C(r) = c'(r)/c(r)$ and $\Delta\theta(r) = C'(r)$.*

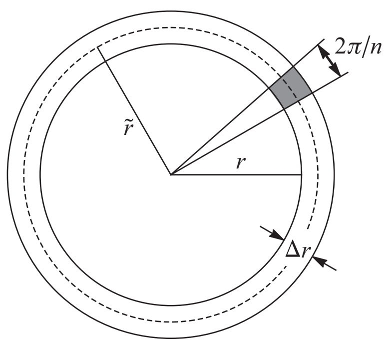


Figure 9 Proof of Proposition 1: $\Delta A = C(\tilde{r})\Delta r$.

To prove Proposition 1, note that the shaded region in Figure 9 (in which n is a sufficiently large positive integer) is nearly rectangular, and so its area is $\frac{C(\tilde{r})}{n}\Delta r$ for some $\tilde{r}$ between r and $r + \Delta r$. Then the annular area in the figure is $\Delta A = C(\tilde{r})\Delta r$, and the result follows.

To prove Proposition 2, we need definitions of geodesic curvature and turning angle, at least for circles. (The definition of geodesic curvature given here differs from the one for curves in a surface $S \subset \mathbb{R}^n$ found in elementary differential geometry texts, e.g., [**14, 15**]. The latter is extrinsic, that is, it depends on the way S is embedded in $\mathbb{R}^n$, and requires more background.)

To motivate definitions that work in all three geometries, consider how the wheels of a wheelchair are related to the geometry of the chair's path. A person in a wheelchair determines how the path curves by controlling the relative speeds of the wheels. If they roll at the same rate, the path is straight, otherwise it is curved. Thus, the amount and rate of turning can be measured from the rolls of the wheels. This can be done in a small region and without reference to any fixed direction. We need to quantify this.

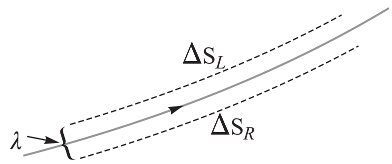


Figure 10 Geodesic curvature in M: $\kappa \approx \frac{\Delta s_R - \Delta s_L}{\lambda \Delta s}$.

Suppose the wheelchair goes straight or follows a circle. In this case the wheels roll in some fixed ratio to each other, and the "parallel" paths they follow are like neighboring lanes of an athletic track. Let Δs_R and Δs_L be the distances traveled by the right and left wheels along some portion of the curve (dotted curves in Figure 10), and let λ be the length of the axle between them. In $\mathbb{R}^2$ it is easy to show that the distance traveled by the middle of the chair is $\Delta s = (\Delta s_R + \Delta s_L)/2$ and that its turning angle (change of direction) is $\Delta\theta = (\Delta s_R - \Delta s_L)/\lambda$. (These formulas are valid even if the path is not a line or circle.) The curvature of the path is then the constant

$$\kappa = \frac{d\theta}{ds} = \frac{\Delta\theta}{\Delta s} = \frac{\Delta s_R - \Delta s_L}{\lambda \Delta s}.$$

It is also easy to show that the path is a circle of radius $1/|\kappa|$ if $\kappa \neq 0$, and a line if $\kappa = 0$. If you work through these exercises, be sure to note your use of similar triangles, similar circular sectors, or your identification of $d\varphi$ and $d\theta$ in Figure 7.

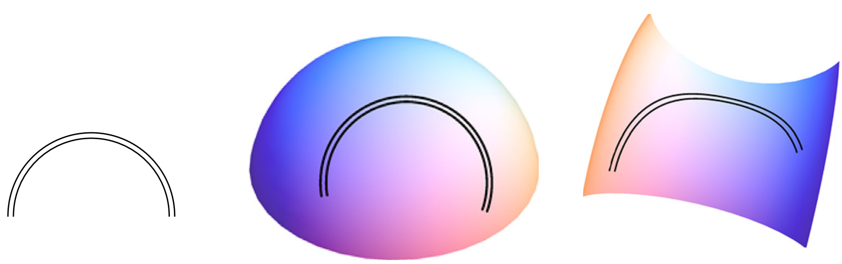


Figure 11 Paper strips with the same curvature and length.

To see what happens in S^2 and H^2, it is instructive to cut a narrow strip of paper following a circular arc. The edges of the strip represent the paths of the wheels. The strip, which is cut from a plane, can easily be applied to a sphere or saddle surface (Figure 11). People following these paths on the different surfaces would agree their paths have the same turning angle and curvature since their wheels roll with the same fixed ratio. This can be verified without needing to know the radius, center of curvature, or even that the path is part of a circle.

The expressions for Δs, $\Delta\theta$, and κ above depend on λ and are approximations in S^2 and H^2. To make them precise, we take the wheelchair to be infinitesimal in size by letting $\lambda \to 0$. We clearly have $\Delta s = \lim_{\lambda\to 0} \frac{1}{2}(\Delta s_R + \Delta s_L)$. More importantly, we define the turning angle along an arc of the curve and the curvature to be

$$\Delta\theta = \lim_{\lambda\to 0} \frac{(\Delta s_R - \Delta s_L)}{\lambda} \quad \text{and} \quad \kappa = \frac{d\theta}{ds} = \frac{\Delta\theta}{\Delta s} = \lim_{\lambda\to 0} \frac{\Delta s_R - \Delta s_L}{\lambda \Delta s}. \qquad (10)$$

With these definitions in hand, the proof of Proposition 2 is immediate. Going around the circle we take $\Delta s_L = C(r - \lambda/2)$, $\Delta s_R = C(r + \lambda/2)$, and $\Delta s = C(r)$. The formulas in (10) then yield

$$\Delta\theta(r) = \lim_{\lambda\to 0} \frac{C(r+\lambda/2) - C(r-\lambda/2)}{\lambda} = C'(r)$$

and $\kappa(r) = \Delta\theta/\Delta s = C'(r)/C(r)$.

With slight modifications the definitions in (10) are valid for other curves in S^2 and H^2 (in fact, in any smooth surface). This simple, intrinsic view of geodesic curvature is common among differential geometers (see the first sections of [**7**] for a nice exposition, albeit in more dimensions), but does not seem to be well represented in the undergraduate literature.

Consequences and related results

In this section we show that some important differential-geometric results (at least special cases for circles) and formulas for $C(r)$, $A(r)$, and $\kappa(r)$ follow directly from the proof of the Pythagorean theorem. (Interested readers will find the more general results in the references.) We also mention how (2) generalizes to the law of cosines.

The second formula in (9) can be written as

$$\Delta\theta(r) + KA(r) = 2\pi,$$

which is the Gauss-Bonnet Theorem for a disk. (The general theorem for a region D homeomorphic to a closed disk in a surface is $\Delta\theta + \iint_D K\,dA = 2\pi$. Here the turning angle $\Delta\theta$ includes $\int_{\partial D} \kappa\,ds$, where κ is the geodesic curvature of ∂D, and the exterior angles at any vertices ∂D may have. For a nice presentation, see [**14**].) A bicyclist who is sure that $\Delta\theta$ is always 2π can conclude from this that $K = 0$. On the other hand, a bicyclist who knows there is a circle for which $\Delta\theta = 0$ that divides his space into two finite, equal areas (Figure 8b) can conclude that $K = 4\pi/A_0$, where A_0 is the area of the whole space.

Combining (7) with Proposition 2 we obtain

$$c'(r) = 1 - Ka(r) \tag{11}$$

for $r > 0$. Multiplying this by $4\pi C(r)$ and using Proposition 1 yields

$$2C(r)C'(r) = 4\pi A'(r) - 2KA(r)A'(r).$$

Integrating leads to

$$C(r)^2 = 4\pi A(r) - KA(r)^2,$$

which is the case of equality in the isoperimetric inequality. (More generally [**16**], if R is a region in M with area A and perimeter C, then $C^2 \geq 4\pi A - KA^2$ with equality if and only if R is a circular disk.)

Differentiating (11) and using Proposition 1 yields

$$c''(r) = -Kc(r). \tag{12}$$

This is a special case of the Jacobi equation, which governs how fast neighboring geodesics spread out. In H^2, two geodesics starting at the same point spread out faster than in $\mathbb{R}^2$. In S^2 they spread out more slowly, then get closer and intersect again.

(See [**15**] for the general Jacobi equation on a surface.) Considering the differential equation (12) with initial conditions $c(0) = 0$ and $c'(0) = 1$ (the limiting value in (11) as $r \to 0^+$) leads to the following formulas for $C(r)$, depending on the sign of K. The formulas for $A(r)$ and $\kappa(r)$ follow from Propositions 1 and 2.

$$\begin{array}{cccc} & S^2\ (K > 0) & \mathbb{R}^2\ (K = 0) & H^2\ (K < 0) \\ C(r) & 2\pi \dfrac{\sin\left(\sqrt{K}r\right)}{\sqrt{K}} & 2\pi r & 2\pi \dfrac{\sinh\left(\sqrt{|K|}r\right)}{\sqrt{|K|}} \\ A(r) & 2\pi \dfrac{1-\cos\left(\sqrt{K}r\right)}{K} & \pi r^2 & 2\pi \dfrac{1-\cosh\left(\sqrt{|K|}r\right)}{K} \\ \kappa(r) & \sqrt{K}\cot\left(\sqrt{K}r\right) & 1/r & \sqrt{|K|}\coth\left(\sqrt{|K|}r\right) \end{array} \tag{13}$$

Note the expected formulas for $\mathbb{R}^2$ when $K = 0$ and for a sphere of radius $R = 1/\sqrt{K}$ when $K > 0$. Readers familiar with hyperbolic geometry will recognize the formulas for a hyperbolic plane of pseudoradius $R = 1/\sqrt{|K|}$ when $K < 0$ [**6, 16**].

In closing, we express the law of cosines in a manner similar to (2).

Theorem 2. (Unified Law of Cosines) *An arbitrary triangle in M (assumed proper in S^2) with side lengths x, y, and z satisfies*

$$A(z) = A(x) + A(y) - \frac{K}{2\pi}A(x)A(y) - \frac{1}{2\pi}C(x)C(y)\cos\gamma,$$

where γ is the angle opposite the side with length z.

A proof in the spirit of this article is left to the reader as an exercise. The general proof in all three geometries is challenging; the proof in $\mathbb{R}^2$ is easier and gives some new insight into the law of cosines. See [**5**] for a proof based on the formulas in (13) and its use to prove a unified formula for cross ratio in these geometries. See [**17**] for the standard formulas for the laws of cosines in S^2 and H^2.

Acknowledgment The author would like to thank the referees for their close readings, thoughtful comments, and useful suggestions.

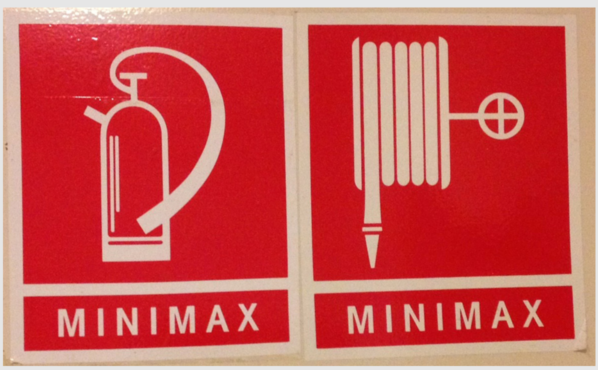
REFERENCES

1. T. M. Apostol, M. A. Mnatsakanian, Subtangents—An aid to visual calculus, *Amer. Math. Monthly* **109** (2002) 525–533, http://dx.doi.org/10.2307/2695442.
2. R. Bonola, *Non-Euclidean Geometry*. Originally published in Italian by Open Court, 1912. English trans. H.S. Carslaw, Dover, New York, 1955.
3. S. Braver, *Lobachevski Illuminated*, Mathematical Association of America, Washington, DC, 2011.
4. R. L. Foote, M. Levi, S. Tabachnikov, Tractrices, bicycle tire tracks, hatchet planimeters, and a 100-year-old conjecture, *Amer. Math. Monthly* **120** (2013) 199–216, http://dx.doi.org/10.4169/amer.math.monthly.120.03.199.
5. R. L. Foote, X. Sun, An intrinsic formula for cross ratio in spherical and hyperbolic geometries, *College Math. J.* **46** (2015) 182–188, http://dx.doi.org/10.4169/college.math.j.46.3.182.
6. M. J. Greenberg, *Euclidean and Non-Euclidean Geometries*. Fourth ed. W. H. Freeman, New York, 2007.
7. M. Gromov, Sign and geometric meaning of curvature, *Milan J. Math.* **61** (1991) 9–123, http://dx.doi.org/10.1007/BF02925201.
8. R. Hartshorne, Non-Euclidean III. 36, *Amer. Math. Monthly* **110** (2003) 495–502, http://dx.doi.org/10.2307/3647905.

9. D. W. Henderson, D. Taimiņa, *Experiencing Geometry*. Third ed. Pearson-Prentice Hall, Upper Saddle River, NJ, 2005.
10. O. Henrici, Report on Planimeters, *British Assoc. Adv. Sci.*. Report of the 64th meeting (1894) 496–523.
11. D. Hilbert, S. Cohn-Vossen, *Geometry and the Imagination*. Originally published in German by Springer, Berlin, 1932. English trans. P. Nemenyi, Chelsea, New York, 1952.
12. M. Levi, *The Mathematical Mechanic: Using Physical Reasoning to Solve Problems*. Princeton Univ. Press, Princeton, NJ, 2009.
13. G. M. Martin, *The Foundations of Geometry and the Non-Euclidean Plane*. Springer-Verlag, New York, 1975.
14. F. Morgan, *Riemannian Geometry, A Beginner's Guide*. Jones and Bartlett, London, 1993.
15. B. O'Neill, *Elementary Differential Geometry*. Academic Press, Orlando, FL, 1966.
16. R. Osserman, Bonnesen-style isoperimetric inequalities, *Amer. Math. Monthly* **86** (1979) 1–29, http://dx.doi.org/10.2307/2320297.
17. J. G. Ratcliffe, *Foundations of Hyperbolic Manifolds*. Graduate Texts in Mathematics Vol. 149, Springer-Verlag, New York, 1994.

Summary. We state a formula for the Pythagorean theorem that is valid in Euclidean, spherical, and hyperbolic geometries and give a proof using only properties the geometries have in common.

ROBERT FOOTE (MR Author ID: 205970) received his Ph.D. from the University of Michigan. He is a professor of mathematics at Wabash College. His research interests include differential geometry and unexplored connections between Euclidean, spherical, and hyperbolic geometries.



In case of fire, use von Neumann's minimax theorem? (This was taken in a small hotel in Basel, Switzerland.)

— contributed by Michael A. Jones, Mathematical Reviews, Ann Arbor, MI.

Euler's Favorite Proof Meets a Theorem of Vantieghem

KONSTANTINOS GAITANAS
National Technical University of Athens
15780 Athens, Greece
raffako@hotmail.com

A famous theorem of Fermat, the so-called little theorem, states that if p is an odd prime number and p does not divide b, then $b^{p-1} \equiv 1 \pmod{p}$. Fermat's little theorem was included in a letter to Frénicle de Bessy in 1640 but it was first proved by Euler in 1736. The Swiss mathematician gave three proofs of this theorem and, according to Dickson [**1**], he preferred his third proof which has a more algorithmic spirit. We will see how this not-so-well-known method and a famous identity, which was introduced by Fermat, prove a theorem of E. Vantieghem.

Sketch of Euler's proof of Fermat's little theorem

Dickson presents Euler's method as the following. There exists a positive integer $\lambda \leq p-1$ that satisfies $b^{\lambda} \equiv 1 \pmod{p}$. (This is not difficult to prove.) If $\lambda = p-1$, then the Fermat's little theorem holds. If not, then $\lambda < p-1$ and there exists a natural number $k < p$ which is not contained in $B_1 = \{b^1 \pmod{p}, b^2 \pmod{p}, \ldots, b^{\lambda} \pmod{p}\}$. Then $B_2 = \{kb^1 \pmod{p}, kb^2 \pmod{p}, \ldots, kb^{\lambda} \pmod{p}\}$ has distinct residues modulo p such that $B_1 \cap B_2 = \emptyset$. If all the residues are contained in $B_1 \cup B_2$, then $\lambda = \frac{p-1}{2}$ and λ divides $p-1$ so that $b^{p-1} \equiv 1 \pmod{p}$, as required. If not, then we construct $B_3 = \{lb^1 \pmod{p}, lb^2 \pmod{p}, \ldots, lb^{\lambda} \pmod{p}\}$ for some l not contained in $B_1 \cup B_2$, and so on. The process terminates so that λ divides $p-1$ and thus $b^{p-1}-1$ is divisible by $b^{\lambda}-1$ and hence by p.

Vantieghem's theorem and applying the idea of Euler's proof

Many similar congruences to Fermat's little theorem also exist. One of the most well known is Wilson's theorem which states that p is prime if and only if $(p-1)! \equiv -1 \pmod{p}$. In 2008, E. Vantieghem [**2**] proved another congruence that characterizes prime numbers. In particular, he proved that $p > 2$ is prime if and only if there exists an integer b with $2 \leq b \leq p-1$ such that

$$\prod_{n=1}^{p-1}(b^n+1) \equiv 1 \left(\bmod \frac{b^p-1}{b-1}\right).$$

As a special case for $b = 2$ we can see that $(2+1)(2^2+1)\cdots(2^{p-1}+1)-1$ is divisible by 2^p-1. If p is composite the congruence does not hold true.

Vantieghem uses the theory of cyclotomic polynomials to prove his theorem. We will prove the "if" case of Vantieghem's theorem without the use of cyclotomic polynomials. Our method of proof will be in the spirit of Euler's idea in his preferred

Math. Mag. **90** (2017) 70–72. doi:10.4169/math.mag.90.1.70.
MSC: Primary 11A07, Secondary 11A51

proof. We will also make use of some basic facts from the theory of congruences and the well-known identity

$$(x^1+1)(x^{2^1}+1)\cdots(x^{2^{n-1}}+1)=\frac{x^{2^n}-1}{x-1}. \tag{1}$$

We recall that the order of $b \pmod p$ is the least exponent r for which the congruence $b^r \equiv 1 \pmod p$ holds.

Theorem 1. *(Vantieghem [**2**]) Let b be a natural number with* $2 \le b \le p-1$. *If* $p > 2$ *is prime, then*

$$\prod_{n=1}^{p-1}(b^n+1) \equiv 1 \left(\bmod \frac{b^p-1}{b-1}\right). \tag{2}$$

Proof. Let p be an odd prime, r be the order of 2 modulo p, and $P=\{1,2,\ldots,p-1\}$. This means that the numbers $1, 2^1, \ldots, 2^{r-1}$ are incongruent $\pmod p$ and from Fermat's little theorem we know that r divides $p-1$. We use an approach similar to Euler's approach by splitting the set $P=\{1,2,\ldots,p-1\}$ into $k=\frac{p-1}{r}$ subsets in the following way. Let $A_1=\{1 \pmod p, 2^1 \pmod p, \ldots, 2^{r-1} \pmod p\}$ be the first set. Let $a_i \in P$ be an integer that is not contained in any of the sets $A_1, \ldots, A_{i-1}$. Then define $A_i=\{a_i\cdot 1 \pmod p, a_i\cdot 2^1 \pmod p, \ldots, a_i\cdot 2^{r-1} \pmod p\}$.

We shall prove that $A_1 \cup A_2 \cup \cdots \cup A_k = P$. It suffices to prove that all the elements of the sets are pairwise incongruent modulo p. If two elements belong in the same set A_i, suppose that $a_i\cdot 2^m \equiv a_i \cdot 2^n \pmod p$ with $n<m$. Since p does not divide evenly into a_i, we obtain $2^n \equiv 2^m \pmod p$. This leads to a contradiction since by definition the numbers $1, 2, \ldots, 2^{r-1}$ are all pairwise incongruent modulo p.

We consider now the case when two elements belong to different sets. Suppose that $a_j\cdot 2^m \equiv a_i\cdot 2^n \pmod p$ with $1 \le m, n \le r-1$ and without loss of generality $i<j$. Multiplying both sides by 2^{r-m}, it follows that $a_j\cdot 2^r \equiv a_i \cdot 2^{r+n-m} \pmod p$ which implies $a_j \equiv a_i\cdot 2^{r+n-m} \pmod p$. But this means that $a_j \in A_i = \{a_i\cdot 1, \ldots, a_i\cdot 2^{r-1}\}$, which is a contradiction since a_j is by definition an integer not belonging in any of the sets $A_1, \ldots, A_i, \ldots, A_{j-1}$. We showed that every natural number not greater than $p-1$ is an element in its reduced form in exactly one of the sets A_i, $1 \le i \le k$, which shows that

$$A_1 \cup A_2 \cup \cdots \cup A_k = P.$$

We can derive that for every $n \in P$, $n \equiv a_i \cdot 2^m \pmod p$, $0 \le m \le r-1$, which is equivalent to saying that $n = a_i\cdot 2^m + x\cdot p$ for some integer x. So, $b^n = b^{a_i\cdot 2^m + x\cdot p} = b^{a_i\cdot 2^m}\cdot (b^p)^x \equiv b^{a_i \cdot 2^m} \pmod{\frac{b^p-1}{b-1}}$. The last congruence comes from the observation that $(b^p)^x \equiv 1 \pmod{\frac{b^p-1}{b-1}}$.

The left-hand side of Equation 2 can take the form

$$\prod_{n=1}^{p-1}(b^n+1) \equiv \prod_{i=1}^{\frac{p-1}{r}} \cdot \prod_{m=0}^{r-1}(b^{a_i\cdot 2^m}+1) \left(\bmod \frac{b^p-1}{b-1}\right),$$

and, by using the identity in Equation 1 for $x=b^{a_i}$ and $n=r$, the last product can be written as $\prod_{m=0}^{r-1}(b^{a_i\cdot 2^m}+1) = ((b^{a_i})^1+1)((b^{a_i})^{2^1}+1)\cdots((b^{a_i})^{2^{r-1}}+1) = \frac{(b^{a_i})^{2^r}-1}{b^{a_i}-1}$.

Since $2^r \equiv 1 \pmod p$ and p does not divide evenly into a_i, the congruence $(b^{a_i})^{2^r} -$

$1 \equiv b^{a_i} - 1 \pmod{\frac{b^p-1}{b-1}}$ holds true; this implies that $\frac{(b^{a_i})^{2^r}-1}{b^{a_i}-1} \equiv 1 \pmod{\frac{b^p-1}{b-1}}$. Thus we see $\prod_{m=0}^{r-1}(b^{a_i \cdot 2^m}+1) \equiv 1 \pmod{\frac{b^p-1}{b-1}}$ and we can obtain immediately:

$$\prod_{n=1}^{p-1}(b^n+1) \equiv \prod_{i=1}^{\frac{p-1}{r}} 1 \equiv 1^{\frac{p-1}{r}} \equiv 1 \left(\mathrm{mod}\ \frac{b^p-1}{b-1} \right),$$

completing the proof. ■

We demonstrate the theorem with the following example.

Example. Let $p = 89$ and $b = 2$. The order of 2 modulo 89 is $r = 11$ and $k = \frac{89-1}{11} = 8$.

subsets A_i	a_i's
{1, 2, 4, 8, 16, 32, 64, 39, 78, 67, 45}	
{3, 6, 12, 24, 48, 7, 14, 28, 56, 23, 46}	3
{5, 10, 20, 40, 80, 71, 53, 17, 34, 68, 47}	5
{9, 18, 36, 72, 55, 21, 42, 84, 79, 69, 49}	9
{11, 22, 44, 88, 87, 85, 81, 73, 57, 25, 50}	11
{13, 26, 52, 15, 30, 60, 31, 62, 35, 70, 51}	13
{19, 38, 76, 63, 37, 74, 59, 29, 58, 27, 54}	19
{33, 66, 43, 86, 83, 77, 65, 41, 82, 75, 61}	33

In the above table the numbers $a_2 = 3, a_3 = 5, a_4 = 9, a_5 = 11, a_6 = 13, a_7 = 19$, and $a_8 = 33$ are the least natural numbers that do not appear in the union of the previous A_i's.

REFERENCES

1. L. E. Dickson, *History of the Theory of Numbers. Vol. I: Divisibility and Primality*. Dover, New York, 2005.
2. E. Vantieghem, On a congruence only holding for primes II (2008), http://arxiv.org/abs/0812.2841.

Summary. This article may be regarded as a connection between the past and the present. In 2008 E. Vantieghem proved a new primality criterion whose proof requires the use of cyclotomic polynomials. Almost two and a half centuries ago, Euler used a not-so-well-known method of proof to prove Fermat's little theorem. In this article we show that Euler's method can be adopted in order to prove Vantieghem's result in a more elementary way.

KONSTANTINOS GAITANAS (MR Author ID: 1103989) got his degree at the National Technical University of Athens, in 2010. He is currently a professor at Enomena Frontistiria, a primary school in Volos, Greece. His interests include polynomials, Egyptian fractions, and prime numbers.

ACROSS

1. Big fusses over little problems
5. Lose it
9. Noisy laundry cycle (or a kind of manifold)
13. Merest puff, as of smoke
14. Owie relative
16. Greek letter that doesn't amount to much
17. Geometry's egg
18. Magnetically archived, say
19. Gleefully gossip
20. It models sigmoidal growth (and it's in love)
23. Shorthand for a shorthand transcriptionist
24. Silicon Intel product
25. Sheepish?
29. Academic url ender
31. Not smaller than any other element in the subset
33. Fundraising school group, for short
36. How unit fractions begin
40. Algebraic number theorist Artin, to friends
41. It vibrates in the direction of propagation (and it's in love)
44. Matrix theorist Taussky-Todd, to friends
45. Three-dimensional chiller
46. Bear's lair
47. Scandalous memoir of sorts
49. One of the seven deadly trig functions
51. Harmonic analyst Elias, or a lager glass
52. Latin "author just mentioned"
56. Cotan kin
60. For f, it is defined as f'/f (and it's in love)
63. Hedge network featured in *The Shining*, e.g.
66. Interstellar gas cloud
67. "You're gonna ___ a bigger boat"
68. πr^2 for a circle
69. Call to the ___ (chastise)
70. Shrek, for one
71. Noted logician, in some circles
72. DC baseball team
73. Dampens

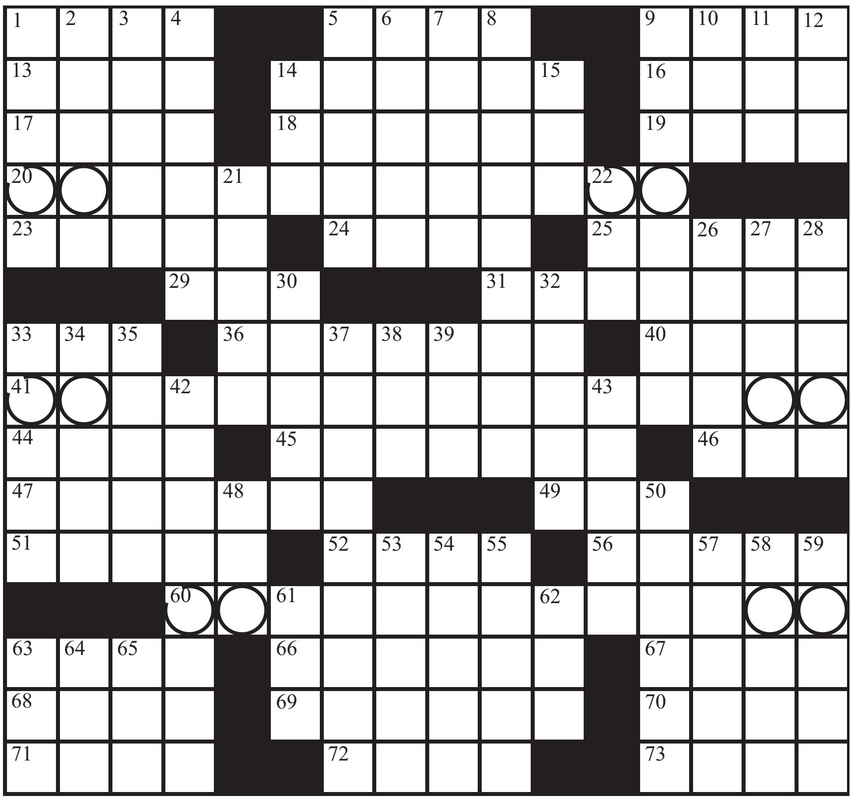
DOWN

1. Base escapists, briefly
2. Dislodged turf clod
3. "August ___ County" (2013 Meryl Streep/Julia Roberts film)
4. Special type of polynomial for interpolation
5. ___ boom
6. Turn it up a ___
7. Early summers?
8. Right-click offering
9. Objects may be closer than they appear in this type of mirror
10. Polynesian taro-based staple
11. "Pure mathematics is, in ___ way, the poetry of logical ideas": Einstein
12. Slangy negative
14. Automated software app
15. "___ the ramparts we watched"
21. Words to second that emotion
22. ___ populi
26. Seether's confession
27. Paul Halmos's "___ Set Theory"
28. Set theorist Mary ___ Rudin
30. ___ the cows come home
32. Women of Yemen, usually
33. Draws points on a coordinate plane
34. Flat rental sign
35. It's between the lines (unless they're parallel)
37. Classical, geometry-wise
38. Expression with a one-variable function and derivatives, briefly
39. Old ___ (London theater)
42. Beyond Copernican
43. German camera brand
48. "That's ___-brainer"
50. "Later, I'm busy"
53. Actress Messing or Winger
54. Blow
55. Road trip accumulation
57. Prolonged military assault
58. Tennis Hall-of-Famer Chris
59. Hands over, legally
61. Vitamin supply chain, briefly
62. Gigantic brewery tub
63. Dallas hoopster, informally
64. "Uneven numbers ___ the gods' delights": Virgil
65. Meditational discipline

Math. Mag. **90** (2017) 73–74. doi:10.4169/math.mag.90.1.73.

Mathematics in Love

TRACY BENNETT
Mathematical Reviews
Ann Arbor, MI
tgp@ams.org



Clues start at left, on page 73. The Solution is on page 57.

Extra copies of the puzzle can be found at the MAGAZINE's website, www.maa.org/mathmag/supplements.

Crossword Puzzle Creators

If you are interested in submitting a mathematically themed crossword puzzle for possible inclusion in MATHEMATICS MAGAZINE, please contact the editor at mathmag@maa.org.

PROBLEMS

Proposals

To be considered for publication, solutions should be received by July 1, 2017.

2011. *Proposed by Souvik Dey (M. Math student), Indian Statistical Institute, Kolkata, India.*

Let S be an open subset of the set $\mathbb{R}$ of real numbers such that S contains at least one positive number and at least one negative number. Show that every real number can be written as a finite sum of (not necessarily distinct) elements of S.

2012. *Proposed by D. M. Bătineţu-Giurgiu, "Matei Basarab" National College, Bucharest, Romania and Neculai Stanciu, "George Emil Palade" School, Buzău, Romania.*

Let f be a continuous real-valued function on $(0, \infty)$ satisfying the identity $f(1/x) = -f(x)$ for all $x > 0$. Given $a > 0$, calculate

$$\int_{\sqrt{2}-1}^{\sqrt{2}+1} \frac{dx}{(1+x^2)(1+a^{f(x)})}.$$

2013. *Proposed by Julien Sorel, PNI, Piatra Neamt, Romania.*

For a positive integer n, let $\mathcal{T}$ be the regular tetrahedron in $\mathbb{R}^3$ with vertices $O(0, 0, 0)$, $A(0, n, n)$, $B(n, 0, n)$, and $C(n, n, 0)$. Show that the number N of lattice points (x, y, z) (i. e., points with integer coordinates x, y, z) lying inside or on the boundary of $\mathcal{T}$ is

$$N = \frac{1}{3}(n+1)(n^2+2n+3).$$

Math. Mag. **90** (2017) 75–82. doi:10.4169/math.mag.90.1.75. © Mathematical Association of America

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must 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.

Effective immediately, 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 hope that this online system will help streamline our editorial team's workflow while still proving accessible and convenient to longtime readers and contributors. We encourage submissions in PDF format, ideally accompanied by LaTeX source. General inquiries to the editors should be sent to* mathmagproblems@maa.org.

2014. *Proposed by Gaitanas Konstantinos, Greece.*

For every integer $n \geq 2$, let $(\mathbb{Z}_n, +)$ be the additive group of integers modulo n. Define an *antimorphism of* $\mathbb{Z}_n$ to be any function $f : \mathbb{Z}_n \to \mathbb{Z}_n$ such that $f(x) - f(y) \neq x - y$ whenever x, y are distinct elements of $\mathbb{Z}_n$. Let an *antiautomorphism of* $\mathbb{Z}_n$ be any bijective antimorphism of $\mathbb{Z}_n$. For what values of n does $\mathbb{Z}_n$ admit an antiautomorphism?

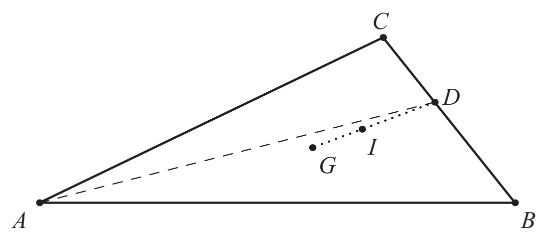
2015. *Proposed by George Stoica, Saint John, New Brunswick, Canada.*

Let K be any field. Let $P(X)$ be any nonconstant polynomial in a single variable X having coefficients in K. If K is a finite field, assume that deg P (the degree of P) is coprime to the characteristic of K. Prove that there exists a polynomial $Q(X)$ with coefficients in K and an integer $m > \deg Q$, such that the polynomial $R(X) = X^m P(X) + Q(X)$ has only simple zeros.

Quickies

Q1067. *Proposed by Abdilkadir Altintaş, Emirdağ, Turkey.*

Let $\triangle ABC$ be a heptagonal triangle with angles $\angle A = \frac{\pi}{7}$, $\angle B = \frac{2\pi}{7}$, $\angle C = \frac{4\pi}{7}$. Let I be the incenter and G the centroid of $\triangle ABC$, and let $\overline{AD}$ be the symmedian of $\triangle ABC$ through A. Show that the points D, I and G are collinear.



Q1068. *Proposed by Wong Fook Sung, Temasek Polytechnic, Singapore.*

A solution (x, y, u, v) to the equation $x^2 + y^3 = u^2 + v^3$ is called *trivial* if either $x^2 = u^2$ and $y^3 = v^3$ or else $x^2 = v^3$ and $y^3 = u^2$. Show that the equation admits infinitely many nontrivial solutions in integers x, y, u, v.

Solutions

A functional equation with quadratic coefficients **December 2015**

1981. *Proposed by Marcel Chirita, Bucharest, Romania.*

Let $n \geq 2$ be an integer and a a real number different from 0 and ± 1. Determine the polynomials $p(x)$ with real coefficients that verify the relation

$$(a^2x^2+1)p(ax) = (a^{2n+2}x^2+1)p(x),$$

for all real numbers x.

Solution by Joseph DiMuro, Biola University, La Mirada, CA.
For any positive integer k, let $q_k(x) = a^{2k}x^2 + 1$, and let $Q(x) = \prod_{m=1}^{n} q_k(x)$. The identity in the statement of the problem reads

$$q_1(x)p(ax) = q_{n+1}(x)p(x). \tag{1}$$

It is straightforward to verify that, for any real constant c, the polynomial $p(x) = cQ(x)$ satisfies (1). We will prove that there are no other solutions.

First, we note that each polynomial $q_k(x)$ is quadratic and irreducible over the reals since $a \neq 0$. For positive integers $k \neq l$, we have $a^{2l}q_k(x) - a^{2k}q_l(x) = a^{2l} - a^{2k}$ is a nonzero constant (since $a \neq 0, \pm 1$), so the polynomials $q_k(x)$ and $q_l(x)$ are coprime.

Let $p(x)$ satisfy (1). Each side of the identity (1) can be factored as a product of polynomials that are irreducible over the reals. Since (1) holds, the irreducible polynomial factors of the left- and right-hand sides must be the same, in some order, up to multiplication by nonzero constants (the ring of polynomials with real coefficients is a unique factorization domain whose units are nonzero constants). The polynomial $q_1(x)$ is a factor of the left-hand side, so it must be a factor of the right-hand side as well. Since $q_1(x)$ and $q_{n+1}(x)$ are coprime, $q_1(x)$ must be a factor of $p(x)$. It follows that $q_2(x) = q_1(ax)$ is a factor of $p(ax)$, so $q_2(x)$ must be factor of both sides of (1). Successively repeating the same argument, we see that, for $m = 1, 2, \ldots, n-1$, if $q_m(x)$ is a factor of $p(x)$, then $q_{m+1}(x)$ is a factor of $p(ax)$ and hence a factor of $p(x)$ (since $q_{m+1}(x)$ is coprime to $q_{n+1}(x)$ when $m < n$). Thus, all the distinct irreducible factors $q_1(x), \ldots, q_n(x)$ of $Q(x)$ are factors of $p(x)$, so we must have $p(x) = r(x)Q(x)$ for some polynomial $r(x)$.

Since $Q(x)$ satisfies the identity $q_1(x)Q(x) = q_{n+1}(x)Q(x)$ and $p(x)$ satisfies (1), $r(x)$ must satisfy the identity $r(ax) = r(x)$. If $r(x)$ is the zero polynomial, we are done; otherwise, the leading term of $r(x)$ is cx^k with $c \neq 0$, and the leading term of $r(ax)$ is ca^kx^k, so we must have $a^k = 1$. Since $a \neq \pm 1$, we must have $k = 0$. In either case, $r(x) = c$ is a constant, so $p(x) = r(x)Q(x) = cQ(x)$ as claimed.

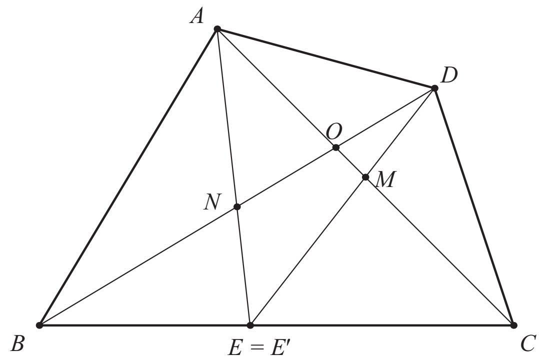
Also solved by Adnan Ali (India), Michel Bataille (France), Brian Bradie, Robert Calcaterra, Eric Egge, Dmitry Fleischman, Eugene A. Herman, Tom Jager, Moubinool Omarjee (France), Achilleas Sinefakopoulos (Greece), and the proposer. There was one incomplete or incorrect solution.

A triangle equivalent to the sum of two others **December 2015**

1982. *Proposed by George Apostolopoulos, Messolonghi, Greece.*

The diagonals of a convex quadrilateral $ABCD$ intersect at a point O such that $OB > OD$, $OC > OA$, and Area($\triangle OAB$) + Area($\triangle OCD$) = Area($\triangle OBC$). Let M and N be the midpoints of the diagonals $\overline{AC}$ and $\overline{BD}$, respectively. Prove that the lines AN and DM intersect at a point that lies on the side $\overline{BC}$.

Solution by Achilleas Sinefakopoulos, Larissa, Greece.



All areas and lengths will be nonnegative (undirected). We use the customary notation $[XYZ]$ to denote the area of triangle $\triangle XYZ$. Since $OC > OA$, the point M lies in the interior of $\triangle BDC$, and since $OB > OD$, the point N lies in the interior of $\triangle BAC$. Thus, the lines AN and DM intersect the side $\overline{BC}$ at points E and E', respectively. We will show that the points E and E' coincide.

From the theorem of Menelaus on triangle $\triangle OBC$ applied to each of the lines ANE and DME', we have

$$\frac{BN}{ON}\cdot\frac{OA}{AC}\cdot\frac{CE}{EB}=1, \quad \text{and} \quad \frac{CM}{OM}\cdot\frac{OD}{DB}\cdot\frac{E'B}{CE'}=1. \tag{1}$$

Since the triangles $\triangle OAB$ and $\triangle OBC$ share their height through B, their areas are in the ratio of their bases: $[OAB]/[OBC] = OA/OC$. Similarly, $[OCD]/[OBC] = OD/OB$. From the hypothesis $[OAB] + [OCD] = [OBC]$, we obtain

$$\frac{OA}{OC}+\frac{OD}{OB}=\frac{[OAB]}{[OBC]}+\frac{[OCD]}{[OBC]}=1 \quad \Rightarrow \quad OA\cdot OB+OC\cdot OD=OB\cdot OC. \tag{2}$$

Since M is the midpoint of $\overline{AC}$, we have $OC - OA = 2\,OM$; from (2), it follows that $OC\cdot OD = (OC - OA)\cdot OB = 2\,OM\cdot OB$. Similarly, we obtain $OA\cdot OB = 2\,ON\cdot OC$. Therefore,

$$\frac{OA}{ON}=2\,\frac{OC}{OB}=4\,\frac{OM}{OD}. \tag{3}$$

Combining (3) with (1) and using the relations $AC = 2\,CM$ and $BN = DB/2$, we get

$$\frac{EB}{CE}=\frac{BN}{AC}\cdot\frac{OA}{ON}=\frac{DB/2}{2\,CM}\cdot\frac{OA}{ON}=\frac{DB}{CM}\cdot\frac{OM}{OD}=\frac{E'B}{CE'}.$$

This proves that the points E and E' both coincide with the point of intersection of AN with DM, completing the proof.

Also solved by Adnan Ali (India), Michel Bataille (France), Robert Calcaterra, Eric Egge, Min Gyu Park (Korea), and the proposer.

Pointwise convergence of iterates of $f(x) = 3 - 1/x$ **December 2015**

1983. *Proposed by Prapanpong Pongsriiam, Silpakorn University, Nakhon Pathom, Thailand.*

Let $y_1 = 0$ and $y_n = 1/(3 - y_{n-1})$ for all $n \geq 2$. Assume that $x_1 \notin \{y_n : n \in \mathbb{N}\}$ and define for $n \geq 2$,

$$x_n = 3 - \frac{1}{x_{n-1}}.$$

Prove that the sequence (x_n) converges.

Solution by Northwestern University Math Problem Solving Group, Evanston, IL.
Note that the inverse of the function $f(x) = 3 - 1/x$ is $f^{-1}(y) = 1/(3 - y)$. The fixed points of these functions (i. e., the solutions to $f(x) = x$) are $\alpha = (3 - \sqrt{5})/2 \approx 0.38$ and $\beta = (3 + \sqrt{5})/2 \approx 2.62$.

If $x_1 = \alpha$, then the sequence (x_n) is constant and hence converges to α. Henceforth, we assume that $x_1 \neq \alpha$. We will prove that $x_n \to \beta$ as $n \to \infty$. From $y_n = 1/(3 - y_{n-1})$, we get

$$y_n - y_{n-1} = \frac{\beta - y_{n-1}}{3 - y_{n-1}} \cdot (\alpha - y_{n-1}) \qquad \text{for } n \geq 2. \tag{1}$$

Since $y_1 = 0 < \alpha < \beta < 3$, it follows from (1) and induction that (y_n) is strictly increasing and bounded above by α. In particular, $y_{n-1} < 3$, so y_n is defined for all $n \geq 1$. Since (x_n) and (y_n) are defined by iteration of the mutually inverse functions f and f^{-1}, a routine inductive argument shows that $x_n = y_1$ if and only if $x_1 = y_n$. Since $x_1 \neq y_n$ for all $n \geq 1$ by hypothesis, we have $x_n \neq y_1 = 0$; in particular, x_n is defined for all $n \geq 1$. Next, consider the sequence (z_n) defined by

$$z_n = \frac{x_n - \beta}{x_n - \alpha} \qquad \text{for } n \geq 1. \tag{2}$$

Note that the assumption $x_1 \neq \alpha$ implies that $x_n \neq \alpha$ for all $n \geq 1$, so (z_n) is well defined. From the recursive definition $x_n = 3 - 1/x_{n-1}$ and the identities $3 - \alpha = 1/\alpha$, $3 - \beta = 1/\beta$, we have

$$\begin{aligned} z_n &= \frac{x_n - \beta}{x_n - \alpha} = \frac{(3 - \frac{1}{x_{n-1}}) - \beta}{(3 - \frac{1}{x_{n-1}}) - \alpha} = \frac{\frac{1}{\beta} - \frac{1}{x_{n-1}}}{\frac{1}{\alpha} - \frac{1}{x_{n-1}}} = \frac{\alpha}{\beta} \cdot \frac{x_{n-1} - \beta}{x_{n-1} - \alpha} \\ &= \frac{\alpha}{\beta} z_{n-1} \approx 0.15 z_{n-1} \qquad \text{for all } n \geq 2. \end{aligned}$$

Hence, $z_n = (\alpha/\beta)^{n-1} z_1 \to 0$ as $n \to \infty$. From (2), we have $x_n = (\beta - \alpha z_n)/(1 - z_n)$, so

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{\beta - \alpha z_n}{1 - z_n} = \frac{\beta - \alpha \cdot 0}{1 - 0} = \beta.$$

Also solved by Adnan Ali (India), Arkady Alt, Michel Bataille (France), Brian Bradie, Stan Byrd, Hongwei Chen, Eric Egge, Joseph DiMuro, David Doster, John N. Fitch, Dmitry Fleischmann, Lixing Han, Jaehyun Kim, Sehyeon Park (Korea), Ioana Mihaila, Henry G. Peterson, Edward Schmeichel, San Francisco University High School Problem Solving Team, Achilleas Sinefakopoulos (Greece), and the proposer. There were four incomplete or incorrect solutions.

A logarithmic improper integral **December 2015**

1984. *Proposed by Timothy Hall, PQI Consulting, Cambridge, MA.*

Evaluate

$$\int_0^\infty \frac{\ln t}{t^2-1}\,dt.$$

Solution by The Bowdoin Math Problem-Solving Seminar, Bowdoin College, ME.
We show that the value of the integral is $I = \pi^2/4$. The change of variables $u = 1/t$ shows that

$$\int_1^\infty \frac{\ln t}{t^2-1}dt = \int_0^1 \frac{\ln u}{u^2-1}du \quad \Rightarrow \quad I = 2\int_0^1 \frac{\ln t}{t^2-1}dt.$$

Since $\ln t = \int_1^t dx/x$, we have

$$\frac{I}{2} = \int_0^1 \frac{\ln t}{t^2-1}\,dt = \int_0^1 \frac{1}{1-t^2}\int_t^1 \frac{1}{x}\,dx\,dt = \int_0^1 \frac{1}{x}\int_0^x \frac{1}{1-t^2}\,dt\,dx.$$

The improper double integrals above are over the region $0 \le t \le x \le 1$. Observe that the integrand is nonnegative, so the exchange of order of integration will be justified once we know the double integral on the right-hand side converges. For $x < 1$, the inner integral may be evaluated by term-wise integration of the geometric series $1/(1-t^2) = \sum_{n=0}^\infty t^{2n}$, giving

$$\int_0^x \frac{dt}{1-t^2} = \sum_{n=0}^\infty \int_0^x t^{2n}\,dt = \sum_{n=0}^\infty \frac{x^{2n+1}}{2n+1},$$

which is (absolutely) convergent for $0 \le x < 1$, justifying the exchange of the sum and integral. Thus,

$$\frac{I}{2} = \lim_{u\to 1^-}\int_0^u \frac{1}{x}\int_0^x \frac{dt}{1-t^2}\,dx = \lim_{u\to 1^-}\sum_{n=0}^\infty \int_0^u \frac{1}{x}\cdot\frac{x^{2n+1}}{2n+1}\,dx = \lim_{u\to 1^-}\sum_{n=0}^\infty \frac{u^{2n+1}}{(2n+1)^2}$$

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

with the exchange of the sum and integral justified by the (absolute) convergence of the latter series, whose value is $I/2 = \sum_{n=1}^\infty n^{-2} - \sum_{n=1}^\infty (2n)^{-2} = (1-1/4)\sum_{n=1}^\infty n^{-2} = (3/4)(\pi^2/6) = \pi^2/8$, so $I = \pi^2/4$.

Also solved by Michel Bataille (France), Brian Bradie, Robert Calcaterra, Hongwei Chen, Ross Dempsey, Eric Egge, Bruce E. Davis, Bill Dunn III, John N. Fitch, Michael Goldenberg & Mark Kaplan, Lixing Han, Christopher Havens, Eugene A. Herman, Tom Jager, Daniel López-Aguayo, Luke Mannion, Andrew Markoe, Donald Jay Moore, Missouri State University Problem Solving Group, Northwestern University Math Problem Solving Group, Moubinool Omarjee (France), Robert Poodiack, San Francisco University High School Problem Solving Group, Edward Schmeichel, Achilleas Sinefakopoulos (Greece), John Zacharias, and the proposer. There were five incomplete or incorrect solutions.

A game with primes **December 2015**

1985. *Proposed by Brooke Tooles (student), Rutgers University, Piscataway, NJ.*

Alice and Bob play the following game. They first agree on a positive integer $n > 1$, called the *target number*, which remains fixed. Throughout, the state of the game is represented by an integer, which is initially 1. Players take turns with Alice going first. Each turn, a player names a prime factor of n, and the state of the game is multiplied by this prime. If a player manages to make the state of the game exactly equal to the target number, n, then that player wins. If the state of the game ever exceeds n, then the game is declared a draw. Determine for which values of n each player has a winning strategy and for which values of n the game should end in a draw.

Solution by San Francisco University High School Problem Solving Group, San Francisco, CA.

We note that the state of the game increases at each turn, so the game always ends in a win or a draw in finitely many steps. By the fundamental theorem of arithmetic, the target number admits a factorization

$$n = p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}$$

as a product of powers of r distinct primes $p_1, p_2, \dots, p_r$ with positive integer exponents $k_1, k_2, \dots, k_r$. Since $n > 1$, there is at least one prime factor p_1, hence $r \geq 1$. Without loss of generality, we assume $1 \leq k_1 \leq k_2 \leq \cdots \leq k_r$. Let $m = k_1 + k_2 + \cdots + k_r$ be the total number of (not necessarily distinct) prime factors of n. The factorization is unique up to order of the factors, so a game leading to a win for some player must end after exactly m turns. If m is odd, Bob has no hope of winning since he only plays in even turns, so his aim is to sabotage Alice's efforts to win; if m is even, Alice cannot win, so her aim is to sabotage Bob.

By the uniqueness of factorization, the game will necessarily end in a draw whenever it reaches a state divisible by $p_j^{k_j+1}$ for some prime factor p_j of n. We prove that the saboteur has a successful strategy leading to a draw under the assumption that n has $r \geq 3$ distinct prime factors. If m is even, Alice is the saboteur, and she forces the draw by naming p_1 on her every turn: Assuming this is the case, since the number of prime factors of n is $m \geq k_1 + k_2 + k_3 \geq 3k_1$, the state of the game will be the product of $2k_1$ prime factors p_j, including at least k_1 factors p_1, right after the end of the $(2k_1)$-th play (Bob's play). The state cannot be equal to n after $2k_1$ plays since $m > 2k_1$; thus, the choice of p_1 by Alice in the $(2k_1 + 1)$-st play will lead to a new state divisible by $p_1^{k_1+1}$, forcing an eventual draw. If m is odd, Bob is the saboteur, and he adopts a similar strategy with a small adaptation that takes into account Alice's first play. Namely, if Alice's first named prime is p_j with exponent $k_j = k_1$ (least among all exponents), then Bob successfully forces a draw by choosing the same p_j on his every turn (the argument above shows that the state of the game will be divisible by $p_j^{k_j+1}$ after $2k_j$ plays since $k_j = k_1$). Otherwise, if Alice's first named prime is p_j with $k_j > k_1$, Bob forces a draw by always choosing p_1 on his turn: After the $(2k_1 + 1)$-st move (Alice's play), the state is a product of $2k_1 + 1$ factors, at least k_1 of which are p_1; moreover, Alice cannot have yet won since $m \geq 3k_1 + 1$ (since $k_j > k_1$ for some j, and $r \geq 3$), whereas the state is a product of only $2k_1 + 1$ prime factors at this point. Bob's choice of p_1 in the $(2k_1 + 2)$-th play leads to a state divisible by $p_1^{k_1+1}$, forcing an eventual draw.

If $r = 1$, so $n = p_1^{k_1}$, there is no strategy to the game at all since p_1 is chosen in every turn: Alice wins if $m = k_1$ is odd; Bob wins if $m = k_1$ is even.

Finally, assume $r = 2$, so $n = p_1^{k_1} p_2^{k_2}$ with $1 \leq k_1 \leq k_2$. Bob has a winning strategy if $k_2 = k_1$: He can always name the prime other than that just chosen by Alice. After the $(2k_1)$-th play, the state of the game is $(p_1 p_2)^{k_1} = n$, so Bob wins. Alice has a winning strategy if $k_2 = k_1 + 1$: She can name p_2 in her first turn; in her subsequent turns, she names the prime other than that just chosen by Bob. After the $(2k_1 + 1)$-st play, the state of the game is $p_2(p_1 p_2)^{k_1} = n$, so Alice wins. If $k_2 - k_1 \geq 2$, then the saboteur's strategy of always choosing p_1 in his/her turn forces a draw.

In summary, Alice has a winning strategy iff $n = p^{2k-1}$ or $n = p^k q^{k+1}$, while Bob has a winning strategy iff $n = p^{2k}$ or $n = p^k q^k$ for some $k \geq 1$ and different primes $p \neq q$. In all other cases, there is a strategy ensuring a draw.

Also solved by Robert Calcaterra, Hyun Jun Cho (Korea), John Christopher, Eric Egge, Joseph DiMuro, Dmitry Fleischman, Patrick Flynn, Tom Jager, Northwestern University Math Problem Solving Group, Pittsburg State University Problem Solving Group, Edward Schmeichel, Achilleas Sinefakopoulos (Greece), Stuart V. Witt, and the proposer. There was one incomplete or incorrect solution.

Answers

Solutions to the Quickies from page 76.

A1067. Let $a = BC$, $b = AC$ and $c = AB$. Introduce baricentric coordinates: $A(1{:}0{:}0)$, $B(0{:}1{:}0)$, and $C(0{:}0{:}1)$. We have $G(1{:}1{:}1)$, $D(0{:}b^2{:}c^2)$, and $I(a{:}b{:}c)$. It is well known that the sides a, b, c of a heptagonal triangle satisfy $c^2 - a^2 = bc$, $c^2 - b^2 = ab$, and $b^2 - a^2 = ac$. Without loss of generality, assume that $\triangle ABC$ has circumradius 1, and let $\theta = \pi/7$. We have $a = 2\sin\theta$, $b = 2\sin 2\theta$, and $c = 2\sin 4\theta$ by the law of sines. Thus, $b - a = 2(\sin 2\theta - \sin\theta) = 4\sin(\theta/2)\cos(3\theta/2)$, and similarly, $b + a = 4\sin(3\theta/2)\cos(\theta/2)$, so $b^2 - a^2 = (b - a)(b + a) = 4\sin(\theta/2)\cos(\theta/2) \cdot 4\sin(3\theta/2)\cos(3\theta/2) = 2\sin\theta \cdot 2\sin 3\theta = ac$. The relations $c^2 - b^2 = ab$ and $c^2 - a^2 = bc$ are obtained similarly by letting $\theta = 2\pi/7$, $\theta = 3\pi/7$, respectively. Thus,

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ 0 & b^2 & c^2 \end{vmatrix} = c^2(b - a) - b^2(c - a) = b(b + a)(b - a) - a(c + a)(c - a)$$

$$= b(b^2 - a^2) - a(c^2 - a^2) = bac - abc = 0,$$

so D, I, and G are collinear.

A1068.
For any number k, let

$$x = 1 + \frac{9k(7k + 5)}{2} \qquad u = 8 + \frac{9k(7k + 5)}{2}$$

$$y = 4 + 7k \qquad v = 1 + 7k.$$

The verification that x, y, u, v satisfy $x^2 + y^3 = u^2 + v^3$ is straightforward; moreover, x, y, u, v are integers whenever k is an integer. We have $x^2 = u^2$ and $y^3 = v^3$ for at most finitely many values of k, namely the roots of $441k^2 + 315k + 63$ $(= u^2 - x^2 = y^3 - v^3)$. Similarly, $x^2 = v^3$ and $y^3 = u^2$ hold for at most finitely many values of k, namely the roots of $3969k^4 + 4298k^3 + 1689k^2 + 96k$ $(= 4(x^2 - v^3) = 4(u^2 - y^3))$. Therefore, as k varies over the integers, we obtain infinitely many nontrivial solutions. (It is routine to verify that the only trivial solution in integers is obtained when $k = 0$, namely $(1, 4, 8, 1)$.)

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.

Thompson, Avery, 5 simple math problems no one can solve,
http://www.popularmechanics.com/science/g2816/5-simple-math-problems/.

Romik, Dan, The moving sofa problem,
https://www.math.ucdavis.edu/~romik/movingsofa/.

Calculus textbooks often include an optimization problem about the longest pole (of no thickness) that can be carried around a corner in a hall. Of course, you can carry a longer pole if you tilt it into the third dimension (a point usually not addressed in the textbook). But what if the "pole" is a sofa, of rectangular or other shape? Even sticking to two dimensions, the answer to the "moving sofa problem," proposed by Leo Moser in 1966, is still unknown. Also unsolved so far: Is there a perfect cuboid (all lengths and diagonals are integers)? Does every plane simple closed curve have an inscribed square? How many points in the plane are needed to be sure that n of them outline a convex n-gon? And then, of course, there's the $3x + 1$ ("Collatz") problem. These are all great instances that you can relate to students and friends to demonstrate how mathematics is not yet "finished."

Lagarias, Jeffrey C., Erdős, Klarner, and the $3x + 1$ problem, *American Mathematical Monthly* 123 (October 2016) 753–776.

Simple functions can produce intriguing questions and considerable theory, as with quadratic functions in dynamical systems and linear functions in number theory. The latter functions are also termed *integer affine maps*, and it is the study of the structure of their orbits that attracted Erdős and others. Author Lagarias presents a history of their efforts, ranging from self-orthogonal Latin squares to orbits of an affine semigroup. He explains how the $3x + 1$ ("Collatz") problem can be framed in terms of the latter—but doing so requires allowing noninteger rational coefficients in the affine functions. A simpler problem, which substitutes $(x + 1)/2$ for $3x + 1$, can be solved in terms of such a formulation. So there may be hope for the $3x + 1$ problem, which Erdős termed "hopeless."

Villani, Cédric, *Birth of a Theorem: A Mathematical Adventure*, Farrar, Straus and Giroux, 2015; 260 pp, \$26, \$15(P). ISBN 978-086547767-4, 978-037453667-1.

This book relates the experience of a Fields Medalist, from diving into a research topic, through several years, to publication of the resulting theorem. It is Villani's answer to "what it's like to be a mathematician—what a mathematician's daily life is like." Included are pages and pages of obscure mathematics from his papers, emails with his research partner, a long account of his favorite music, vignettes expositing concepts in mathematics, brief biographies of mathematicians he meets, and lots of lecturing and travel for him (and moving for his family), and dealing with the demands of fame. He writes engagingly; you don't need to understand the details of the mathematics (you can't). What comes through is the passion.

Math. Mag. **90** (2017) 83-84. doi:10.4169/math.mag.90.1.83.

Colquhoun, David, The perils of p-values: Why more discoveries are false than you thought, *Chalkdust* 4 (Autumn 2016) 32–38.

Colquhoun, David, An investigation of the false discovery rate and the misinterpretation of p-values, investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society of Open Science* 1 (2014): 140216. http://dx.doi.org/10.1098/rsos.140216 .

The statistical practice and philosophy of p-values has been under assault for years. Some have emphasized effect size instead (e.g., Stephen T. Ziliak and Deirdre N. McCloskey in *The Cult of Statistical Significance*, 2008); "p-hackers" have violated both the letter and the spirit of the methodology by picking cherries from an orchard of p-values; the public has always been justifiably confused about what "significance" means; and some researchers—whose favorite research hypotheses just don't sink below .05, acting as "sore losers" in the "statistics game"—have tried to banish p-values from journals. Author Colquhoun points out that a p-value does not answer—or even address—the key question: Is the result true (e.g., is Drug X better than placebo)? A p-value doesn't even offer an estimate of how likely it is true. Colquhoun goes on to claim that "p-values exaggerate the strength of the evidence": A claim based on a p-value of .05 has a 30% chance of being a "false discovery." He bases that calculation on a presupposed probability of 10% of a "real effect," offering a tree diagram like those used for assessing the predictive value of a medical or other test. Colquhoun admits that the "prevalence" of true real effects is unknown and unknowable, but asserts that it would not make sense for it to be greater than 50% (if so, then why experiment?); and a Bayesian might have no qualms about assigning such a probability.

Stigler, Stephen M., *The Seven Pillars of Statistical Wisdom*, Harvard University Press, 2016; 230 pp, $22.95(P). ISBN 978-0-674-08891-7.

Stigler's title echoes *The Seven Pillars of Wisdom* by T. E. Lawrence ("Lawrence of Arabia") but harks further back to an origin in Proverbs 9:1. Stigler, a renowned historian of statistics, assesses the "disciplinary foundations" (often dating back to antiquity) of statistics: aggregation (data summaries), information (e.g., the $\sqrt{n}$ rule), likelihood (probability), intercomparison (partitioning of variance, bootstrapping), regression, design (of experiments), and residual (model diagnostics, graphical displays). The book offers a sense—often missing in introductory courses—of statistics as a discipline, an art, and a philosophical approach to striving for truth.

Katz, Victor J. (ed.), *Sourcebook in the Mathematics of Medieval Europe and North Africa*, Princeton University Press, 2016; xvi + 574 pp, $95. ISBN 978-0-691-15685-9.

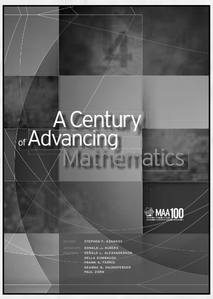
This volume of translated excerpts from key texts in Latin, Hebrew, and Arabic covers the period 800–1450. The authors focus not just on extended excerpts but on the interchange among cultures, as illustrated by the treatment of the same problem in varied sources. (I wish that the index were more extensive; the terms *al-jabr* and *al-muqābala* and other transliterated terms do not appear in it.)

Lipton, R. J. and Regan, K. W., Going for two, https://rjlipton.wordpress.com/2016/10/16/going-for-two/ .

American football offers options to go for 1 point (kick) or 2 points (run or pass) after a touchdown. The dilemma of what a team should do when behind by 14 points is sometimes analyzed in probability classes with the help of a tree diagram of possible outcomes and accompanying estimated probabilities, and coaches often use a chart that suggests strategy for various score situations. Here Richard Karp (of P vs. NP fame) offers a new "Fundamental Theorem of Football," for optimization from the beginning of a game: *Always go for 2. If after* $2t - 1$ *tries you have succeeded* t *times (so that you are then ahead of what just kicking would have brought), switch to kicking.* Of course, this statement is just the conclusion of the theorem; the hypotheses are that the kick is always successful (virtually true in the National Football League), the 2-point conversion succeeds half of the time (true there recently), and you will be scoring lots of touchdowns (e.g., unboundedly many). And there are caveats: "So if you miss a few two-point tries and the game is deep into the second half, you're left holding the bag of foregone points." The "theorem" does not address changing the strategy based on the score, nor the game-theoretic aspect of interaction with the opponent's strategy; and the proof is left to the sports fan reader.

Congratulations to MAA Award-winning Books of 2016!

Selected by ALA CHOICE Magazine as Outstanding Academic Books

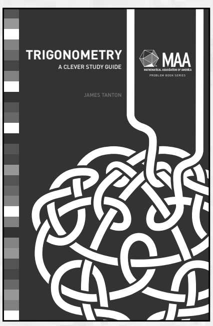


A Century of Advancing Mathematics

Stephen Kennedy, Editor
Donald J. Albers, Gerald L. Alexanderson, Della Dumbaugh, Frank A. Farris, Deanna B. Haunsperger, & Paul Zorn, Associate Editors

Catalog Code: CAM
392 pp., Hardbound, 2015
ISBN: 978-0-88385-588-1
eBook, Free with MAA membership

List: $60.00
MAA Member: $45.00



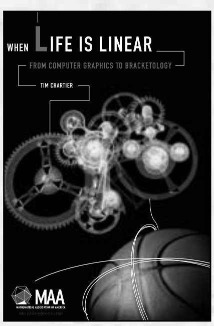
Trigonometry: A Clever Study Guide

By James Tanton

Catalog Code: CLP-1
232 pp., Paperbound, 2015

List: $19.95
ISBN: 978-0-88385-836-3

MAA Beckenbach Book Prize Winner



When Life is Linear: From Computer Graphics to Bracketology

By Tim Chartier

Catalog Code: NML-45
140 pp., Paperbound, 2015
ISBN: 978-0-88385-649-9

List: $52.00
MAA Member: $39.00

Get these award-winning titles and more.
Call 1-800-331-1622 or order online at:
store.maa.org/site.

MAA PRESS

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

CONTENTS

ARTICLES

PROBLEMS AND SOLUTIONS

REVIEWS