LETTER FROM THE EDITOR

With this issue, I complete my first year as editor of *Mathematics Magazine*, and we close out 2020 with our usual bumper crop of expository excellence.

Our lead article comes from Rick Mabry and Laura McCormick. They solve an interesting open problem in knot theory. Specifically, they show, by construction, that there is a Brunnian link whose minimal projection is a simple symmetric Venn diagram of order seven. The jargon is all explained in the article, but you may know more about it than you realize. You are probably familiar with the Borromean rings, which involve three linked circles. This link is Brunnian because if you remove any one of the circles, the other two are unlinked. And if you project the rings to a plane, the result is the standard Venn diagram on three circles. Mabry and McCormick are basically doing *that*, but with seven links instead of three.

The Borromean rings arise again in Dave Auckly's article, where they are linked, pardon the pun, with the icosahedron and the Poincaré homology sphere. Auckly spins a fascinating, and wonderfully lucid, tale through the thickets of algebraic topology and three-dimensional geometry. Any paper that involves a flying dog is just fine with me!

Euclidean geometry is also well-represented in this issue. Andrés Navas presents a charming and original proof of the Pythagorean theorem. Dixon J. Jones, for his part, proves some surprising results about the altitudes and perpendicular bisectors of triangles. Euclidean geometry seems to be a bottomless pit of interesting mathematics. People have been studying it for millennia, but it seems there is still more to find.

If calculus and analysis are more to your liking, then have a go at the article by John Engbers, Andy Hammett, and Ian Hogan. They present a novel approach to hyperbolic trigonometry. This is a subject that has largely disappeared from the standard calculus sequence, which is a pity, since it is really quite beautiful. Sticking with calculus, the short note by Leonard Van Wyk will make you smile. He points to an amusing situation in which a combination of two common wrong approaches to a differentiation problem somehow leads to the correct answer.

Marcin Mazur generalizes a challenging Putnam problem. Challenging, mind you, by *Putnam* standards! Matthew Durey develops the proper strategy for using your life lines on *Who Wants to be a Millionaire*. Roger Nelsen considers almost equilateral Heronian triangles, meaning triangles with integer area whose sides are consecutive integers. And Jeremy Alm wraps up this month's articles with a probability-based derivation for the formula for the sum of a geometric series.

We close out the issue with Problems, Reviews, Proofs Without Words, and the presentation of the annual Allendoerfer Awards. Fun for the whole family!

It has been an interesting challenge for me to learn how to do this job at the same time that life beyond the magazine has intruded more and more. Let me take one more opportunity to thank my editorial board, and all the people at the MAA who work behind the scenes to make this magazine happen. Even with all the distractions, the magazine has made significant progress with regard to efficiency and smooth operation. I look forward to extending this progress in 2021, and to continue bringing you the first-rate mathematical exposition for which the magazine is known.

Jason Rosenhouse, Editor

Math. Mag. **93** (2020) 322. doi:10.1080/0025570X.2020.1819017

Branko's M2 Is Brunnian*

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

LAURA MCCORMICK
Discovery High School
Lake Alfred, FL 33850
lmccormick83@gmail.com

Dedicated to the late Branko Grünbaum (1929–2018), whose friendly humor, patience, and encouragement will be missed. With this note, Branko is further Venn-erated for his discovery of M2.

A problem and its proof

Problem. *Prove that there exists a Brunnian link whose minimal projection is a symmetric Venn diagram of order 7.*

Proof Without Words:

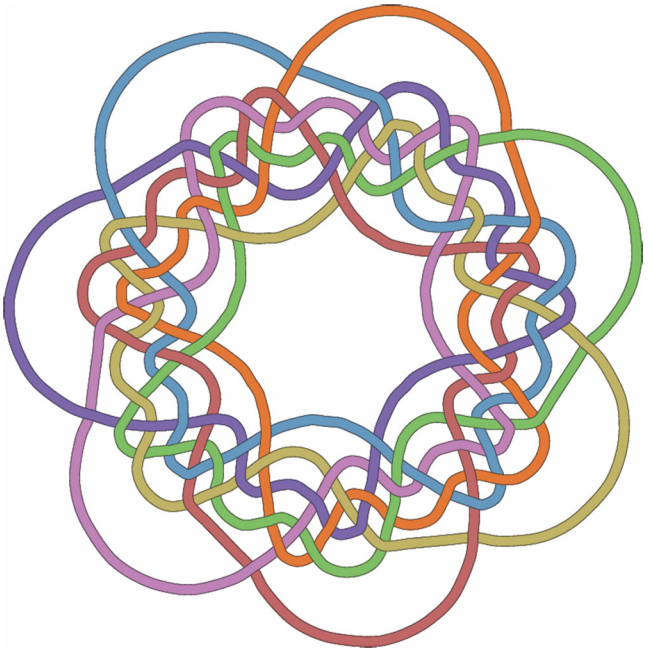


Figure 1 The link "AltM2."

*The online version of this paper has color graphics that may be easier to follow.

Math. Mag. **93** (2020) 323–329. doi:10.1080/0025570X.2020.1823767
MSC: 57M25, 05C10, 05A18

HINT WITHOUT WORDS:

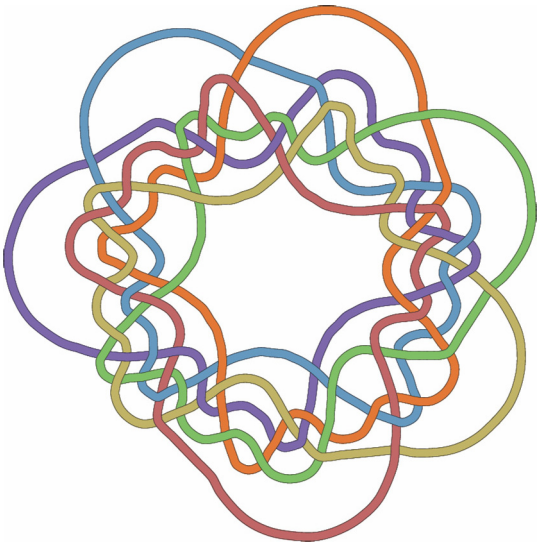


Figure 2 The link "AltM2-minus-one."

Remarks with words

It would go without saying (if no words were to be used) that the statement of the problem requires us to show that the following conditions are satisfied.

C1 The link in Figure 1 is Brunnian.
C2 Its 2-dimensional rendering is a minimal projection.
C3 The projection is a symmetric Venn diagram of order 7.

But the words above, if used, should be defined! We'll briefly sketch the ideas needed, starting with three dimensions and ending with two.

Before doing so, it certainly helps to mention a very familiar, less complicated example: the Borromean rings (See Figure 3). We'll put it the same way:

Prove that there exists a Brunnian link whose minimal projection is a symmetric Venn diagram of order 3.

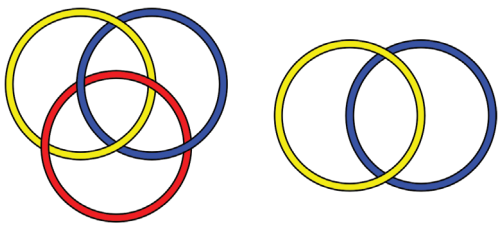


Figure 3 A proof without words (left) and a hint without words (right) for a simpler version of the opening problem.

We'll get to the cryptic title of this note in due course. First, we discuss the main ideas relevant to the problem itself.

Background

Regarding condition C1, the Brunnian and other 3-dimensional notions used here belong to a component of *knot theory*. Rigorous and detailed versions of these ideas can be found, for example, in the (highly recommended) book by Colin Adams [**1**]. For us, it suffices to say that a *knot* is a non-self-intersecting closed curve in $\mathbb{R}^3$ and a *link* is a finite union of such knots, no two of which intersect. We can assume for our special purposes that the curves are nice, smooth or polygonal being nice enough.

Each of the individual closed curves comprising a link is called a *component* of the link. Two links are considered equivalent if one can be continuously deformed into the other without causing any of the intermediate components to intersect one another or themselves. A link with n components is called *unlinked* if it is equivalent to a link of n components, each of which can be contained in its own sphere (or box—any pairwise disjoint open containers will do). A link of n components is called *trivial* if it is equivalent to n unlinked circles in $\mathbb{R}^3$. Such a link is often called an *unlink*, which then consists of n *unknots*.

Finally, a link of n components is called *Brunnian*, after Hermann Brunn, 1862–1939, if it is linked, but every sub-link of $n-1$ components is unlinked. Loosely speaking, the link is Brunnian if when you pick it up, you can't shake, wiggle, or pull it apart without breaking things, but when any one component is removed, the remaining components can slip apart.

The Borromean rings clearly constitute a Brunnian 3-component link. Two linked components rather trivially form a 2-component Brunnian link. (The simplest of these is called a *Hopf link*, equivalent to two linked circles.) Brunn himself gave a recursive construction of n-component Brunnian links for all n [**3**]. Brunn's, and other recursive and infinite families of Brunnian links, are shown in Jablan [**8**], which can be easily found online. You can also check out the fashions by Adams, Fleming, and Koegel [**2**].

With that as background, let's return to our opening diagrams. Figure 1 claims to be a Brunnian link of 7 components. Figure 2 is what we obtain when we remove one of the components. (We have nicknamed the first figure "AltM2" for reasons to be clarified later.)

Moving on to condition C2, the printed figures we are discussing are, of course, plane renderings of spatial objects. We needn't get too technical here, but this passage from three to two dimensions is important for our particular result.

Let us concede that the renderings of the components in the figures are certainly beefed-up versions of one-dimensional closed curves! But no apologies are needed for such embellishments, since without some such thickening, we could never pretend to view curves at all. Ignoring this aspect of such renderings (as we always do), what we are looking at are renderings of typical orthogonal *projections* of curves in $\mathbb{R}^3$ onto a plane. In such pictures, we include extra information that let's us visualize which strands go over or under each other when viewed from a chosen direction. Such representations are usually called *link diagrams*.

Now, when non-intersecting curves in $\mathbb{R}^3$ are projected onto a plane, the resulting plane curves can suddenly intersect. Such a point of intersection is called a *multiple point*. The multiplicity in question is the number of points on the space curves that project to the multiple point. In general, that number can be infinite. But we will soon be assured that the number will never be greater than two in our case.

Naturally, a given union of nonintersecting space curves (like our link) can produce different numbers of multiple points depending on the particular projection. In fact, it could happen that a multiple point is the projected image of *three or more* points in the link. We expressly forbid that and require things be oriented so that any multiple points

are *double points*. That is, their multiplicities must be 2. Furthermore, we require that at such a double point the associated plane curves must cross *transversely*. This means no tangent curves, no touching of two different projected polygons at their vertices; no kissing whatsoever is allowed (as certain Venn diagrammarians have put it). These double points are then called *crossings*. We want only finitely many crossings. With the smooth or polygonal curves that we stipulated, such *regular projections* always exist, and are even hard to avoid [**4**].

This brings us to the next definition: a projection of a link is a *minimal projection* if it results in the fewest possible crossings.

Observe also that our link AltM2 is an *alternating link*, which means that as we traverse any curve in the link, the crossings in the projection (more precisely, in the link diagram) always alternate: over, under, over, under, etc. (The "Alt" part of "AltM2" is now revealed.)

Now that we are looking at intersecting plane curves, we get to condition C3, concerning Venn diagrams. Venn himself gave a recursive construction of n-component Venn diagrams for all n [**14**]. We note that if a link is projected to a Venn diagram, then the components of the link are all actually unknots, since the curves comprising a Venn diagram are required to be closed curves with no crossings. That is, they are curves homeomorphic to circles.

Here we are interested in *symmetric Venn diagrams of order n*. The order is simply the number of sets depicted in the Venn diagram, while symmetry here means *rotational symmetry* about a point (center). More precisely, we require symmetry that can be obtained from the diagram by a continuous transformation of the plane. (The symmetry is important only in a "geombinatoric" sense.) For symmetric Venn diagrams of order n to exist, it is necessary that n be prime [**15**]. A comprehensive reference for symmetric Venn diagrams is the article by Ruskey and Weston [**11**].

The last definition we need is that of a *simple* Venn diagram, which is one for which all crossings (of the planar curves representing the regions) are transverse. There are indeed non-simple diagrams and they figure prominently in the history of symmetric Venn diagrams. But since we are looking at links with regular projections that are Venn diagrams, we certainly must require those diagrams to be simple. It is not yet known if simple symmetric Venn diagrams exist for all prime n.

The proof

We can finally move on to the proof without words!

The easiest part is condition C3, the symmetric Venn-ness of the projection. For this, we do not consider the projection of AltM2 to be a link diagram, just a true projection of one-dimensional curves. The symmetry of the projection is clear. The Venn-ness of the projection can easily be verified by hand using an appropriate labeling of its $128 = 2^7$ regions. (It is convenient to use 7-digit binary representations of the integers 0 through 127 for this.)

For Brunnianism C1, we need to verify two things:

(a) Unlinked-ness: This is proved starting with the 6-component "Hint" in Figure 2, where a single component has been removed from AltM2. It requires nothing but patience to verify that the resulting link is actually unlinked. (Use some yarn. Or a computer graphics program. Better still, some greased plastic tubing. Upload the video to YouTube.) Notice also that symmetry implies that the removal of just one of the seven links is sufficient. *That is because the figure AltM2 is fully three-dimensionally*

rotationally symmetric. (The projected Venn diagram inherits its 7-fold rotational symmetry from that of AltM2.) A link with such 3D symmetry—a *periodic link*—was not required in the statement of the problem, but it is a sweet bonus.

(b) Linked-ness: The 7-component link is itself not unlinked. This might be very difficult to prove, were it not for a famous theorem, one of the three famous *Tait conjectures* presented at the turn of the twentieth century by Peter Guthrie Tait [**13**] and proved by Thistlethwaite, Kauffman, and Murasugi in 1987. The one we need states that *any reduced diagram of an alternating link has the fewest possible crossings.* See, for instance, Kauffman [**9**, p. 212]. For a very different, very recent proof, see [**5**]. This assures us that the 126 crossings shown in the projection of our rendering cannot be reduced, and so certainly cannot be reduced to 0, which would be the case if the seven components could be unlinked.

Okay, but now what is a *reduced diagram*? Let's not get too technical and leave it at this: the usual "figure eight" **8** is not reduced, but with a twist it becomes so: **O**. Our link diagram of AltM2 is clearly reduced, since its components are projected to topological circles.

In any case, knowing that Tait's conjecture is true makes it possible to wordlessly decide by inspection whether small, reduced diagrams of alternating links are indeed linked or not: *A reduced alternating link is linked if and only if it has crossings.*

In the course of demonstrating Brunnianism (condition C1), we also verified that ours is a minimal projection, thereby fulfilling C2.

With that, the proof is finished.

Notes and history

Branko and M2. The figure AltM2 solves the seventh "Open Problem" of Ruskey and Weston [**11**]. (It is the 7th problem on the 2005 revision, which is the most recent as of this writing. The problem first appeared in the 1998 revision—the first author proposed it!)*

It is not difficult to verify that the projection of AltM2 is equivalent to the diagram called "M2" in Ruskey and Weston [**11**], one of the 56 known simple symmetric Venn diagrams on seven sets. (Click your way to the portion, *Symmetric Diagrams / small n / n = 7 / Symmetric simple monotone Venn diagrams without polar symmetry*.) The "Tutte embedding" shown there was used to generate AltM2 by marking coordinates of the crossings and using *Mathematica* to generate interpolating curves, albeit clumsily.

In 1976, Branko Grünbaum was awarded the MAA Lester R. Ford prize for his expository paper about Venn diagrams [**6**]. In it he conjectured that no simple symmetric Venn diagram on 7 sets could be found. Over a decade later he was proven wrong—by Branko Grünbaum!—with his own discovery of M2 [**7**]. In that paper, he reversed course and conjectured that simple symmetric Venns of all prime orders exist, a problem that so far remains unsolved.

The case of $n = 5$. For $n = 5$ we have the links shown in Figure 4, filling the gap between $n = 3$ and $n = 7$.

The alternating symmetric link on the left in Figure 4 projects to the only simple symmetric Venn diagram of order 5, which we will call "V5." Grünbaum was the first to discover V5, which he showed using five congruent ellipses [**6**]. Later he gave a construction with five congruent equilateral triangles [**7**]. The Brunnian connection to the Venn diagram seems to have first been noted in 1996 by Mabry, who then learned

*See http://www.combinatorics.org/files/Surveys/ds5/VennOpenEJC.html

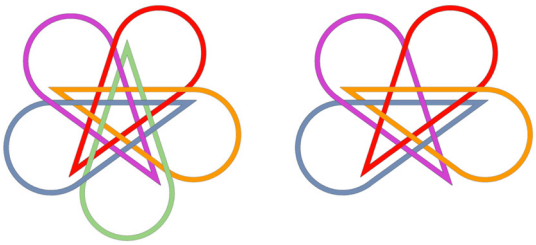


Figure 4 A proof without words and a hint without words for the case $n = 5$.

from Frank Ruskey that the Brunnian link itself had been noted by Ian Stewart [12, p. 106]. The figure had been sent to Stewart by C. van de Walle of Evreux "in a letter, no explanation of the source" (private comm., 1997). See *Symmetric Diagrams / small n* in Ruskey and Weston [11] for more.

Symmetric Venns of order $n > 7$. Khalegh Mamakani and Frank Ruskey have recently constructed simple symmetric Venn diagrams of orders 11 and 13 [10]. That breakthrough automatically opens the problem of whether there are Brunnian examples among them. Given that there are "more than 200,000" unearthed examples of order 11, how hard could it be find one that is Brunnian? (Conjecture: Very.)

Personal notes. In 1996, a doodle by the first author led to his independent discovery of the 5-Venn in Figure 4. After realizing it was indeed a Venn diagram, he searched the literature, finding the works by Branko Grünbaum and Frank Ruskey. He learned that the Venn diagram of order 5 was already known and that simple symmetric Venn diagrams of order 7 exist. Given the realization that the alternating versions of the symmetric Venns for $n = 2, 3, 5$ produce Brunnian links, it was imperative to check the order 7 cases! First examined were the 23 simple symmetric *monotone* Venn diagrams of order 7 labeled P1–P6 and M1–M17 in Ruskey and Weston [11] (where 'P' is for *polar* and 'M' is for monotone-but-not-polar). Some of these were mercifully easy to check, but only because they are so quickly seen *not* to be Brunnian, because of the obvious linkedness of certain *pairs* of curves. Checking pairs was also relatively easy to automate, but the rest seemed hard and would have to wait ...

In 2009, the problem was mentioned by the first author in an undergraduate "Friday Seminar" at LSU Shreveport, at which the second author, then an undergraduate student, was present and intrigued. This breathed new life into the problem. The first author over-reached and promised the second that if the problem could be solved, its rendering would surely be printed on the cover of a mathematics magazine. The second author taught the first (and subsequently the seminar) about link invariants. In 2011 the problem was ultimately solved using a combination of computer automation and eyestrain. What we discovered can be stated as follows. It turns out that Branko's M2 is very special, indeed.

Claim. *(a) Among all the 23 simple symmetric monotone 7-Venn diagrams, M2 alone produces a Brunnian link when alternated. (b) None of the so-far known non-monotone examples N1–N33 are Brunnian when alternated. (c) Among all 56 of the known simple symmetric 7-Venns and all of their 3-dimensionally symmetric link diagrams, only the alternating version of M2 is Brunnian.*

Regarding (b), it is presently unknown if there exist more than the 33 non-monotones specimens listed in [11], where it is conjectured not. As for (c), every simple symmetric 7-Venn can be woven to generate $2^{17} = 131\,072$ distinct, 3-

dimensionally symmetric link diagrams. The corresponding links themselves, which are 7-periodic, needn't be distinct. Parts (b) and (c) were realized in 2019.

On March 8, 2011, our discovery of AltM2 was sent to a handful of mathematicians known or suspected to have some interest in such things. That presentation was dedicated "to Branko Grünbaum, the father of symmetric Venn diagrams and discoverer of M2." Branko replied the next day with his typical enthusiasm and good humor. Included were a few words about a wordless proof: "Once you have your drawing, as you say, the verification is simple," and he signed off as *"papa" Branko*.

Supplement. The first two statements of the Claim in the previous section can be straightforwardly checked. An online search should turn up a "Supplement" prepared by the authors that contains useful details enabling such verification, which can be done by hand and eye (and some labor). Included there is also a more leisurely version of our proof without words, revealed as a long sequence of images in which the link AltM2-minus-one in Figure 2 is unraveled step-by-step. Other related items are also included, such as a *non*-symmetric Brunnian link that projects to V5 and likewise another that projects to M2. A few related questions are posed.

Acknowledgment The first author is grateful to the Sklar Foundation for generous support during the years 2010–2015.

REFERENCES

[1] Adams, C. C. (2004). *The Knot Book*. Providence, RI: American Mathematical Society.
[2] Adams, C., Fleming T., Koegel, C. (2004). Brunnian clothes on the runway: not for the bashful. *Amer. Math. Monthly.* 111(9): 741–748. doi.org/10.1080/00029890.2004.11920138
[3] Brunn, H. (1892). Über Verkettung. *Sitzungsber. Bayer. Akad. Wiss. Math. Naturwiss. Abt.* 22: 77–99.
[4] Crowell, R. H., Fox, R. H. (1963). *Introduction to Knot Theory*. Boston: Ginn and Co.
[5] Greene, J. E. (2017). Alternating links and definite surfaces. *Duke Math. J.* 166(11): 2133–2151. doi.org/10.1215/00127094-2017-0004
[6] Grünbaum, B. (1975). Venn diagrams and independent families of sets. *Math. Mag.* 48(1): 12–23. doi.org/10.2307/2689288
[7] Grünbaum, B. (1992). Venn diagrams II. *Geombinatorics.* 2(2): 25–32.
[8] Jablan, S. V. (1999). Are Borromean links so rare? *Forma.* 14(4): 269–277.
[9] Kauffman, L. H. (1988). New invariants in the theory of knots. *Amer. Math. Monthly.* 95(3): 195–242. doi.org/10.1080/00029890.1988.11971990
[10] Mamakani, K., Ruskey, F. (2014). New roses: simple symmetric Venn diagrams with 11 and 13 curves. *Discrete Comput. Geom.* 52(1): 71–87. doi.org/10.1007/s00454-014-9605-6
[11] Ruskey, F., Weston, M. (1997, revised 2005). A survey of Venn diagrams. *Electron. J. Combin.* 4(1): (electronic). doi.org/10.37236/26
[12] Stewart, I. (1989). *Game, Set, and Math*. Oxford: Basil Blackwell.
[13] Tait, P. G. (1900). On knots I, II, III. In: *Scientific Papers*, Vol. 1. London: Cambridge University Press, pp. 273–347.
[14] Venn, J. (1880). On the diagrammatic and mechanical representation of propositions and reasonings. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 10(59): 1–18.
[15] Wagon, S., Webb, P. Venn symmetry and prime numbers: a seductive proof revisited. *Amer. Math. Monthly*. 115(7): 645–648.

Summary. We show a Brunnian link whose minimal projection is a simple symmetric Venn diagram of order 7, solving a problem that first appeared in 1998. The Venn diagram itself was discovered by Branko Grünbaum in 1992.

RICK MABRY (MR Author ID: 117420) is an emeritus professor at LSU Shreveport, where he held his only academic position for a very short quarter century.

LAURA MCCORMICK (MR Author ID: 888302) is a mathematics teacher at Discovery High School in Lake Alfred, Florida. At the time of research she was completing her bachelor's degree in Mathematics at LSU Shreveport.

Folklore, the Borromean Rings, the Icosahedron, and Three Dimensions*

DAVE AUCKLY
Kansas State University
Manhattan, KS 66506
dav@math.ksu.edu

There are many bits of mathematical folklore—results that have been known for a long time—that are not covered in a typical undergraduate curriculum. We will give a proof that the Borromean rings are linked, but our proof will not be the shortest possible. Indeed, we give a number of unnecessary, but loosely related, stories to expose more folklore. In particular, we will see that the Borromean rings are related to the icosahedron and to something called the Poincaré homology sphere. We begin with a cautionary tale.

Around 1900, Poincaré made many important contributions to topology. He was considering spaces which looked a bit like 3-space, but which were not necessarily just $\mathbb{R}^3$. The same phenomena may be seen with shapes that look like 2-space. For instance, a small neighborhood on the surface of the earth looks like $\mathbb{R}^2$, even though the surface of the Earth is not just a flat plane. The surface of the Earth is better modeled by the 2-sphere $S^2 := \{v \in \mathbb{R}^3 \mid |v|^2 = 1\}$. The same could be said about the surface of a donut: neighborhoods of it look like $\mathbb{R}^2$, but it is not $\mathbb{R}^2$. Shapes that look like $\mathbb{R}^2$ are called 2-manifolds, and shapes that look like $\mathbb{R}^3$ are called 3-manifolds.

Poincaré noticed that a compact, "does not extend to infinity" (roughly speaking), 2-manifold that has no holes in it must be the two-sphere, S^2. He conjectured that the same thing was true for 3-manifolds, where the measure of holes in a manifold M is given by something that we now call the first homology, $H_1(M)$. In 1904, he found a counter-example to this conjecture. To see this, he introduced the notion of the fundamental group of a space, $\pi_1(M)$, and constructed a counter-example called the Poincaré homology sphere that may be described using an icosahedron. He then revised his conjecture to state that the only compact 3-manifold with trivial fundamental group is S^3 [3]. More than one hundred years later, Perelman gave the first correct proof of this conjecture [2].

In fact, before the appearance of a correct proof, a number of incorrect proofs were given of the Poincaré conjecture. In many incorrect proofs, the only hypothesis that is used is that the homology is trivial. Of course, any such proof is doomed to fail, as the Poincaré homology sphere shows. John Stallings wrote a nice paper about this called *How Not to Prove the Poincaré Conjecture* [5]. Had these authors paid a bit more attention to the folklore about the difference between homology and fundamental group, there may have been fewer false proofs. The first homology $H_1(M)$ is the abelianization of the fundamental group, $\pi_1(M)$. The difference between trivial homology and trivial fundamental group, while understood, can be subtle. We will explain the fundamental group in greater detail in the first section below.

The fundamental group

Imagine you are walking a slightly skittish dog on a leash in a forest. Most of the time the dog will stay by your side. Suddenly the dog thinks it detects a squirrel and sniffs

*The online version of this paper has color graphics that may be easier to follow.

Math. Mag. **93** (2020) 330–342. doi:10.1080/0025570X.2020.1823766
MSC: 57-01

its way around a tree twice. You are stuck until you can convince the dog to unwind from the tree. There is a group here. The dog could sniff 2 times around the tree, or 13 times, or could go the other way to wrap -7 times around the tree. The amount the dog winds around the tree can be measured by an integer. The leash tangled around two trees is displayed in Figure 1.

A more precise description of this may be given using the fundamental group. Let X be a topological space, that is, a space for which the notion of continuity is defined, and let x_0 be a point in X. Elements of the fundamental group are equivalence classes of continuous paths, $\gamma : [0, 1] \to X$ with $\gamma(0) = \gamma(1) = x_0$. Two paths, γ_0 and γ_1 are equivalent if there is a deformation (*homotopy*) $H : [0, 1] \times [0, 1] \to X$ so that $H(0, s) = \gamma_0(s)$, $H(1, s) = \gamma_1(s)$, $H(t, 0) = x_0$, and $H(t, 1) = x_0$. In the dog walk example, the space X is all of the ground that is not covered by the tree. The base point x_0 is the point that you are standing on. The leash may be viewed as a continuous path γ. In the equivalence relation, one can view the parameter t in the function H as time. Thus, $H(0, s) = \gamma_0(s)$ requires that the leash is along the path γ_0 at time zero. Fixing t, the function $H(t, \cdot)$ represents the position of the leash at time t. The condition $H(t, 0) = x_0$ corresponds to the handle of the leash being with you at all times, and the condition $H(t, 1) = x_0$ corresponds to your skittish dog sitting on your foot at all times. This is captured in the definition:

$$\pi_1(X, x_0) := \{\gamma : [0, 1] \to X \mid \gamma(0) = \gamma(1) = x_0\}/\text{homotopy}\,.$$

The group operation is called *concatenation*. Let γ represent the position of the leash after one short exploration of the dog, and let δ represent the position of the leash after another short exploration. The product, $\gamma * \delta$ is what would happen if the dog first did γ, and then did δ. The formula is:

$$(\gamma * \delta)(t) = \begin{cases} \gamma(2t) & \text{if } t \in [0, 1/2], \\ \delta(2t - 1) & \text{if } t \in [1/2, 1]. \end{cases}$$

Going further on this walk, you now approach a pair of close trees, and there is a squirrel in them. In complete excitement, your dog races counterclockwise around the left tree L back to you, then clockwise around the tree on the right R^{-1} and back to you, clockwise around the tree on the left L^{-1}, and counterclockwise around the tree on the right R. Figure 1 shows the leash and how it may be pulled a bit tighter to an equivalent (homotopic) path. This path will be denoted by $LR^{-1}L^{-1}R = L * R^{-1} * L^{-1} * R$.

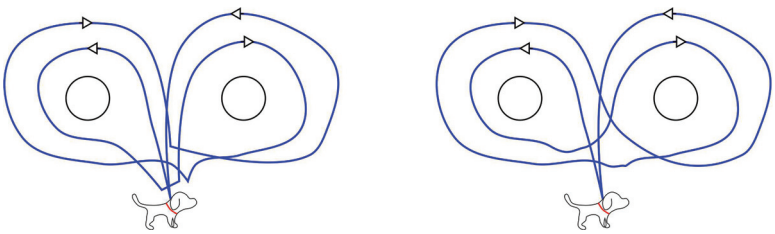


Figure 1 A tangled leash.

There is something interesting about this particular example. In total, the dog has wrapped around the left tree zero times, so the leash would pull free if the right tree was not there. In addition the dog walked around the right tree a total of zero times. Yet, as long as the squirrel continues to scare your dog and your dog does not leave your foot, the leash will be linked with the trees. This is exactly the difference between

the first homology and the fundamental group. The first homology is the abelianization of the fundamental group (in other words $XY = YX$ in homology). In homology $LR^{-1}L^{-1}R = LL^{-1}R^{-1}R = 1$. We typically write the operation in the fundamental group multiplicatively (XY) and the operation in the homology additively $(X + Y)$, to remind us that the homology is abelian.

The fundamental group of the complement of two disjoint disks in the plane (the forest floor away from the two trees) is what is known as a *free group* on two generators. The elements of this group are a trivial element 1 and all finite "words" that may be made from the letters L, R, L^{-1}, R^{-1}, with all cancellations of X next to X^{-1}. Multiplication is just the process used to create a compound word. Thus, the product of $LRLR^{-1}$ with $RRRLL$ is $LRLRRLL$.

The false proofs of the Poincaré conjecture asserted that when every loop in a compact 3-manifold wrapped zero times around any hole in the abelian sense as measured in homology, the manifold had to be trivial. This turned out to be false. Now that we have a new trick for our dog, let's find a better place to tangle our leash.

The Borromean rings

Start by making a roughly 34 inch by 55 inch rectangle. Really a 34 by $34 \cdot (1 + \sqrt{5})/2$ rectangle. The number $(1 + \sqrt{5})/2$ is the *golden ratio*. It is the length of the diagonal of a side length 1 pentagon. These numbers are close to adjacent numbers in the Fibonacci sequence (34 and 55) and this is a nice size for public display. We could prepare and make two rectangles as on the left of Figure 2 and later move them into position as on the right of Figure 2.

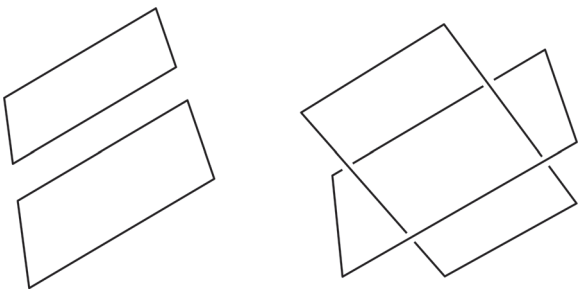


Figure 2 Two sets of unlinked rectangles.

A more precise description of the right-hand side starts with three of these rectangles. We put one with center at the origin and line of symmetry parallel to the long side (call this the rectangle axis) matching the x-axis of the $x - y$ plane. Similarly, we could put one with the axis matching the y-axis of the $y - z$ plane and one with the axis matching the z-axis of the $z - x$ plane. The result is the very specific model of the Borromean rings on the right of Figure 3. The Borromeo family used a version of this link on its coat of arms. The version from the coat of arms is the traditional image of the Borromean rings, and it is displayed in Figure 6 later.

This brings us to a question:

Question. *If we set our three rectangles out on the floor as on the left of Figure 3, would we be able to move them into the desired configuration on the right of Figure 3 without taking one apart?*

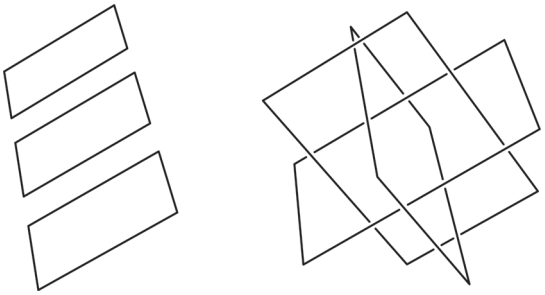


Figure 3 Assembling three rectangles.

The answer is that it is not possible to assemble three rectangles into the Borromean configuration without breaking one of the rectangles. We will ultimately give a proof of this fact using the fundamental group of the complement of the rectangles.

A *link* is a disjoint collection of topological circles in 3-space. Given a picture of a link, one may label some loops by putting arrows crossing under various strands of the link as on the left in Figure 4.

We just discovered that your dog is a super dog! He now flies up to a point above a crossing. Each arrow in Figure 4 represents the loop of leash generated when your dog starts in the air (at the base point), flies to the tail of the arrow, follows the arrow, and returns to the floating base point. The arrows labeled L and R are below the horizontal strand but above the height of the vertical strand on the left in Figure 4. Thus, the loops represented by arrows labeled L and R are equivalent because one may be deformed into the other. However, the loops labeled F and B are not equivalent, because the horizontal component of the link gets in the way of the deformation. This is shown on the right of Figure 4. The inverse of a loop, is the same loop, with reversed orientation.

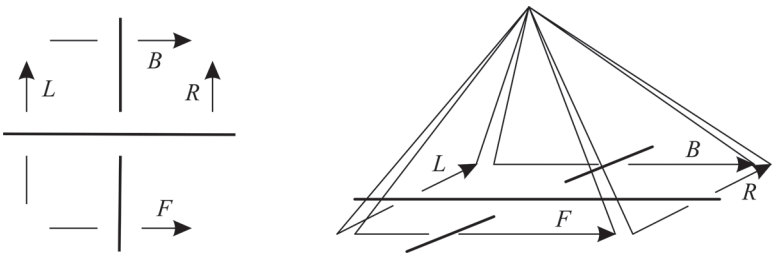


Figure 4 Loops near a crossing. On the left we have the top view. On the right is the "super leash" view.

Superdog will now show us a trick while trying to tie its leash around the strands. Superdog flies

1. down then along the arrow on the right,
2. back up, back down,
3. follows the arrow in the back in the opposite direction,
4. back up, then down,
5. follows the arrow on the left in the opposite direction,
6. back up, then down,
7. follows the front arrow in the indicated direction, and
8. finally flies back up to the base point.

The result is the product $RB^{-1}L^{-1}F$ displayed on the right in Figure 4. However, this is the trivial element of the fundamental group. To see this, note that the portions of the leash generated by the "back up, then down" portions 2, 4, and 6 of the flight may be deformed down to give the representative in Figure 5.

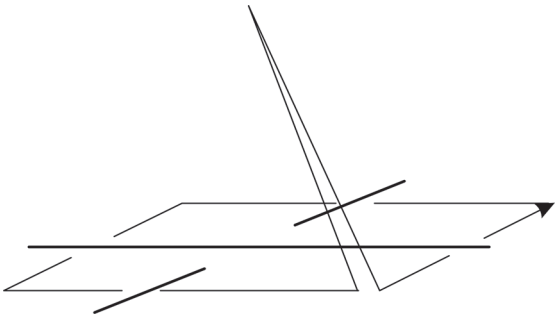


Figure 5 The "deformed" version of the link in Figure 4.

From here the square portion of the leash loop on the bottom can shrink, and the entire loop can be pulled back up to the base point. It is a fact that the fundamental group of the complement of any link may be described as the group with one generator for each strand and one relation similar to the one we just described for each crossing. This is called the *Wirtinger presentation*. See the book by Dale Rolfsen for a proof of this [**4**].

How could this possibly help us? We are trying to show that we cannot move the unlinked rectangles on the left in Figure 3 into the configuration on the right in Figure 3. Imagine that we could, and fill up the space around the rectangles with honey. When one pushes the set of rectangles on the left over to the position on the right, all of the honey will move as well. The correspondence between the starting location of a molecule of honey and the ending location of that molecule of honey gives a homeomorphism between the complement of the configuration on the left and the complement of the configuration on the right in Figure 3. Thus the fundamental groups of the complements of these two configurations would be the same. The fancy name for this idea is the "isotopy extension theorem."

The fundamental group of the complement of the trivial 3-component link on the left in Figure 3 is the group with three generators and no relations (one generator for each component, and no relations because there are no crossings):

$$\pi_1(\text{Complement of unlink}) = \langle P, Q, R \rangle .$$

The three rectangles in the configuration on the right in Figure 3 may be continuously deformed, without crossing, to the configuration in Figure 6. Thus the fundamental groups of the complements of these configurations are the same.

To compute the fundamental group of the complement of the Borromean rings we label arrows under the strands in Figure 6 with T, U, V, X, Y, Z. The strand on the upper right of this figure has an X arrow running under it. The loop represented by this arrow starts at a base point above the figure (say your nose), travels down to the tail of the arrow, follows the arrow and returns to the base point. This loop may be slid (homotoped) to other positions along this strand as long as one does not try to slide past an undercrossing. We label two loops equivalent to X near the top crossing to make it easier to see the relation arising from this crossing. Following the procedure in Figure 4 starting at the head of the T and proceeding clockwise we read the relation

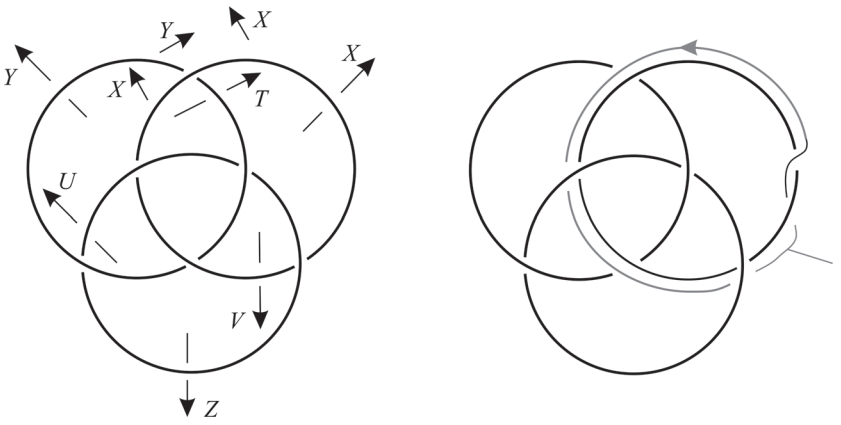


Figure 6 On the left, the Borromean rings. On the right, a "glue curve."

$T^{-1}XYX^{-1} = 1$. Thus $T = XYX^{-1}$. Similarly, the outside crossing on the left shows that $U = YZY^{-1}$ and the outside crossing on the right shows that $V = ZXZ^{-1}$. Using the three inner crossings leads to the presentation:

$$\begin{aligned}\pi_1(\text{Complement of the Borromean rings}) = \\ \langle X, Y, Z \mid YZY^{-1}XYZ^{-1}Y^{-1}ZX^{-1}Z^{-1} = 1, \\ XYX^{-1}ZXY^{-1}X^{-1}YZ^{-1}Y^{-1} = 1, \\ ZXZ^{-1}YZX^{-1}Z^{-1}XY^{-1}X^{-1} = 1\rangle\,.\end{aligned}$$

Indeed, the relation coming from the inner crossing closest to arrow Y going counterclockwise starting with the U on the top reads:

$$1 = UXU^{-1}V^{-1} = (YZY^{-1})X(YZ^{-1}Y^{-1})(ZX^{-1}Z^{-1}).$$

To prove that the two links are different, we could just show that these two groups are different. Instead we will use a slightly weird argument. We choose a weird argument because it will allow us to discuss the relationship between the Borromean rings, the Poincaré homology sphere and the icosahedron. Sometimes the long way around is more scenic.

From the Borromean rings to the Poincaré sphere

This is the most technical section of this paper. A quick summary is that if we add the relation $P = 1$, and the corresponding relations for the other components ($Q = R = 1$), to the fundamental group of the complement of the trivial 3-component link, we would get a trivial group. If the trivial 3-component link was equivalent (isotopic) to the Borromean rings, then the relation corresponding to $P = 1$ would be $X^{-1}U^{-1}Z = 1$. Adding this to the fundamental group of the complement of the Borromean rings together with the analogous relations for the other two components results in a nontrivial group, so the two links are not equivalent. Right now you should probably have no idea where the relation $X^{-1}U^{-1}Z = 1$ comes from. We will turn this around and start with the relation $X^{-1}U^{-1}Z = 1$ and see that it leads to $P = 1$. We will also construct a space with the corresponding fundamental group—the Poincaré homology sphere.

We can perform something called surgery on the trivial 3-component link and on the Borromean rings. If these two links were the same, the result of the corresponding surgeries would be the same. We begin with a warm-up. Consider the two "links" in Figure 7. Each is two points put into a space consisting of two 2-spheres.

Figure 7 Two instances of two points in a space consisting of two 2-spheres.

If there was a deformation taking the trivial 2-point link on the left to the 2-point link on the right, there would be a homeomorphism taking the complement of the two points on the left to the complement of the two points on the right. (This is the isotopy extension theorem again. Fill the complement with honey and see where the honey molecules would go after pushing the one link to the other.)

We will now glue the same thing to each complement, namely a cylinder. Notice that $S^1 \times (0, 1)$ is homeomorphic to the open unit ball with the origin deleted. Indeed, just consider the coordinate in S^1 as an angle, and the coordinate in $(0, 1)$ as a radius, and $S^1 \times (0, 1)$ will be the punctured open ball expressed in polar coordinates. Similarly $S^1 \times (-1, 0)$ is also homeomorphic to a deleted open ball. Thus, we can glue the open cylinder, $S^1 \times (-1, 1)$ to the complement of either link. The homeomorphism that we get by assuming that the links are equivalent extends and would imply that the result of gluing in the open cylinder to the left side would be homeomorphic to the result of gluing the open cylinder to the right side. However, we display the result of this gluing in Figure 8, and we can tell that the two sides are not homeomorphic because one is disconnected, and the other is connected. This process is called surgery because we are cutting something out and sewing something else back in, a bit like an organ transplant.

Figure 8 Surgery on the points shown in Figure 7.

We are now going to do the analogous thing to the Borromean rings. We cut out the rings and glue in solid tori (donuts). This is displayed in Figure 9.

To specify how an open torus is glued to the complement of the Borromean rings, we will keep track of one curve. The thin black curve on the left side of the boundary of the solid torus on the left of Figure 9 bounds a disk. However, it is not part of the open solid torus. To be clear the open solid torus is given by

$$\{(v, w) \in \mathbb{R}^2 \times \mathbb{R}^2 \,|\, |V| < 1,\ |w| = 1\}.$$

The red curve* is parallel to the thin black curve. When we glue the open solid torus to the complement of the Borromean rings we will make sure that the thin red curves in Figure 9 match. The thin red curve in the complement of the Borromean rings is called

*Since the magazine prints in black and white, this curve is rendered in gray here. The online version of this article shows the curve as red.

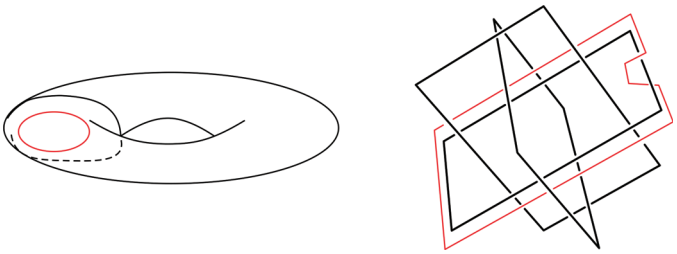


Figure 9 The Poincaré homology sphere. On the left, a donut. On the right, a rectangular glue curve.

a "surgery curve." The reason we keep track of this curve is because the corresponding loop will be trivial in the new manifold. We do the same gluing with each of the other components.

The result of attaching three open solid tori to the complement of the Borromean rings is equivalent to removing one point from the Poincaré homology sphere. We can take this as the definition of the Poincaré homology sphere.

We could fill in the missing point by gluing in a hemisphere from the 3-dimensional sphere, that is, the set of points one unit from the origin in $\mathbb{R}^4$. A schematic of this is displayed in Figure 10. Here, the left hemisphere with a small copy of the Borromean

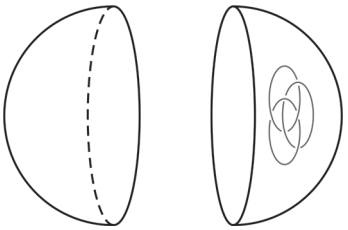


Figure 10 Two balls.

rings represents the result of removing the Borromean rings from a solid ball and gluing in donuts. The right hemisphere represents a second solid ball. Gluing these together yields the Poincaré homology sphere. We denote it by Σ. To compute the fundamental group of the Poincaré homology sphere, we just need to add relations indicating that the loop corresponding to the thin (red) circle and the analogous loops are trivial. The thin (red) loop in Figure 9 corresponds to the loop on the right in Figure 5. Reading the relation we get $X^{-1}U^{-1}Z = X^{-1}YZ^{-1}Y^{-1}Z = 1$. Adding the analogous relations for the other two components gives:

$$\begin{aligned}\pi_1(\Sigma) = \langle X, Y, Z \mid\ & YZY^{-1}XYZ^{-1}Y^{-1}ZX^{-1}Z^{-1} = 1,\\ & XYX^{-1}ZXY^{-1}X^{-1}YZ^{-1}Y^{-1} = 1,\\ & ZXZ^{-1}YZX^{-1}Z^{-1}XY^{-1}X^{-1} = 1,\\ & X^{-1}YZ^{-1}Y^{-1}Z = 1,\\ & Y^{-1}ZX^{-1}Z^{-1}X = 1,\\ & Z^{-1}XY^{-1}X^{-1}Y = 1\rangle\,.\end{aligned}$$

This is the fundamental group of the Poincaré homology sphere. We can simplify the expression for this group. The last relation in $\pi_1(\Sigma)$ implies $Z = XY^{-1}X^{-1}Y$.

Substituting this into the other relations allows us to write the group without using the generator Z. Using this, the relation $Y^{-1}ZX^{-1}Z^{-1}X = 1$ is seen to be equivalent to

$$XY^{-1}X^{-1}YX^{-1}Y^{-1}X = 1,$$

and the relation $X^{-1}YZ^{-1}Y^{-1}Z = 1$ is seen to be equivalent to

$$YX^{-1}Y^{-1}XY^{-1}X^{-1}Y = 1.$$

The other relations all follow from these two and the expression for Z. Thus,

$$\pi_1(\Sigma) = \langle X, Y \mid XY^{-1}X^{-1}YX^{-1}Y^{-1}X = 1, YX^{-1}Y^{-1}XY^{-1}X^{-1}Y = 1\rangle .$$

To write this in an interesting way, set $X = A^{-1}BA^{-2}B$ and $Y = A^{-1}B$. Notice that this implies that $A = YX^{-1}Y$ and $B = YX^{-1}Y^2$, so we can either use generators X and Y, or generators A and B, to describe the group. Using this substitution, the relation

$$YX^{-1}Y^{-1}XY^{-1}X^{-1}Y = 1$$

becomes $AB^{-2}A^2 = 1$, or just $B^2 = A^3$. Now

$$XY^{-1}X^{-1}YX^{-1}Y^{-1}X = 1$$

becomes

$$A^{-1}BA^{-1}B^{-1}A^2B^{-1}A^2B^{-1}A^{-1}B = 1.$$

Using $A^2 = B^2A^{-1}$ this becomes $(A^{-1}B)^4B^{-2}(A^{-1}B) = 1$ or $(A^{-1}B)^5 = B^2$. Thus we have

$$\pi_1(\Sigma) = \langle A, B \mid (A^{-1}B)^5 = A^3 = B^2\rangle .$$

Now imagine that the Borromean rings could be deformed ("isotoped" is the technical word) to the trivial link. Where would the surgery curves go? Notice that the surgery curve is on the boundary of a solid torus centered on one of the rings. It would still be on the boundary of such a solid torus after the deformation. The surgery curve is parallel to the link component in this solid torus. It would still be parallel after deformation. It also links the component once. After the deformation, it would still have to link once.

In fact, one way to define how many times a loop links a simple ring is to look at the class of the loop in the fundamental group of the complement. The fundamental group of the complement of a single ring with no crossings has just one generator and no relations. Call the generator F. Since the only words that may be made in this group are F^n, it is isomorphic to the integers, and we say a loop homotopic to F^n links the ring n times.

The tail of the loop represented by the surgery curve in Figure 5 might get deformed into something $CRAZY$, so the resulting loop that would have to be killed in the fundamental group of the complement of the trivial 3-component link would have the form $(CRAZY)P^{\pm 1}(CRAZY)^{-1}$, and this will be trivial exactly when $P^{\pm 1} = 1$, that is, $P = 1$.

Thus if the Borromean rings could be deformed into the trivial 3-component link, the fundamental group of the Poincaré homology sphere would have to be

$$\langle P, Q, R \mid P = Q = R = 1\rangle .$$

In other words, it would be the trivial group.

Let's see, we can abelianize the fundamental group of the Poincaré homology sphere. This means we are assuming that $XY = YX$, etc. Using this in the last three relations in $\pi_1(\Sigma)$ as expressed via X, Y, and Z generators, gives $X^{-1} = 1$, $Y^{-1} = 1$ and $Z^{-1} = 1$, and this implies that the abelianization of the fundamental group of the Poincaré homology sphere is trivial. Maybe we should assume that this means that this space is just a 3-sphere? We wouldn't be the first to make this guess. Perhaps it is possible to unlink the Borromean rings after all. Hmm, our yappy friend is reminding us of a leash that represented a trivial loop in an abelianization, but which was not trivial in the fundamental group. We need to think.

Well, the fundamental group of the Poincaré sphere does not look trivial. Is this good enough? Consider a different group presentation. The following group is trivial, but it takes work to prove it. (Try!)

$$\langle C, D \mid C^{-1}D^2C = D^3,\ D^{-1}C^2D = C^3 \rangle .$$

To see that the fundamental group of the deleted Poincaré homology sphere is not trivial, it would suffice to construct a surjective homomorphism from it to some other non-trivial group. We could do so now and end the paper. However, sometimes it is entertaining when the old man on the porch starts telling stories. This reminds us of something related to an icosahedron, so we'll tell a few more stories.

The icosahedron

Now return to the three linked rectangles from Figure 3, and add 34-inch struts connecting the corner of each rectangle to the four corners of the other rectangles that are closest to it. The result is the structure displayed on the left in Figure 11. If one removes the long edges of each original rectangle, the remaining figure is called an "icosahedron." It is the structure represented by the MAA logo, and it is displayed on the right in Figure 11.

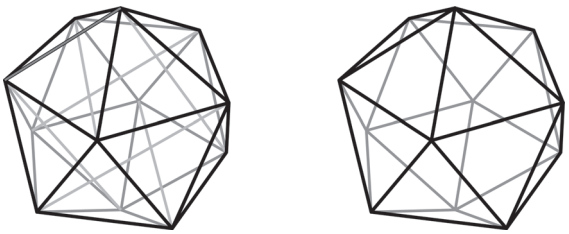


Figure 11 Connecting the corner of each rectangle to the four corners of the other rectangles closest to it leads to the structure on the left. Removing the long edges of each original rectangle leads to the familiar icosahedron, shown on the right.

The icosahedron is one of the five Platonic solids. These solids have captured the imaginations of people for a long time. One early model of the orbits of the planets was based on placing one Platonic solid inside of the next. It started by placing the octahedron in the icosahedron. Why don't we try to do the same thing in one way?

The convex hull of three congruent, mutually perpendicular line segments meeting at their centers is a regular octahedron. The axes of the three original rectangles exactly meet this condition, so we can add the edges of a regular octahedron to our figure by connecting the midpoints of the short edges of the original three rectangles. This

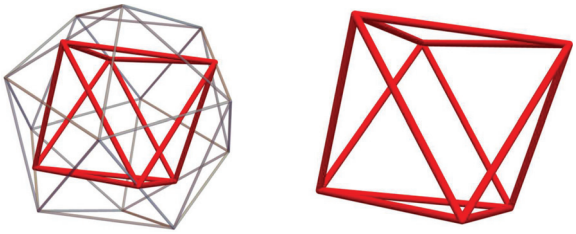


Figure 12 Constructing an octahedron within an icosahedron.

construction is displayed on the left in Figure 12, with the octahedron by itself on the right.

This is fun, so why stop? We started with three rectangles, and thus had six short edges. We put the six vertices of a red octahedron at the midpoints of these first six edges. We then added more short edges until each rectangle corner met a total of five edge ends. This gave a total of 4 corners per rectangle times 3 original rectangles times 5 (original short edge plus four new short edges) short edge ends. Thus, there are 60 short edge ends and 30 short edges. It looks like we can fit $5 = 30/6$ octahedra inside the icosahedron in this way. Let's try with

1. a red octahedron,
2. an orange octahedron,
3. a yellow octahedron,
4. a green octahedron, and
5. a blue octahedron.

It works! The result is an icosahedral compound of octahedra as displayed in Figure 13.

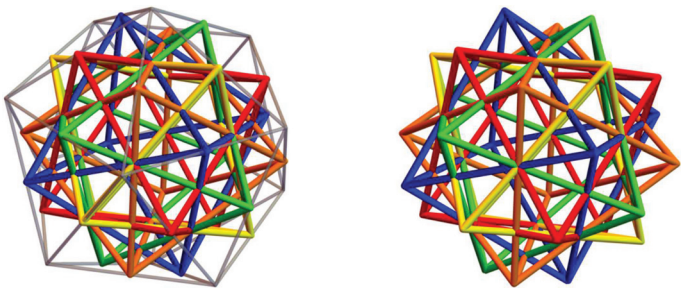


Figure 13 Compound of octahedra in and out of an icosahedron.

Consider the orientation-preserving symmetries of the icosahedron. These are all rotations, since they must fix the center of the icosahedron. There is a $1/5$th right hand rotation about the directed line connecting the bottom vertex to the top vertex. This rotation takes: 1—the red octahedron to 2—the orange octahedron, which moves to the location of 3—the yellow octahedron, which moves to 4—the green octahedron, which moves to 5—the blue octahedron, which moves to 1—the red octahedron. We will summarize this by labeling this rotation (12345).

There is also a $1/3$ rd right-hand rotation about the directed line connecting the centroid of the bottom front triangle to the centroid of the opposite triangle. This rotation takes 1—the red octahedron, to 2—the orange octahedron, to 3—the yellow octahedron, which moves to 1—the red. This should be labeled by (123).

Finally, there is a $1/2$ rotation about the line connecting the top of the red octahedron to the bottom of the red octahedron. This interchanges the 4—green, and the 5—blue octahedra, as well as interchanges the 2—orange, and the 3—yellow octahedra. We label it by $(23)(45)$.

Every rotational symmetry of the icosahedron must be one of these three types because the axis of rotation will meet the surface of the icosahedron in two opposite fixed-points. The only possibilities for one of these surface fixed-points are a vertex, the center of a triangle, or the midpoint of an edge. There are 12 vertices, so 6 opposite vertex pairs. Each vertex pair has $1/5$, $2/5$, $3/5$, and $4/5$ rotations for a total of $6 \times 4 = 24$ rotations. As we saw in the construction, there are 30 edges, so 15 pairs of opposite edges, and 15 more rotations. Each vertex meets 5 triangle corners for a total of 5×12 triangle corners and $60/3 = 20$ triangular faces. Each of the $20/2 = 10$ opposite triangle pairs contributes $1/3$ and $2/3$ rotations for a total of 20 more rotations. Combined with the trivial motion (1), we see that there are a total of $24 + 15 + 20 + 1 = 60$ orientation-preserving symmetries.

It is not difficult to see the group operations in terms of the labels. For example, the label (153) corresponds to a map taking 5 to 3 and 3 to 1, so it is clear that the inverse and the cube are the maps corresponding to

$$(153)^{-1} = (135) \quad \text{and} \quad (153)^3 = (1).$$

Similarly,

$$(135) \circ (12)(34) = (12345) \quad ((12)(34))^2 = (1) \quad \text{and} \quad (12345)^5 = (1).$$

Thus,

$$\left((153)^{-1}(12)(34)\right)^5 = (153)^3 = ((12)(34))^2 .$$

The set of all functions permuting the numbers $1, \ldots, 5$ forms a group known as the permutation group on 5 symbols. It is denoted by $\mathfrak{S}_5$ and has $5! = 120$ elements. The labeling scheme demonstrates that the orientation-preserving symmetry group of the icosahedron is isomorphic to the index-2 subgroup of $\mathfrak{S}_5$ known as the alternating group of 5 symbols, A_5.

Oh, this reminds us where we were going. We were going to explain why the fundamental group of the Poincaré homology sphere is non-trivial. We just saw that $a = (153)$ and $b = (12)(34)$ satisfy the same relations as A and B in the simplified presentation of the fundamental group of the Poincaré homology sphere. Recall these were $(A^{-1}B)^5 = A^3 = B^2$. Yet the alternating group on 5-symbols is not trivial. In fact this shows that the fundamental group of the Poincaré homology sphere surjects onto the symmetry group of the icosahedron.

The Poincaré homology sphere is a remarkably beautiful space that appears in many different contexts. Read about several descriptions of it in *Eight Faces of the Poincaré Homology Sphere* [1]. To learn more about computing fundamental groups of knot and link complements, cutting out tubes and gluing back in donuts, and many other foundations of low-dimensional topology, Rolfsen's book is a good place to start [4].

Acknowledgments The author would like to thank the referees and Bob Burckel for very helpful comments on an earlier draft. Partially supported by Simons Foundation grant 585139, and National Science Foundation grants DMS 1952755 and INCLUDES 1744474.

REFERENCES

[1] Kirby, R. C., Scharlemann M. G. (1979). Eight faces of the Poincaré homology 3-sphere. In: *Geometric Topology: Proc. Georgia Topology Conf., Athens, GA, 1977*. New York: Academic Press, pp. 113–146.

[2] Perelman, G. (2002). The entropy formula for Ricci flow and its geometric applications. Preprint.
[3] Poincare, H. (1900). Second complement a l'analysis situs. *Proc. Lond. Math. Soc.* 32(1): 277–308. doi.org/10.1112/plms/s1-32.1.277
[4] Rolfsen, D. (1990). *Knots and Links*. Mathematics Lecture Series, Vol. 7. Houston: Publish or Perish Inc.
[5] Stallings, J. (1966). How not to prove the Poincaré conjecture. In: *Topology Seminar, Wisconsin, 1965. Ann. of Math Stud.*, Vol. 60. Princeton: Princeton University Press, pp. 83–88.

Summary. There is a relationship between the Borromean rings, the icosahedron, and something called the Poincaré homolgy sphere. This relationship is explored in a wandering path that introduces fundamental ideas from topology and a geometric construction of an icosahedral compound of octahedra. This exploration results in proofs that the orientation-preserving symmetry group of an icosahedron is an alternating group of five symbols, the fact that the Borromean rings are linked, and background related to the Poincaré conjecture.

DAVE AUCKLY (MR Author ID: 305151) is amazed by how fast his children are growing up and treasures time with his family. He has been actively involved in mathematical outreach for his entire career. He served as the Associate Director of MSRI for three years, and is one of the co-founders of the Navajo Nation Math Circles Project. This paper is based on an activity used at the Navajo Nation Math Circles Summer Camp (see Figure 14).

Figure 14 The icosahedral tent at math camp (Photo by Amanda Serenevy).

The Pythagorean Theorem Via Equilateral Triangles

ANDRÉS NAVAS*
Universidad de Santiago de Chile
Alameda 3363, Santiago, Chile
andres.navas@usach.cl

Throughout history, there have been many different proofs of the Pythagorean theorem. Here we propose another one, which is nonstandard in that we use neither squares nor similarity of polygons. To proceed, consider the right triangle in Figure 1, with equilateral triangles drawn on each of its sides.

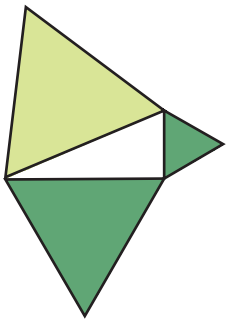


Figure 1 A right triangle with equilateral triangles drawn on its sides.

We will prove the following theorem:

Theorem 1. In Figure 1, the areas of the small equilateral triangles sum up to the area of the larger one.

We prove this by the classical method of rearrangement. Let A, B, C be the vertices of the triangle, and let a, b, c be the lengths of the corresponding opposite sides. Rotate the triangle $\triangle ABC$ counter-clockwise by 60^o at the vertex A, and clockwise by 60^o at B. Let C_1, B_1 and C_2, A_2 be the images of the vertices under these transformations, as shown in Figure 2. Notice that

$$|BA_2| = c = |AB_1|, \qquad \text{and} \qquad \angle ABA_2 = \angle BAB_1 = 60^o.$$

Hence, A_2 and B_1 coincide. Denoting this point by D, we have that the vertices A, B and D determine an equilateral triangle of side length c.

Also, notice that $\triangle BCC_2$ and $\triangle ACC_1$ are equilateral triangles of side length a and b, respectively. Moreover, triangles $\triangle BC_2D$ and $\triangle AC_1D$ are both congruent to triangle $\triangle ABC$, as shown in Figure 3.

Now, referring to the areas, we have

$$\begin{aligned} ABC_2DC_1 &= ABD + BC_2D + AC_1D \\ &= ACC_1 + BCC_2 + ABC + C_2DC_1C. \end{aligned}$$

*Funded by Anillo Project 1415 "Geometry at the Frontier."

Math. Mag. **93** (2020) 343–346. doi:10.1080/0025570X.2020.1817696
MSC: 01A99, 97G40, 97G50

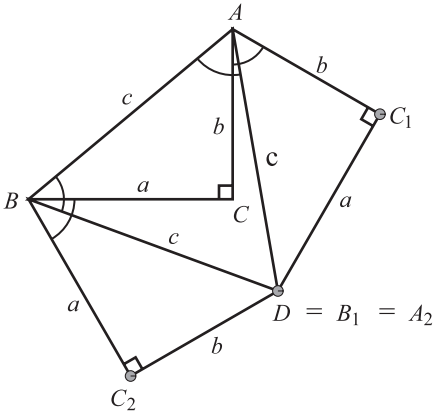


Figure 2 The result of rotating $\triangle ABC$ counter-clockwise 60° about A, and 60° clockwise about B.

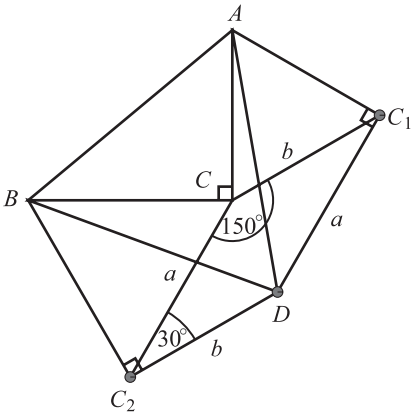
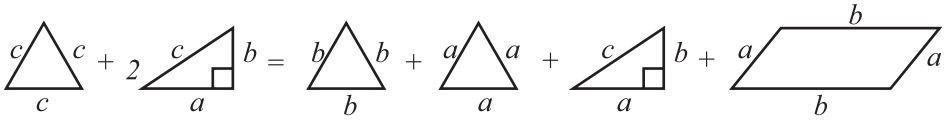
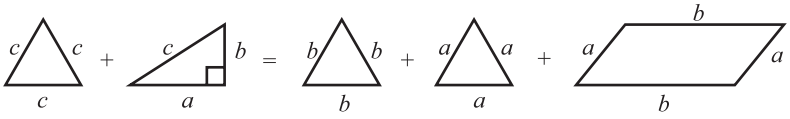


Figure 3 Our construction implies that $\triangle BCC_2$ and $\triangle ACC_1$ are both equilateral. Moreover, $\triangle BC_2D$ and $\triangle AC_1D$ are both congruent to $\triangle ABC$.

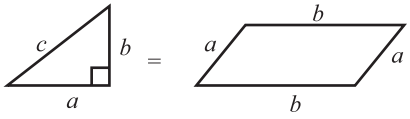
In pictures,



which yields



Hence, we are left to prove that



To do this, notice that C_2DC_1C is a parallelogram of side lengths a and b. Also, since $\angle BCA = 90^\circ$ and

$$\angle ACC_1 = \angle BCC_2 = 60^\circ,$$

we have that $\angle C_1CC_2$ must equal 150°. Moreover,

$$\angle CC_2D = \angle BC_2D - \angle BC_2C = 30^\circ,$$

and similarly $\angle DC_1C = 30^\circ$. Thus,

$$\text{area}(C_1CC_2D) = ab\sin(30^\circ) = \frac{1}{2}ab = \text{area}(ABC),$$

as desired.

In the last step above, one can certainly avoid the use of trigonometry just by looking at Figure 4. Indeed, the height of a parallelogram with angles 30° and 150° equals half the length of the side that is not being considered as the base. This is simply because a triangle of angles 30°, 60°, and 90° is half of an equilateral triangle. Hence, the length of its smallest side is half of the length of the largest one.

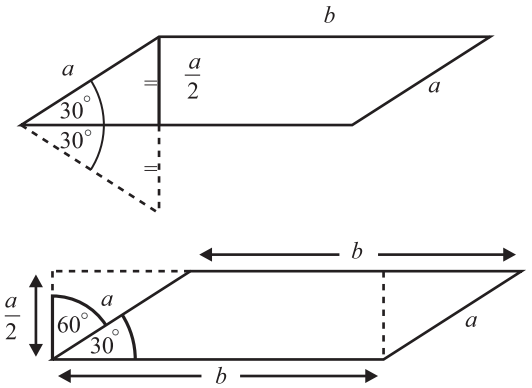


Figure 4 Another approach to finding the area of parallelogram C_1CC_2D.

It is straightforward to show that Theorem 1 implies the Pythagorean theorem.

Since the area of an equilateral triangle of side length ℓ equals $\ell^2\sqrt{3}/4$, the equality proven above may be rewritten as

$$\frac{a^2\sqrt{3}}{4} + \frac{b^2\sqrt{3}}{4} = \frac{c^2\sqrt{3}}{4}.$$

This implies the Pythagorean equality $a^2 + b^2 = c^2$ just by canceling the common factor $\sqrt{3}/4$.

Actually, if we start with the Pythagorean identity and multiply each term by $\sqrt{3}/4$, we obtain the equality between the area of the equilateral triangle built on the hypotenuse and the sum of those built on the legs. The two claims are therefore equivalent.

Finally, the Pythagorean identity $a^2 + b^2 = c^2$ also implies that the area of a regular n-gon of side length c is the sum of the areas of two regular n-gons of side lengths a and b, respectively (see Figure 5). To see this, as above, it suffices to multiply each term of this identity by an appropriate constant (namely, the area of the regular n-gon

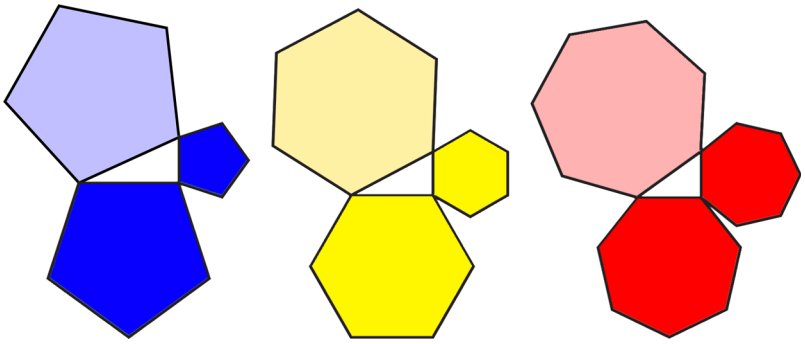


Figure 5 The area of a regular n-gon of side length c is equal to the sum of the areas of two regular n-gons of side lengths a and b, respectively.

of side length 1). Actually, this statement, for any given $n \geq 3$, also implies (and hence it is equivalent to) the Pythagorean theorem, just by reversing this argument.

Certainly, a more geometric proof for regular polygons different from equilateral triangles or squares would be desirable. Of course, in this framework, the Wallace–Bolyai–Gerwien decomposition theorem applies, hence there should be a decomposition of the two small n-gons into polygonal pieces that, after rearrangement, yield the larger one. Nevertheless, the number of required pieces seems to be quite large in general.

Acknowledgments I wish to thank Nicolé Geyssel for drawing the pictures of this note.

Summary. We prove the Pythagorean theorem by constructing equilateral triangles on each side of a right triangle. Our approach is unusual in that it uses neither squares nor similarity of polygons.

ANDRÉS NAVAS (MR Author ID: 696638, ORCID 0000-0002-0853-5136) was born in Santiago de Chile in 1976. He obtained a Ph.D. in Pure Mathematics at École Normale Supérieure de Lyon (France) in 2003, and since 2007 he is a full professor at Universidad de Santiago de Chile. His major areas of interest are dynamical systems and group theory, on which he gave an invited lecture in the last International Congress of Mathematicians (Rio de Janeiro, 2018).

What's New With Altitudes and Perpendicular Bisectors?

DIXON J. JONES
Bellingham, WA
djjones@alaska.edu

A triangle's non-parallel altitudes and perpendicular bisectors intersect in six points. In this note, we describe three apparently little-known properties of these points. The first, which recently appeared in Jones [2], can be informally summarized as follows:

Theorem 1 (Inverse similarity theorem)**.** *In a non-degenerate, oriented triangle ABC, the six intersection points of the non-parallel altitudes and perpendicular bisectors are the vertices of two congruent triangles which are inversely similar to ABC.*

Here, triangle ABC may be *positively oriented*, meaning that we measure its interior angles and label its vertices counterclockwise, or *negatively oriented*, meaning that we measure and label clockwise. Two coplanar triangles whose corresponding angles have equal magnitudes are *directly similar* if the triangles have identical orientations, and are *inversely similar* if they have opposite orientations (see Johnson [1]).

Let's begin with a positively oriented triangle ABC, whose sides a, b, and c are opposite vertices A, B, and C, respectively. Let h_a, h_b, and h_c be the altitudes perpendicular to a, b, and c, respectively, concurrent at orthocenter H. Likewise, let o_a, o_b, and o_c be the respective perpendicular bisectors of a, b, and c, concurrent at circumcenter O. Let $\ell_1 \wedge \ell_2$ denote the intersection of lines ℓ_1 and ℓ_2. Following Jones [2], we call the intersections of the altitudes and perpendicular bisectors the *minor periambic points*, labeled as follows:

$$\begin{aligned} A_p &= h_c \wedge o_b & \qquad A_q &= h_b \wedge o_c \\ B_p &= h_a \wedge o_c & \qquad B_q &= h_c \wedge o_a \\ C_p &= h_b \wedge o_a & \qquad C_q &= h_a \wedge o_b \,. \end{aligned}$$

In Jones [2], the minor periambic triangles $A_pB_pC_p$ and $A_qB_qC_q$ are shown to be inversely similar to triangle ABC (Figure 1); this more precisely expresses the inverse similarity theorem. The minor periambic triangles are congruent and radially symmetric around the midpoint of the Euler segment HO. Furthermore, they each have negative area. The trigonometric formula for triangle area is $\frac{1}{2}|x||y|\sin\theta$, where x and y are triangle sides, $|x|$ and $|y|$ are their lengths, and θ is the interior angle between them. In a positively oriented triangle, each vertex angle is positive and less than $\frac{\pi}{2}$, so $\sin\theta > 0$ and the area is positive; in a negatively oriented triangle, each vertex angle is negative and greater than $-\frac{\pi}{2}$, yielding a negative area. To emphasize that they may be positive or negative, we say that triangle areas are *signed*.

The inverse similarity theorem provides a springboard for two more intriguing results. We construct images of the minor periambic triangles under a reflection transformation: Each minor periambic point is reflected across the side of ABC containing its defining perpendicular bisector's foot (illustrated for A_p, B_p, and C_p in Figure 2). Let A'_p be the reflection of A_p across side b, let B'_p be the reflection of B_p across side c, let C'_p be the reflection of C_p across side a, and so on for A'_q from A_q and c, B'_q from B_q and a, and C'_q from C_q and b. We then have:

Math. Mag. **93** (2020) 347–351. doi:10.1080/0025570X.2020.1822703
MSC: 51M04

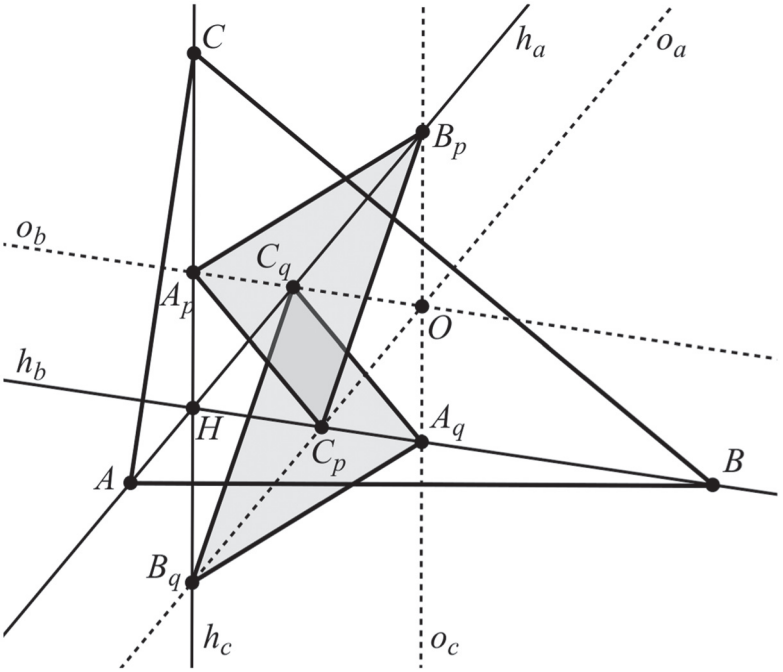


Figure 1 The minor periambic triangles $A_pB_pC_p$ and $A_qB_qC_q$.

Theorem 2 (Direct similarity theorem)**.** *Triangles $A'_pB'_pC'_p$ and $A'_qB'_qC'_q$ are congruent and directly similar to triangle ABC.*

Theorem 3 (Constant-sum area theorem)**.** *In signed areas, where* area(*ABC*) *is positive,*

$$\begin{aligned}\text{area}(A_pB_pC_p) + \text{area}(A'_pB'_pC'_p)\\ = \text{area}(A_qB_qC_q) + \text{area}(A'_qB'_qC'_q) = \text{area}(ABC).\end{aligned}$$

As with the proof of the inverse similarity theorem, these two theorems are proved by computing elementary but lengthy arithmetical expressions (admittedly, symbolic algebra software helps with this). Without loss of generality, place triangle *ABC* in the Cartesian plane with

$$A = (0, 0), \qquad B = (1, 0), \qquad \text{and} \qquad C = (c_1, c_2),$$

where $c_2 > 0$. This gives a positively oriented triangle whose altitudes and perpendicular bisectors are

$$\begin{aligned} h_a &: y = \frac{1 - c_1}{c_2}x & o_a &: y = \frac{1 - c_1}{c_2}x + \frac{c_1^2 + c_2^2 - 1}{2c_2}\\ h_b &: y = -\frac{c_1}{c_2}x + \frac{c_1}{c_2} & o_b &: y = -\frac{c_1}{c_2}x + \frac{c_1^2 + c_2^2}{2c_2}\\ h_c &: x = c_1 & o_c &: x = \frac{1}{2}\end{aligned}$$

from which it follows that

$$A_p = h_c \wedge o_b = \left(c_1, \frac{c_2^2 - c_1^2}{2c_2}\right)$$

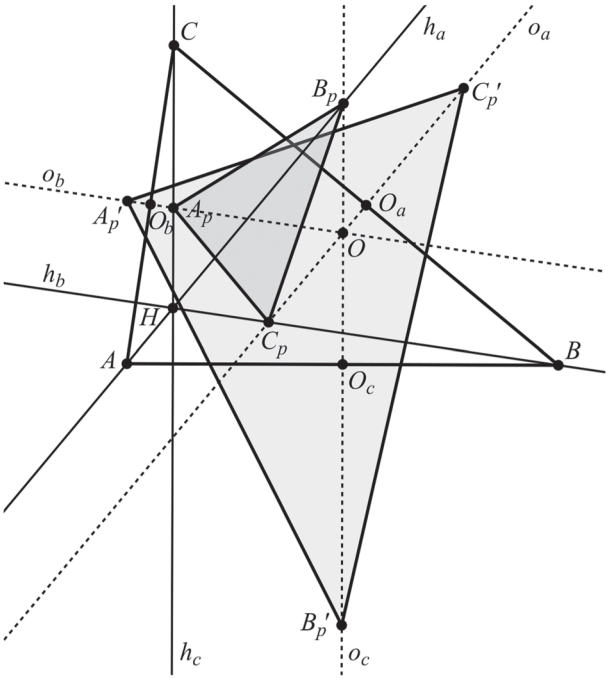


Figure 2 The minor periambic triangle $A_pB_pC_p$ and its side-reflected image $A'_pB'_pC'_p$, where $A'_pO_b = O_bA_p$, $B'_pO_c = O_cB_p$, and $C'_pO_a = O_aC_p$.

$$B_p = h_a \wedge o_c = \left(\frac{1}{2}, \frac{1-c_1}{2c_2}\right)$$
$$C_p = h_b \wedge o_a = \left(\frac{1+2c_1-c_1^2-c_2^2}{2}, \frac{c_1-2c_1^2+c_1^3+c_1c_2^2}{2c_2}\right).$$

For real numbers m and d, the reflection of a point (x_0, y_0) across the line $y = mx + d$ is

$$(x_0', y_0') = \left(\frac{(1-m^2)x_0 + 2m(y_0-d)}{1+m^2}, \frac{2mx_0 - (1-m^2)y_0 + 2d}{1+m^2}\right).$$

Applying this to A_p, B_p, and C_p, we obtain reflected points

$$A'_p = \left(0, \frac{c_1^2+c_2^2}{2c_2}\right)$$
$$B'_p = \left(\frac{1}{2}, \frac{c_1-1}{2c_2}\right)$$
$$C'_p = \left(\frac{c_1^2+c_2^2+1}{2}, -\frac{c_1^3-2c_1^2+c_2^2c_1+c_1-2c_2^2}{2c_2}\right).$$

To show that triangles $A'_pB'_pC'_p$ and ABC are directly similar, we first calculate the squares of the side lengths of the latter:

$$(AB)^2 = 1$$

$$(BC)^2 = 1 - 2c_1 + c_1^2 + c_2^2$$
$$(CA)^2 = c_1^2 + c_2^2 .$$

The squares of the side lengths of triangle $A'_pB'_pC'_p$ work out to be

$$(A'_pB'_p)^2 = \frac{1}{4c_2^2}\left[c_1^4 - 2c_1^3 + c_1^2(2c_2^2+3) - 2c_1(c_2^2+1) + c_2^4 + 3c_2^2 + 1\right]$$
$$(B'_pC'_p)^2 = \frac{1}{4c_2^2}\left[(c_1^2 - 2c_1 + c_2^2 + 1)(c_1^4 - 2c_1^3 + c_1^2(2c_2^2+3) - 2c_1(c_2^2+1) + c_2^4 + 3c_2^2 + 1)\right]$$
$$(C'_pA'_p)^2 = \frac{1}{4c_2^2}\left[c_1^6 - 2c_1^5 + 3c_1^4(c_2^2+1) - 2c_1^3(2c_2^2+1) + c_1^2(3c_2^4 + 6c_2^2 + 1) - 2c_1(c_2^4 + c_2^2) + c_2^6 + 3c_2^4 + c_2^2\right] .$$

Expanding the preceding three equalities, we find that

$$\left(\frac{A'_pB'_p}{AB}\right)^2 = \left(\frac{B'_pC'_p}{BC}\right)^2 = \left(\frac{C'_pB'_p}{CA}\right)^2 = \frac{c_1^4}{4c_2^2} - \frac{c_1^3}{2c_2^2} + \frac{3c_1^2}{4c_2^2} + \frac{c_1^2}{2} - \frac{c_1}{2c_2^2} - \frac{c_1}{2} + \frac{c_2^2}{4} + \frac{1}{4c_2^2} + \frac{3}{4} ,$$

which shows that the sides of triangles ABC and $A'_pB'_pC'_p$ are proportional.

The area of triangle $A_pB_pC_p$ is computed as

$$\begin{aligned}\text{area}(A_pB_pC_p) &= \frac{1}{2}\left|\begin{matrix}\vec{B}_p - \vec{A}_p \\ \vec{C}_p - \vec{A}_p\end{matrix}\right| \\ &= \frac{1}{2}\begin{vmatrix} \frac{1}{2} - c_1 & \frac{1 - c_1 + c_1^2 - c_2^2}{2c_2} \\ \frac{1 - c_1^2 - c_2^2}{2} & \frac{c_1 - c_1^2 + c_1^3 + c_1c_2^2 - c_2^2}{2c_2} \end{vmatrix} \\ &= \frac{1}{2}\left(-\frac{c_1^4}{4c_2} + \frac{c_1^3}{2c_2} - \frac{c_1^2c_2}{2} - \frac{3c_1^2}{4c_2} + \frac{c_1c_2}{2} + \frac{c_1}{2c_2} - \frac{c_2^3}{4} - \frac{1}{4c_2} + \frac{c_2}{4}\right),\end{aligned} \tag{1}$$

while the area of triangle $A'_pB'_pC'_p$, calculated in the same way using vectors $\vec{B}'_p - \vec{A}'_p$ and $\vec{C}'_p - \vec{A}'_p$, turns out to be

$$\text{area}(A'_pB'_pC'_p) = \frac{1}{2}\left(\frac{c_1^4}{4c_2} - \frac{c_1^3}{2c_2} + \frac{c_1^2c_2}{2} + \frac{3c_1^2}{4c_2} - \frac{c_1c_2}{2} - \frac{c_1}{2c_2} + \frac{c_2^3}{4} + \frac{1}{4c_2} + \frac{3c_2}{4}\right) . \tag{2}$$

The area formulas (1) and (2) are negatives of each other except for their last terms. We conclude that

$$\text{area}(A_pB_pC_p) + \text{area}(A'_pB'_pC'_p) = \frac{c_2}{2} = \text{area}(ABC) .$$

In our positively oriented triangle ABC, area$(ABC) > 0$, and we know from Jones [2] that area$(A_pB_pC_p) < 0$. It follows that area$(A'_pB'_pC'_p) > 0$, and consequently that triangles

$$A'_pB'_pC'_p \qquad \text{and} \qquad ABC$$

are directly similar. Repeating these calculations using A_q, B_q, and C_q yields corresponding results for triangles $A_qB_qC_q$ and $A'_qB'_qC'_q$, and shows that

$$A_pB_pC_p \qquad \text{and} \qquad A_qB_qC_q$$

are congruent.

Undoubtedly, simpler proofs of all three theorems will be found. And, frankly, one can scarcely believe that the inverse similarity theorem has not previously appeared in the multi-millennium history of plane geometry. Despite these cautions, the results presented here suggest that more geometric gems may yet be found among the venerable perpendiculars of a triangle.

REFERENCES

[1] Johnson, R. A. (2007). *Advanced Euclidean Geometry*. Mineola: Dover Publications, pp. 11–18.

[2] Jones, D. J. (2017). The periambic constellation: altitudes, perpendicular bisectors, and other radical axes in a triangle. *Forum Geom.* 17: 383–399. forumgeom.fau.edu/FG2017volume17/FG201737.pdf

Summary. The six intersection points of a triangle's non-parallel altitudes and perpendicular bisectors are the vertices of two congruent *minor periambic triangles* which are inversely similar to the original. We show that elementary transformations of the minor periambic triangles yield a second pair of congruent triangles which are directly similar to the original, and that the signed areas of a minor periambic triangle and its transformed image sum to the area of the original.

DIXON J. JONES (MR Author ID 305906) published his first mathematical work in 1976 in this MAGAZINE. He studied mathematics and design at the Illinois Institute of Technology, earning a BA in 1979, and retired in 2017 after 28 years as a graphic designer at the Elmer E. Rasmuson Library of the University of Alaska Fairbanks.

Inextricably Linked: A Trigonometric Coupling

JOHN ENGBERS
Marquette University
Milwaukee, WI 53233
john.engbers@marquette.edu

ADAM HAMMETT
Cedarville University
Cedarville, OH 45314
ahammett@cedarville.edu

IAN HOGAN
Independent Researcher
Xenia, OH 45385
hoganmathematics@gmail.com

When first learning calculus, there is often a large emphasis on understanding and using (polar) trigonometry, which is absolutely ubiquitous as a fundamental problem-solving tool throughout the entire calculus sequence. On the other hand, *hyperbolic* trigonometry, usually defined in the typical calculus textbook development via exponentials and computed in the same manner (see, for example, Edwards and Larson [**1**, pp. 365–371]), rarely gets much more than a cursory mention on the way to "more important" topics.

So it may come as a surprise to discover that these two trigonometries are, in fact, inextricably linked in a rather elementary way! Of course, the mere fact that they are related is not so surprising in light of formulas one learns in trigonometry involving complex numbers, specifically Euler's formula $e^{ix} = \cos x + i \sin x$ $(x \in \mathbb{R})$, but the connection we articulate below only requires very basic analytic geometry. Complex numbers are completely avoided.

The thrust of what we shall do is this: following a classical approach that goes back to at least an 1895 article by Mellon W. Haskell [**4**], we will *define* the polar and hyperbolic trigonometric functions via *areas* in a natural and unified way. Figures 1 and 2 tell most of the story, and provide about as close to a "proof without words" of the argument as one can hope to get. Using this geometric connection, we will establish an interdependence between the polar and hyperbolic trigonometries. In particular, if we start by assuming a subset of the basic rules about the hyperbolic sine and cosine functions, we can directly show that the standard integral formulas for the *inverse polar* trig functions hold. These formulas, in tandem with the fundamental theorem of calculus and the inverse function theorem, will deliver all the derivatives of the polar trig functions. This is a brand new approach to the result presented in Arnold Insel's classroom capsule [**5**]. Conversely, if we instead start by assuming a subset of the basic rules about the *polar* sine and cosine functions, we similarly obtain the parallel results for the *hyperbolic* trig functions and their inverses. This last part was the subject of a recent classroom capsule [**2**], the inspiration for the current note.

There is no question that our approach through geometric intuition is more difficult than the classical approach through complex numbers, but our method, in addition to avoiding complex numbers, also produces a coordinate system that involves parameterizing by *hyperbolas* rather than by circles as in the usual polar coordinate system.

Math. Mag. **93** (2020) 352–362. doi:10.1080/0025570X.2020.1826847
MSC: Primary 26A06, Secondary 97G60

These *hyperbolic coordinates* enable us to give new integral formulas (see $\mathcal{H}_3$ and $\mathcal{H}_4$) that are analogous to those for polar coordinates.

Parameterizing by area

We begin with a classic, alternative way to describe these trigonometries that will perfectly suit our needs. Start with a point (x, y) on the unit circle $x^2 + y^2 = 1$, and let $\theta/2$ be defined as the area of the circular sector subtended by the segment joining the origin to (x, y) and the positive x-axis. Then we *define* the cosine and sine functions by $x = \cos\theta$ and $y = \sin\theta$ (see Figure 1). It is clear from the formula for the area of a circular sector that this definition coincides with the usual one for the polar trigonometric functions. In other words, for the usual polar parameterization of the unit circle, the argument θ of the polar trig functions is equal to twice the area of the subtended circular sector. Of course, θ *also* happens to be the central angle (measured in radians) of this sector!

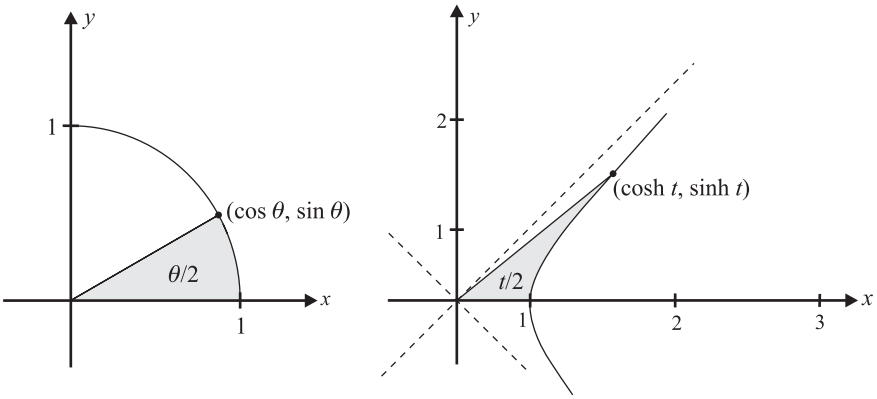


Figure 1 The unit circle and unit hyperbola parameterized by area.

Turning to the hyperbolic trigonometry, we define the "unit hyperbola" to be the right branch of the hyperbola $x^2 - y^2 = 1$, so that $x \geq 1$. Given (x, y) on the unit hyperbola, let's consider the segment joining the origin to this point. This segment, together with the positive x-axis and the unit hyperbola encloses a region of a particular area, which we denote by $t/2$ (see Figure 1). For simplicity, we initially only consider points (x, y) with $x > y \geq 0$, so that $t/2 \geq 0$ (to agree with integration). Points elsewhere in the plane will be considered later. Proceeding as we did for the polar trig functions, we *define* the hyperbolic cosine and sine functions by $x = \cosh t$ and $y = \sinh t$, where (to repeat) $t/2$ is the area of the "hyperbolic sector" in Figure 1. We again remark that this parameterization by area is well-known and goes back to at least an 1895 article by Haskell [**4**]. Note that in contrast with the area parameter θ in the case of the polar trigonometry, t does *not* represent the radian angle that the segment joining the origin to $(\cosh t, \sinh t)$ makes with the positive x-axis.

It is natural to wonder whether the parameterization by area for the hyperbolic trigonometry coincides with the usual definition. Indeed it does! Consider the hyperbolic sector in Figure 1: by our definition of t and integrating with respect to the vertical direction, we have

$$t = 2\int_0^{\sinh t} \sqrt{1+v^2}\,dv - \cosh t \sinh t \qquad (v = \tan u)$$

$$= 2\int_0^{\arctan(\sinh t)} \sec^3 u\, du - \cosh t \sinh t$$

$$= \Big(\ln|\sec u + \tan u| + \sec u \tan u\Big)\Big|_0^{\arctan(\sinh t)} - \cosh t \sinh t$$

$$= \ln(\cosh t + \sinh t)\,; \tag{1}$$

here, we have used the fact that $\cosh^2 t - \sinh^2 t = 1$ (since we are on the unit hyperbola), as well as $\sec(\arctan\alpha) = \sqrt{1+\alpha^2}$. From (1) and

$$1 = \cosh^2 t - \sinh^2 t = (\cosh t - \sinh t)(\cosh t + \sinh t)$$

the familiar identities $\cosh t = (e^t + e^{-t})/2$ and $\sinh t = (e^t - e^{-t})/2$ now follow; see also Engbers and Hammett [2] or Greenspan [3].

By contrast, our goal below is to *circumvent entirely* these exponential identities for the hyperbolic trig functions, and to instead use only basic polar trig derivatives, plus the hyperbolic area parameter t and its relationship to the polar area parameter θ, to obtain the hyperbolic trig calculus properties. In the reverse direction, our aim will be the same: to use only basic hyperbolic trig derivatives, plus the polar area parameter θ and its connection with the hyperbolic area parameter t, to obtain the polar trig calculus properties.

A geometric coupling

Now that we have laid the foundation for how we will view the polar and hyperbolic trigonometries, namely through their respective area parameters θ and t, we endeavor to link θ and t together. This will involve developing a new coordinate system utilizing hyperbolas rather than circles, as we do in polar coordinates. We remind the reader that, in this initial discussion, we only consider points (x, y) in the plane with $x > y \geq 0$. These are points in the first quadrant lying strictly below the line $y = x$, or on the positive x-axis, and we denote this infinite sector of points by $\mathcal{S}$.

The idea here is quite simple, and Figure 2 tells most of the story. Consider an arbitrary point $(x, y) \in \mathcal{S}$. Just as (x, y) sits on a unique circle $x^2 + y^2 = r^2$ and thus has a polar coordinate representation $x = r\cos\theta$, $y = r\sin\theta$, so does this point sit on a unique *hyperbola* $x^2 - y^2 = \rho^2$ and thus has a *hyperbolic coordinate* representation $x = \rho\cosh t$, $y = \rho\sinh t$. In addition, we distinguish two points on the ray emanating from $(0, 0)$ through (x, y): one on the unit circle and the other on the unit hyperbola, with the property that the point on the unit circle is obtained from (x, y) via scaling by the factor $1/r$. And the point on the unit circle can be scaled by r/ρ to obtain the point on the unit hyperbola, which can in turn be scaled by ρ to get back to (x, y). Once again, Figure 2 is a geometric rendering of all these facts. Note well that in this figure, because of the scaling factors that are being applied to move among regions, it is possible to write down the areas of these regions in terms of the area parameters θ and t. As such, in Figure 2 we have listed various equations concerning the areas of these regions implicit in these scaling relationships. These will benefit us in what follows.

By our definition, it is clear that we have the following relations that allow us to move among rectangular, polar and hyperbolic coordinates:

$$\begin{aligned} &x = r\cos\theta = \rho\cosh t, \quad y = r\sin\theta = \rho\sinh t, \quad y/x = \tan\theta = \tanh t, \\ &x^2 - y^2 = \rho^2 = r^2(\cos^2\theta - \sin^2\theta), \quad x^2 + y^2 = r^2 = \rho^2(\cosh^2 t + \sinh^2 t). \end{aligned} \tag{2}$$

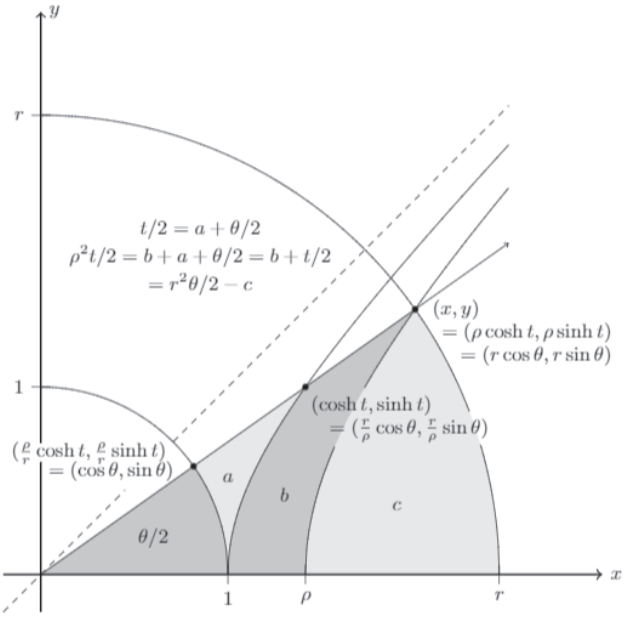


Figure 2 Converting between rectangular, polar, and hyperbolic coordinates.

As a first application of the coordinate relations in (2), consider the special case of $(x, y) = (\cosh t, \sinh t) \in \mathcal{S}$ a point on the unit hyperbola (i.e., $\rho = 1$). Then the fourth equation in (2) delivers

$$r^2 \cos^2\theta - r^2 \sin^2\theta = 1 \implies r^2 = \frac{1}{\cos^2\theta - \sin^2\theta} = \frac{\sec^2\theta}{1 - \tan^2\theta}. \tag{3}$$

On the other hand, in the special case $(x, y) = (\cos\theta, \sin\theta) \in \mathcal{S}$ a point on the unit circle (i.e., $r = 1$), the fifth equation in (2) gives

$$\rho^2 \cosh^2 t + \rho^2 \sinh^2 t = 1 \implies \rho^2 = \frac{1}{\cosh^2 t + \sinh^2 t} = \frac{\operatorname{sech}^2 t}{1 + \tanh^2 t}. \tag{4}$$

Equations (3) and (4) will be very important below.

A logical coupling

In this section, we distinguish a small number of properties for the polar and hyperbolic systems of trigonometry that we will use to show their calculus equivalence. Let's start with what we will assume about the polar trigonometry. Recall that the formula for the area A of a polar sector enclosed by the rays $\theta = \alpha$ and $\theta = \beta$ emanating from the origin, and the curve $r = r(\theta)$ for $\alpha \le \theta \le \beta$ is

$$A = \frac{1}{2} \int_\alpha^\beta r^2 \, d\theta.$$

This formula, along with two other fundamental polar trig calculus facts, will be our starting point:

$\mathcal{P}_1$: $\frac{d}{d\theta}(\sin\theta) = \cos\theta$,
$\mathcal{P}_2$: $\frac{d}{d\theta}(\cos\theta) = -\sin\theta$, and

$\mathcal{P}_3$: the area of the region bounded by $r = r(\theta)$, $\alpha \le \theta \le \beta$, is given by $A = \frac{1}{2}\int_\alpha^\beta r^2\, d\theta$.

The hyperbolic analogues of $\mathcal{P}_1$ and $\mathcal{P}_2$ are the derivatives of $\sinh t$ and $\cosh t$, respectively. However, the existence of a hyperbolic analogue of $\mathcal{P}_3$ is not obvious. Define a *hyperbolic sector* to be the region enclosed by the ray emanating from $(0, 0)$ through $(\cosh a, \sinh a)$, the ray from $(0, 0)$ through $(\cosh b, \sinh b)$, and the curve $(\rho(t)\cosh t, \rho(t)\sinh t)$ for $a \le t \le b$. This last curve we can identify by the scaling factor $\rho = \rho(t)$, $a \le t \le b$, with the understanding that the locus of points defining this curve comes from applying the scaling factor $\rho(t)$ to the corresponding point on the unit hyperbola:

$$\rho(t)(\cosh t, \sinh t) = (\rho(t)\cosh t, \rho(t)\sinh t).$$

This notion is consistent with polar curves of the form $r = r(\theta)$, $\alpha \le \theta \le \beta$. Indeed, the locus of points defining such curves is obtained by scaling points on the unit circle:

$$r(\theta)(\cos\theta, \sin\theta) = (r(\theta)\cos\theta, r(\theta)\sin\theta).$$

And since $\tan\theta = \tanh t$ (relations (2)), just as we can identify the angle θ with the ray from $(0, 0)$ through the point $(\cos\theta, \sin\theta)$, we can identify the area parameter t with the ray from the origin through $(\cosh t, \sinh t)$ (Figure 2 again).

To help the reader intuit these matters, consider the example in Figure 3. For the portion of the hyperbola $(\rho_0\cosh t, \rho_0\sinh t)$, $a \le t \le b$, (i.e., $\rho(t) = \rho_0$) the area of the corresponding hyperbolic sector subtended by the segments joining the origin to $(\rho_0\cosh a, \rho_0\sinh a)$ and to $(\rho_0\cosh b, \rho_0\sinh b)$ is precisely $\rho_0^2(b-a)/2$.

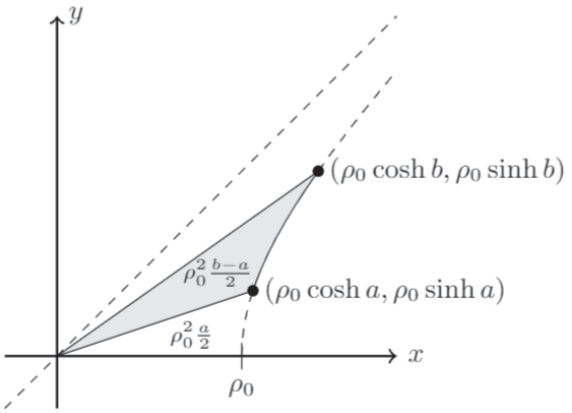


Figure 3 A hyperbolic sector determined by the hyperbola $x^2 - y^2 = \rho_0^2$.

To integrate in such coordinates, we build differentials as follows. Letting $a = t$ and $b = t + \Delta t$ in the setup just considered in Figure 3, we determine the area of the hyperbolic trapezoid $ABCD$ in Figure 4. It is easy to see that there are four distinct hyperbolic sectors in Figure 4, each with a distal coordinate of A, B, C, and D respectively. If we call the areas of these respective sectors $\mathcal{H}(A)$, $\mathcal{H}(B)$, $\mathcal{H}(C)$, and $\mathcal{H}(D)$ then the area of $ABCD$ is $\mathcal{H}(C) - \mathcal{H}(B) - \mathcal{H}(D) + \mathcal{H}(A)$.

That is, analogous to the area computation just done in Figure 3, we obtain

$$\begin{aligned}\text{Area}(ABCD) &= (\rho + \Delta\rho)^2\frac{(t+\Delta t) - t}{2} - \rho^2\frac{(t+\Delta t) - t}{2}\\ &= (2\rho + \Delta\rho)\Delta\rho\frac{\Delta t}{2} \to \rho\, d\rho\, dt, \text{ as } \Delta\rho, \Delta t \to 0.\end{aligned}$$

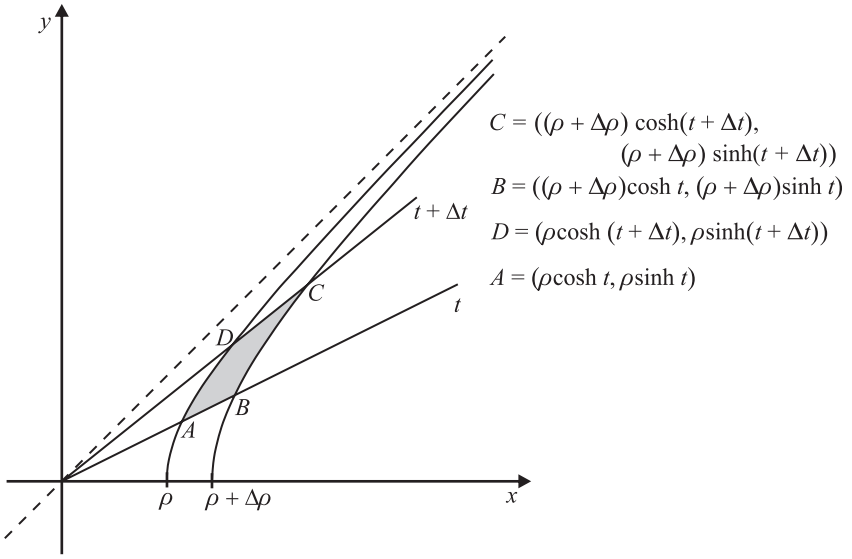


Figure 4 Hyperbolic differentials.

Integrating over these infinitesimal regions gives

$$\int_a^b \int_0^{\rho(t)} \rho \, d\rho \, dt = \frac{1}{2} \int_a^b \rho(t)^2 dt$$

as the area of the region bounded by $\rho(t)$, $a \le t \le b$, in hyperbolic coordinates. Thus, in analogy to $\mathcal{P}_1$–$\mathcal{P}_3$, we have our hyperbolic trigonometric starting point:

$\mathcal{H}_1$: $\frac{d}{dt}(\sinh t) = \cosh t$,

$\mathcal{H}_2$: $\frac{d}{dt}(\cosh t) = \sinh t$, and

$\mathcal{H}_3$: the area of the region bounded by $\rho = \rho(t)$, $a \le t \le b$, is given by $A = \frac{1}{2}\int_a^b \rho^2 \, dt$.

The similarities between $\mathcal{P}_i$ and $\mathcal{H}_i$ beckon us to derive each set from the other, and so we will do this shortly. Specifically, we will use $\mathcal{P}_1$–$\mathcal{P}_3$ to derive $\mathcal{H}_1$ and $\mathcal{H}_2$ and we will use $\mathcal{H}_1$–$\mathcal{H}_3$ to derive $\mathcal{P}_1$ and $\mathcal{P}_2$. And as we have seen, the remaining properties $\mathcal{P}_3$ and $\mathcal{H}_3$ are derived from the area interpretation and a limit process in each system.

To help illustrate area formula $\mathcal{H}_3$, we give two quick examples. Fix $\theta_0 \in [0, \pi/4)$, and consider the (unit) polar sector and triangle, each with vertex angle θ_0, as depicted in Figure 5. As we have seen in the relations (2), there is a unique $t_0 \ge 0$ such that $\tanh t_0 = \tan \theta_0$. This means that the polar integral limits $0 \le \theta \le \theta_0$ correspond to the hyperbolic integral limits $0 \le t \le t_0 = \operatorname{arctanh}(\tan \theta_0)$. Denoting the area of this sector by $S(\theta_0)$ and that of the triangle by $T(\theta_0)$, in the first case identity (4) together with $\mathcal{H}_1$–$\mathcal{H}_3$ deliver

$$\begin{aligned} S(\theta_0) &= \frac{1}{2}\int_0^{\operatorname{arctanh}(\tan\theta_0)} \frac{\operatorname{sech}^2 t}{1 + \tanh^2 t}\, dt \qquad \left(u = \tanh t;\ du = \operatorname{sech}^2 t \, dt\right) \\ &= \frac{1}{2}\int_0^{\tan\theta_0} \frac{du}{1+u^2} = \frac{1}{2}\Big(\arctan u\Big)\Big|_0^{\tan\theta_0} = \theta_0/2, \end{aligned}$$

as they should! The triangle has bounding curve $x = 1$, $0 \le y \le \tan \theta_0$, so since relations (2) imply $1 = x = \rho \cosh t$, this curve is parameterized in hyperbolic coordinates

by $\rho = 1/\cosh t = \operatorname{sech} t$, $0 \le t \le \operatorname{arctanh}(\tan\theta_0)$. Thus $\mathcal{H}_1$–$\mathcal{H}_3$ yield

$$T(\theta_0) = \frac{1}{2}\int_0^{\operatorname{arctanh}(\tan\theta_0)} \operatorname{sech}^2 t\, dt = \frac{1}{2}\Big(\tanh t\Big)\Big|_0^{\operatorname{arctanh}(\tan\theta_0)} = \frac{1}{2}\tan\theta_0,$$

which is again what we would expect.

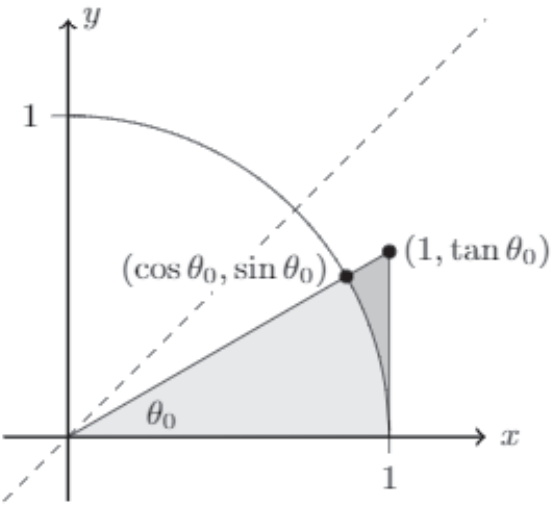


Figure 5 The polar sector and triangle with vertex angle $\theta_0 \in [0, \pi/4)$.

Before departing from this section, we remark that the manner in which we proved area formula $\mathcal{H}_3$ signals that a more general double integral formula is true. Indeed, by $\mathcal{H}_1$ and $\mathcal{H}_2$ the Jacobian of the transformation $x = \rho\cosh t$, $y = \rho\sinh t$ is given by

$$J(\rho, t) = \left|\begin{vmatrix} x_\rho & x_t \\ y_\rho & y_t \end{vmatrix}\right| = \left|\begin{vmatrix} \cosh t & \rho\sinh t \\ \sinh t & \rho\cosh t \end{vmatrix}\right| = \rho|\cosh^2 t - \sinh^2 t| = \rho.$$

Hence, we obtain the double integral formula

$$\mathcal{H}_4 : \iint_R f(x, y)\, dA = \iint_R f(\rho\cosh t, \rho\sinh t)\, \rho\, d\rho\, dt.$$

With all of these tools at our disposal, we are now ready to prove direct integral formulas for the inverse trig functions.

Direct proofs of the inverse trig integral formulas

Let $(x, y) = (\cos\theta, \sin\theta) \in \mathcal{S}$ be a point on the unit circle. Then applying the area interpretation of $\theta/2$, hyperbolic hypotheses $\mathcal{H}_1$–$\mathcal{H}_3$ together with relations (2) and identity (4) give

$$\begin{aligned} \theta &= \int_0^{\operatorname{arctanh}(y/x)} \frac{\operatorname{sech}^2 t}{1 + \tanh^2 t}\, dt \qquad \big(u = \tanh t;\ du = \operatorname{sech}^2 t\, dt\big) \\ &= \int_0^{y/x} \frac{du}{1 + u^2} \qquad \left(1 + u^2 = \frac{1}{1 - v^2};\ \frac{du}{1 + u^2} = \frac{(v/u)\, dv}{1 - v^2}\right) \end{aligned} \tag{5}$$

$$= \int_0^y \frac{dv}{\sqrt{1-v^2}} \qquad \left(1 - v^2 = w^2;\ \frac{dv}{w} = -\frac{dw}{v}\right) \tag{6}$$

$$= \int_1^x \frac{-dw}{\sqrt{1-w^2}}. \tag{7}$$

Since $z := \tan\theta = y/x$ (relations (2) again), from (5)–(7) we obtain

$$\arccos x = \int_1^x \frac{-ds}{\sqrt{1-s^2}}, \quad \arcsin y = \int_0^y \frac{ds}{\sqrt{1-s^2}},$$

$$\arctan z = \int_0^z \frac{ds}{1+s^2}. \tag{8}$$

The integral expressions for $\operatorname{arccosh} x$, $\operatorname{arcsinh} y$ and $\operatorname{arctanh} z$ are obtained in a similar manner in Engbers and Hammett [**2**], by using the area interpretation of $t/2$ together with relations (2), polar hypotheses $\mathcal{P}_1$–$\mathcal{P}_3$ and identity (3). These formulas are given by

$$\operatorname{arccosh} x = \int_1^x \frac{ds}{\sqrt{s^2-1}}, \quad \operatorname{arcsinh} y = \int_0^y \frac{ds}{\sqrt{1+s^2}},$$

$$\operatorname{arctanh} z = \int_0^z \frac{ds}{1-s^2}. \tag{9}$$

Extending the formula definitions

Thus far we have assumed we are within the infinite sector $\mathcal{S}$, so we now extend the formulas (8) and (9) to their usual domains. The case of (9) was handled in Engbers and Hammett [**2**], so we need only treat the formulas in (8). For $(x, y) = (\cos\theta, \sin\theta) \in \mathcal{S}$, we see that (8) is valid by our argument for

$$x \in \left(\tfrac{1}{\sqrt{2}}, 1\right], \quad y \in \left[0, \tfrac{1}{\sqrt{2}}\right), \quad \text{and} \quad z = y/x \in [0, 1).$$

Note that the functions $\arccos x$ and $\arcsin y$ given by the expressions in (8) are continuous for $x, y \in (-1, 1)$, as is $\arctan z$ for $z \in \mathbb{R}$. So we can easily extend them by continuity at $x = y = \frac{1}{\sqrt{2}}$ and $z = 1$.

In particular, at angle $\theta = \pi/4$ (i.e., $x = y = \frac{1}{\sqrt{2}}$ and $z = y/x = 1$), (8) delivers the definite integral values

$$\frac{\pi}{4} = \int_1^{\frac{1}{\sqrt{2}}} \frac{-ds}{\sqrt{1-s^2}} = \int_0^{\frac{1}{\sqrt{2}}} \frac{ds}{\sqrt{1-s^2}} = \int_0^1 \frac{ds}{1+s^2}.$$

By adding the first two of these integrals together, we obtain

$$\frac{\pi}{2} = \int_1^{\frac{1}{\sqrt{2}}} \frac{-ds}{\sqrt{1-s^2}} + \int_0^{\frac{1}{\sqrt{2}}} \frac{ds}{\sqrt{1-s^2}} = \int_0^1 \frac{ds}{\sqrt{1-s^2}}. \tag{10}$$

And so by symmetry we also get

$$\pi = 2 \cdot \frac{\pi}{2} = 2\int_0^1 \frac{ds}{\sqrt{1-s^2}} = \int_{-1}^1 \frac{ds}{\sqrt{1-s^2}}. \tag{11}$$

The definite integral values in (10) and (11) will be critical elements in continuing to extend the formulas in (8).

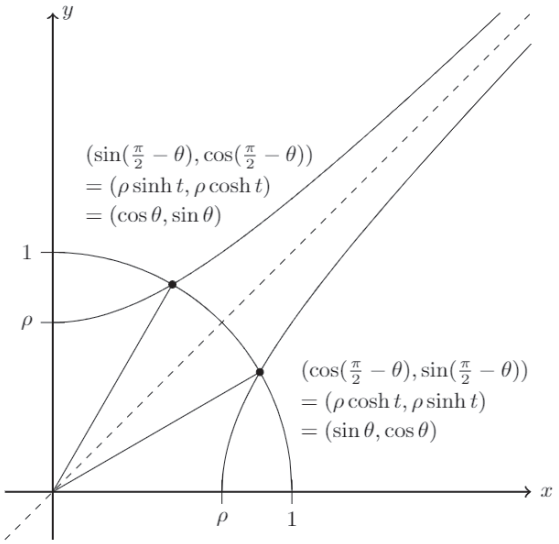


Figure 6 Extending via symmetry.

Next, take a point on the unit circle $(x, y) = (\cos\theta, \sin\theta)$ with $\theta \in (\pi/4, \pi/2]$. Then we have

$$x \in \left[0, \tfrac{1}{\sqrt{2}}\right), \quad y \in \left(\tfrac{1}{\sqrt{2}}, 1\right], \quad \text{and} \qquad z = \tan\theta = y/x \in (1, \infty).$$

Now, Figure 6 makes clear that $x = \cos\theta = \sin(\pi/2 - \theta)$ and consequently

$$\arcsin x = \frac{\pi}{2} - \theta = \frac{\pi}{2} - \arccos x.$$

Working from this identity, and invoking (8) and (10) we get

$$\begin{aligned} \arccos x = \frac{\pi}{2} - \arcsin x &= \int_0^1 \frac{ds}{\sqrt{1-s^2}} - \int_0^x \frac{ds}{\sqrt{1-s^2}} \\ &= \int_x^1 \frac{ds}{\sqrt{1-s^2}} = \int_1^x \frac{-ds}{\sqrt{1-s^2}}, \end{aligned}$$

thus obtaining the extension of (8) to $x \in [0, \frac{1}{\sqrt{2}})$. A very similar computation verifies the extension of (8) to $y \in (\frac{1}{\sqrt{2}}, 1]$, and because $z = \tan\theta = y/x \in (1, \infty)$, (8) is now established for $x, y \in [0, 1]$ and $z \geq 0$.

And what about $x, y \in [-1, 0)$? Here, we observe that for $x \in [-1, 0)$ we have $\arccos x = \pi - \arccos(-x)$ and then apply (8) (since $-x \in (0, 1]$, and we have shown the formula holds for such values) and (11) to get

$$\begin{aligned} \arccos x = \pi - \arccos(-x) &= \int_{-1}^1 \frac{ds}{\sqrt{1-s^2}} - \int_1^{-x} \frac{-ds}{\sqrt{1-s^2}} \\ &= \int_{-1}^{-x} \frac{ds}{\sqrt{1-s^2}} = \int_1^x \frac{-du}{\sqrt{1-u^2}}. \qquad (s = -u;\ ds = -du) \end{aligned}$$

Similar reasoning shows that (8) holds for $y \in [-1, 0)$. Finally, since we have shown the validity of (8) for $x, y \in [-1, 1]$, it is also true for $z = \tan\theta = y/x < 0$ (by taking $y \in [-1, 0)$ and $x \in (0, 1]$, for example).

Trigonometric derivatives

We conclude by finally showing that we can obtain the derivatives of the polar and hyperbolic trig functions from (8) and (9), respectively. First, we need to extend the definitions of these trig functions to their usual domains, as our parameters $t \geq 0$ and $\theta \in [0, 2\pi)$ have thus far been thought of in terms of area.

In the case of t, by reflecting the first quadrant point $(\cosh t, \sinh t)$ in the x-axis, we get the point $(\cosh t, -\sinh t)$, which determines a symmetrical hyperbolic sector of area $-t/2 < 0$ (to agree with integration). Denoting this point by $(\cosh(-t), \sinh(-t))$ in order to agree with our definition in the first quadrant, we see that $\cosh t = \cosh(-t)$ and $-\sinh t = \sinh(-t)$. So the core hyperbolic trig functions are now defined for all $t \in \mathbb{R}$.

And for θ, we note that $\cos\theta$ and $\sin\theta$ may be defined for $\theta \notin [0, 2\pi)$ by $\cos\theta = \cos(\theta + 2\pi a)$ and $\sin\theta = \sin(\theta + 2\pi a)$ for any $\theta \in \mathbb{R}$, where $a \in \mathbb{Z}$ is (uniquely) chosen so that $\theta + 2\pi a \in [0, 2\pi)$. This indeed fits with our definition of these functions in terms of the area parameter θ, and with these extended definitions the functions $\cos\theta$ and $\sin\theta$ have period 2π, as we would expect.

Now on to the derivatives. Note that the fundamental theorem of calculus (FTC) applied to the integral equations in (8) and (9) (differentiating with respect to the variable in the limit of integration) imply existence of the various inverse trig function derivatives on their domains. And thus by the inverse function theorem (IFT), this in turn implies existence of the various trig function derivatives on the ranges of their corresponding inverse function.

Specifically, in the case of the hyperbolic functions, applying the FTC and IFT to the equations in (9) gives

$$\begin{aligned}
\frac{d}{dt}(\sinh t) &= \sqrt{1 + \sinh^2 t} = \cosh t, \\
\frac{d}{dt}(\cosh t) &= \sqrt{\cosh^2 t - 1} = \sinh t, \\
\frac{d}{dt}(\tanh t) &= 1 - \tanh^2 t = \operatorname{sech}^2 t.
\end{aligned} \tag{12}$$

Here, having applied the IFT, it follows that the first and third identities hold for all $t \in \mathbb{R}$, but the second requires a bit more care. Indeed, for this we are applying the IFT with $f^{-1}(t) = \cosh t$ and $f(t) = \operatorname{arccosh} t$, and the former function only permits an inverse for $t \geq 0$ (this is why we know that $|\sinh t| = \sinh t$ when taking the square root), and so under the IFT this second formula is only guaranteed for $t > 0$. So how do we show that $\frac{d}{dt}(\cosh t) = \sinh t$ for $t \leq 0$ as well?

For this, we note that for $t \neq 0$ the first and third equations in (12) imply

$$\frac{d}{dt}(\cosh t) = \frac{d}{dt}\left(\frac{\sinh t}{\tanh t}\right) = \frac{\cosh^2 t - 1}{\sinh t} = \sinh t,$$

and so the second formula in (12) holds for $t \neq 0$. But then $\frac{d}{dt}(\cosh t) = \sinh t$ for $t \neq 0$ implies that in fact this formula holds for all $t \in \mathbb{R}$ by the mean value theorem: the difference quotient $(\cosh(h) - \cosh(0))/h$ is equal to $\sinh(x_h)$ for some x_h between 0 and h, so as $h \to 0$ the difference quotient approaches $\sinh(0)$ by continuity of $\sinh t$.

Thus $\mathcal{H}_1$ and $\mathcal{H}_2$ are established for all $t \in \mathbb{R}$, and any other hyperbolic trig derivative is now easily obtained from these! Similarly, from (8) we obtain the derivatives of

the polar trig functions:

$$\frac{d}{d\theta}(\sin\theta) = \sqrt{1-\sin^2\theta} = \cos\theta, \quad \frac{d}{d\theta}(\cos\theta) = -\sqrt{1-\cos^2\theta} = -\sin\theta,$$

$$\frac{d}{d\theta}(\tan\theta) = 1+\tan^2\theta = \sec^2\theta. \tag{13}$$

Yet again, having applied the IFT it follows that these identities are only guaranteed for $\theta \in (-\pi/2, \pi/2)$ in the first and third formulas, and $\theta \in (0,\pi)$ in the second. To extend these derivatives, note that for $\theta \in (\pi, 2\pi)$ we have $\theta - \pi \in (0,\pi)$ and by symmetry $\cos\theta = -\cos(\theta-\pi)$. So the derivative of (for example) $\cos\theta$ on the interval $(\pi, 2\pi)$ follows from (13) because

$$\frac{d}{d\theta}(\cos\theta) = -\frac{d}{d\theta}(\cos(\theta-\pi)) = \sin(\theta-\pi) = -\sin\theta.$$

But then it follows that $\frac{d}{d\theta}(\cos\theta) = -\sin\theta$ holds for $\theta \neq n\pi$ $(n \in \mathbb{Z})$ since cosine has period 2π. And finally, this derivative formula also holds for these isolated points $\theta = n\pi$ by the mean value theorem. A very similar argument can be used to extend the formula $\frac{d}{d\theta}(\sin\theta) = \cos\theta$ to all $\theta \in \mathbb{R}$: first define the derivative for $\theta \neq \pi/2 + n\pi$, and then apply the mean value theorem for these isolated points.

Thus $\mathcal{P}_1$ and $\mathcal{P}_2$ are proved for all $\theta \in \mathbb{R}$, and once again all other derivatives of the polar trig functions now follow easily.

Acknowledgments We thank the anonymous referees, whose helpful guidance has significantly improved the article's readability. The first author is supported by the Simons Foundation grant 524418.

Summary. Inspired by the classroom capsule [5] and its follow-up [2], we present direct geometric derivations of the integral formulas for both the inverse hyperbolic and inverse polar trig functions, the latter being a new approach independent of Insel's work [5]. In so doing, we show that these two trigonometric settings are linked in a natural analytical geometric way, establishing a coordinate system determined by hyperbolas that is akin to polar coordinates. We then use the integral formulas to obtain the derivatives of the various hyperbolic and polar trig functions.

REFERENCES

[1] Edwards, B., Larson, R. (2014). *Calculus: Early Transcendental Functions*, 6th ed. Boston: Brooks Cole.

[2] Engbers, J., Hammett, A. (2016). A direct proof of the integral formulae for the inverse hyperbolic functions. *College Math. J.* 47(4): 297–299. doi.org/10.4169/college.math.j.47.4.297

[3] Greenspan, I. (2009). Deriving the hyperbolic trig functions. Presented at the MMC Conference of Workshops, The University of Chicago Laboratory of Schools, Chicago, IL, January 24. talks.isaacgreenspan.com/DerivingTheHyperbolicTrigFunctions.pdf

[4] Haskell, M. W. (1895). On the introduction of the notion of hyperbolic functions. *Bull. Amer. Math. Soc.* 1(6): 155–159. doi.org/10.1090/S0002-9904-1895-00266-9

[5] Insel, A. (1989). A direct proof of the integral formula for arctangent. *College Math. J.* 20(3): 235–237. doi.org/10.1080/07468342.1989.11973238

JOHN ENGBERS (MR Author ID 973477) earned his Ph.D. in graph theory from the University of Notre Dame in 2013. Currently, he is an assistant professor at Marquette University. When not immersed in mathematics, he loves spending time with his wife Ruth and kids Sam, Luke, and Joel, riding his bicycle, and enjoying nature.

ADAM HAMMETT (MR Author ID 840691) earned his Ph.D. in combinatorial probability from The Ohio State University in 2007. Ever since, he has taught mathematics at the college level, and has served as associate professor at Cedarville University since 2015. He enjoys overseeing undergraduate research projects, reading, and spending time outdoors with his wife Rachael and their four children Isabelle, Madison, Daniel, and Esther.

IAN HOGAN (MR Author ID 1396352) completed his Ph.D. in computational algebra in 2017 at Kent State University. He currently works as a software developer focusing on the applications of graph theory to business process optimization. Outside of his broad mathematical interests, he plays Go, is working on cycling through every county in Ohio, and sings with various ensembles.

An Etude on a Putnam Problem

MARCIN MAZUR
Binghamton University
Binghamton, NY 13892-6000
mazur@math.binghamton.edu

The purpose of this article is to answer the following question: what is the smallest constant C_n such that the inequality

$$\int_0^1 |f(t)|dt \le C_n \sup_{t\in[0,1]} |f(t)|$$

holds for all real polynomials f of degree at most n which have a root in $[0, 1]$? Our investigation into this problem was inspired by the special case when $n = 3$, which was posed as problem A6 on the 2016 Putnam Competition [2]. It turned out to be one of the hardest Putnam problems, solved by only 2 of the top 187 contestants. We will see that

$$C_{2k} = 1 - 1/(k + 1)^2$$

and

$$C_{2k-1} = 1 - 1/(k(k + 1))$$

for all natural numbers k.

The restriction to the interval $[0, 1]$ is superficial. A simple change of variables shows that

$$\int_a^b |f(t)|dt \le (b - a)C_n \sup_{t\in[a,b]} |f(t)| \tag{1}$$

for every polynomial f of degree at most n which has a root in the interval $[a, b]$, and C_n is smallest with this property. Knowing C_n allows us to give a nice estimate of the integral of any polynomial of degree at most n, as the following simple theorem shows.

Theorem 1. *Let f be a polynomial of degree at most n, and let M, m be its largest and smallest values on the interval $[a, b]$. Then*

$$C_n m + (1 - C_n)M \le \frac{1}{b - a}\int_a^b f(t)\, dt \le C_n M + (1 - C_n)m.$$

Proof. Note that the polynomials $M - f$ and $f - m$ are non-negative on $[a, b]$, have a root in $[a, b]$ and their largest value on $[a, b]$ is $M - m$. The inequality (1) applied to $f - m$ yields the upper bound, and the lower bound is derived by applying it to $M - f$. ∎

We end the introduction with a short outline of the strategy to compute the number C_n. In the first section we introduce a concept of an n-integrating sequence. Roughly speaking, it is a sequence

$$(x_0, a_0), (x_1, a_1), \ldots, (x_s, a_s)$$

Math. Mag. **93** (2020) 363–373. doi:10.1080/0025570X.2020.1818524
MSC: 26C05

of pairs of real numbers which encodes the integration of polynomials of degree up to n. We show that each n-integrating sequence yields an upper bound for C_n. Specifically, we have

$$C_n \leq |a_1| + \cdots + |a_s|.$$

This leads to the definition of a strong n-integrating sequence, the one for which the upper bound becomes an equality. Lemma 3 establishes key properties of strong n-integrating sequences, including the equality $C_n = 1 - a_0$.

The second section uses linear algebra and basic properties of polynomials to characterize all n-integrating sequences in Theorem 2. In particular, it is proved that the numbers $a_0, \ldots, a_s$ in an n-integrating sequence with $s \leq n$ are determined by the numbers $x_0, \ldots, x_s$, and explicit formulas are given.

The third section builds on the results of the second section and shows that there exists at most one strong n-integrating sequence for each n. Several properties of strong n-integrating sequences are established in Lemma 4, including the value of a_0. This yields the desired expression for C_n provided a strong n-integrating sequence exists. The existence is shown in the last section.

Strategy

We start by making some straightforward but useful observations. Let D_n be the smallest constant such that the inequality

$$\left| \int_0^1 f(t)\, dt \right| \leq D_n \sup_{t \in [0,1]} |f(t)|$$

holds for all real polynomials f of degree at most n such that $f(0) = 0$. Clearly $D_n \leq C_n$. Let g be a polynomial of degree at most n such that $g(a) = 0$. For any b, the polynomial $f(x) = g(a + (b - a)x)$ has degree at most n and satisfies $f(0) = 0$. A simple integration by substitution yields

$$\begin{aligned} \left| \int_a^b g(t)\, dt \right| &= |b - a| \left| \int_0^1 f(t)\, dt \right| \\ &\leq |b - a| D_n \sup_{t \in [0,1]} |f(t)| \\ &= |b - a| D_n \sup_{t \in [a,b]} |g(t)|. \end{aligned}$$

Note that this works for $b < a$ as well as for $b \geq a$. Suppose f is a polynomial of degree at most n and

$$0 \leq r_1 < r_2 < \cdots < r_s \leq 1$$

are all the roots of f in $[0, 1]$, $s \geq 1$ (so f has at least one root in $[0, 1]$). Set $r_0 = 0$ and $r_{s+1} = 1$, and apply the above observation to f relative to the interval $[r_i, r_{i+1}]$, where f is either non-negative or non-positive. We get

$$\begin{aligned} \int_{r_i}^{r_{i+1}} |f(t)|\, dt &= \left| \int_{r_i}^{r_{i+1}} f(t)\, dt \right| \\ &\leq (r_{i+1} - r_i) D_n \sup_{t \in [r_i, r_{i+1}]} |f(t)|. \end{aligned}$$

Since $\sup_{t\in[r_i,r_{i+1}]} |f(t)| \le \sup_{t\in[0,1]} |f(t)|$, we have

$$\begin{aligned}\int_0^1 |f(t)|\,dt &= \sum_{i=0}^{s} \int_{r_i}^{r_{i+1}} |f(t)|\,dt \\ &\le D_n \sup_{t\in[0,1]} |f(t)| \sum_{i=0}^{s} (r_{i+1}-r_i) \\ &= D_n \sup_{t\in[0,1]} |f(t)|.\end{aligned}$$

It follows that $C_n \le D_n$ and consequently $C_n = D_n$.

The equality $C_n = D_n$ allows us to focus on integrals of polynomials which vanish at 0. In order to estimate such integrals we introduce the following concept: A sequence

$$(x_0, a_0), (x_1, a_1), \ldots, (x_s, a_s)$$

of pairs of real numbers is called *n-integrating* if

$$0 = x_0 < x_1 < \cdots < x_s \le 1$$

and

$$\int_0^1 f(t)\,dt = \sum_{i=0}^{s} a_i f(x_i) \tag{2}$$

for every polynomial f of degree at most n.

Our strategy to compute C_n is based on the following observation.

Lemma 1. *Suppose that* $(x_0, a_0), \ldots, (x_s, a_s)$ *is an n-integrating sequence. Then* $C_n \le \sum_{i=1}^{s} |a_i|$.

Proof. For polynomials f of degree at most n such that $f(0) = 0$ we have

$$\begin{aligned}\left|\int_0^1 f(t)\,dt\right| &= \left|\sum_{i=1}^{s} a_i f(x_i)\right| \\ &\le \left(\sum_{i=1}^{s} |a_i|\right) \sup_{t\in[0,1]} |f(t)|. \end{aligned} \tag{3}$$

Thus $C_n = D_n \le \sum_{i=1}^{s} |a_i|$. ∎

Example 1. It is easy to check that the equality

$$\int_0^1 f(t)dt = \frac{1}{4} f(0) + \frac{3}{4} f(2/3)$$

holds for any polynomial of degree at most 2. Thus $\left(0, \frac{1}{4}\right), \left(\frac{2}{3}, \frac{3}{4}\right)$ is a 2-integrating sequence and $C_2 \le 3/4$.

It is not hard to see that our proof of Lemma 1 would yield the equality

$$C_n = \sum_{i=1}^{s} |a_i|$$

if the inequality in (3) was actually an equality for some polynomial $f \neq 0$. This observation prompts the following definition: A *strong n-integrating sequence* is an n-integrating sequence $(x_0, a_0), \ldots, (x_s, a_s)$ for which there is a non-zero polynomial h of degree at most n such that $h(0) = 0$ and

$$\int_0^1 h(t)\, dt = \sum_{i=1}^{s} |a_i| \sup_{t\in[0,1]} |h(t)|. \tag{4}$$

Any such h will be called an *extreme polynomial* for the n-integrating sequence.

Lemma 2. *If $(x_0, a_0), \ldots, (x_s, a_s)$ is a strong n-integrating sequence then $C_n = \sum_{i=1}^{s} |a_i|$.*

Proof. We already know that $C_n \leq \sum_{i=1}^{s} |a_i|$. On the other hand, if h is an extreme polynomial for our n-integrating sequence then

$$\begin{aligned} C_n \sup_{t\in[0,1]} |h(t)| &\geq \int_0^1 |h(t)|\, dt \\ &\geq \int_0^1 h(t)dt = \sum_{i=1}^{s} |a_i| \sup_{t\in[0,1]} |h(t)|. \end{aligned}$$

Thus $C_n \geq \sum_{i=1}^{s} |a_i|$. It follows that $C_n = \sum_{i=1}^{s} |a_i|$ and that

$$\begin{aligned} \int_0^1 |h(t)|\, dt &= \int_0^1 h(t)\, dt \\ &= \sum_{i=0}^{s} a_i h(x_i) = \sum_{i=1}^{s} |a_i| \sup_{t\in[0,1]} |h(t)|. \end{aligned} \tag{5}$$

■

Example 2. *Consider $h(t) = t(4/3 - t)$. It is easy to verify that $h(0) = 0$,*

$$\sup_{t\in[0,1]} |h(t)| = h\left(\frac{2}{3}\right) = \frac{4}{9},$$

and

$$\int_0^1 h(t)\, dt = \frac{1}{3} = \frac{3}{4} h\left(\frac{2}{3}\right).$$

Thus the sequence in Example 1 is strong 2-integrating and $C_2 = 3/4$.

The equalities in (5) impose constraints on the numbers x_i, a_i, and the extreme polynomial h. We list them in the following lemma.

Lemma 3. *Suppose that $(x_0, a_0), \ldots, (x_s, a_s)$ is a strong n-integrating sequence with extreme polynomial h. Then*

(i) *$a_0, \ldots, a_n$ are non-negative and $a_0 + a_1 + \cdots + a_s = 1$,*
(ii) *$C_n = 1 - a_0$,*
(iii) *$h(t) \geq 0$ for all $t \in [0, 1]$ and*

$$h(x_1) = \cdots = h(x_s) = \sup_{t\in[0,1]} |h(t)|,$$

(iv) *either* $x_s = 1$ *and* $2s \le n+1$ *or* $2s \le n$.

Proof. From (2) applied to a non-zero constant polynomial we get

$$a_0 + a_1 + \cdots + a_s = 1$$

for any n-integrating sequence.

The first equality in (5) shows that h is non-negative on $[0, 1]$. From the last equality in (5) we then get that $a_1, \ldots, a_s$ are non-negative and

$$h(x_1) = \cdots = h(x_s) = \sup_{t\in[0,1]} |h(t)|.$$

Thus

$$C_n = a_1 + \cdots + a_s = 1 - a_0$$

by Lemma 2 . Since $C_n \le 1$, we get $a_0 \ge 0$. This establishes (i), (ii), and (iii).

From (iii) we see that h has local maxima at $x_1, \ldots, x_{s-1}$, and it must have a local minimum in each of the intervals (x_i, x_{i+1}), $i = 1, \ldots, s-1$. Furthermore, either $x_s = 1$ or h has a local maximum at x_s. Putting these together, we see that either $x_s = 1$ and h' has at least $2(s-1)$ roots in $(0, 1)$ or h' has at least $2(s-1)+1$ roots in $(0, 1)$. Since a non-zero polynomial has no more roots than its degree, we see that $2(s-1) \le n-1$ in the former case and $2(s-1)+1 \le n-1$ in the latter case. This proves part (iv). ■

Using Lemma 3, it is not hard to find strong n-integrating sequences and extreme polynomials for small values of n.

Example 3. *When* $n = 1$ *then* $(0, 1/2)$, $(1, 1/2)$ *is a strong* 1*-integrating sequence with extreme polynomial* $h(t) = t$. *Thus* $C_1 = 1/2$.

The case $n = 2$ *is discussed in Examples 1 and 2.*

When $n = 3$*,we have that*

$$\left(0, \frac{1}{6}\right), \quad \left(\frac{1}{2}, \frac{2}{3}\right), \quad \left(1, \frac{1}{6}\right)$$

is a strong 3*-integrating sequence with extreme polynomial* $h(t) = t^3 - 2t^2 + 5t/4$. *Thus* $C_3 = 5/6$. *The reader is encouraged to write down a detailed argument that* $C_3 = 5/6$ *before reading further.*

While it is a straightforward exercise to check that Example 3 lists strong n-integrating sequences for $n = 1, 2, 3$, it is far from obvious how to proceed for larger values of n. In order to gain some insight into this problem we move our attention to n-integrating sequences. It turns out that some basic linear algebra provides a very convenient language for this purpose.

A helping hand from linear algebra

The set of all real polynomials of degree at most n is a vector space P_n of dimension $n+1$. We set $L(f) = \int_0^1 f(t)\,dt$. We can think of L as a linear functional

$$L : P_n \longrightarrow \mathbb{R}$$

on P_n. For each $u \in [0, 1]$, define $L_u(f) = f(u)$, the evaluation of f at u. Then L_u is also a linear functional on P_n. Recall that the set of all linear functionals on P_n is also

a vector space of dimension $n+1$, called the dual space of P_n and denoted by P_n^*; see [1, Section 2.6].

Consider now an n-integrating sequence $(x_0, a_0), \ldots, (x_s, a_s)$. The condition (2) means that L is a linear combination

$$L = a_0 L_{x_0} + a_1 L_{x_1} + \cdots + a_s L_{x_s}$$

of the functionals $L_{x_0}, L_{x_1}, \ldots, L_{x_s}$. We leave it as an exercise for the reader to show that if $s \leq n$, then the functionals $L_{x_0}, L_{x_1}, \ldots, L_{x_s}$ are linearly independent. In particular, when $s \leq n$, the numbers a_i are uniquely determined by $x_0, \ldots, x_s$. Moreover, if $s = n$, then for any choice of

$$0 \leq x_0 < x_1 < \cdots < x_n \leq 1$$

the functionals $L_{x_0}, L_{x_1}, \ldots, L_{x_s}$ form a basis of P_n^*. Thus, when $x_0 = 0$, there are unique numbers $a_0, \ldots, a_n$ such that $(x_0, a_0), \ldots, (x_n, a_n)$ is an n-integrating sequence. This shows that there are many n-integrating sequences.

The following proposition, whose proof is left as another exercise for the reader, allows us to characterize all n-integrating sequences.

Proposition 1. *Let V be a finite dimensional vector space, and let $L, L_0, \ldots, L_s$ be linear functionals on V. Then L is a linear combination of $L_0, \ldots, L_s$ if and only if whenever*

$$L_0(v) = L_1(v) = \cdots = L_s(v) = 0,$$

we have that $L(v) = 0$. In other words, L is a linear combination of $L_0, \ldots, L_s$ if and only if the intersection of the kernels of $L_0, \ldots, L_s$ is contained in the kernel of L.

Note that the polynomials $f(t)$ such that $L_{x_i}(f) = 0$ for $i = 0, \ldots, s$ are exactly those divisible by

$$(t - x_0)(t - x_1) \ldots (t - x_s).$$

Assume from now on that $s < n$. Then the polynomials

$$(t - x_0)(t - x_1) \ldots (t - x_s)t^j$$

for $j = 0, 1, \ldots, n - s - 1$ form a basis of the space of all polynomials on which each of the functionals $L_{x_0}, L_{x_1}, \ldots, L_{x_s}$ vanishes. This observation together with Proposition 1 give us the following result:

Theorem 2. *Let $s < n$. Given $0 \leq x_0 < x_1 < \cdots < x_s \leq 1$ the equality (2) is possible if and only if*

$$\int_0^1 (t - x_0)(t - x_1) \ldots (t - x_s)t^j \, dt = 0 \tag{6}$$

for $j = 0, 1, \ldots, n - s - 1$. Let

$$g(t) = (t - x_0)(t - x_1) \ldots (t - x_s).$$

The coefficients $a_0, \ldots, a_s$ in (2) are determined by the following formulas:

$$a_i = \frac{1}{g'(x_i)} \int_0^1 \frac{g(t)}{t - x_i} \, dt. \tag{7}$$

Proof. Everything except for the formula (7) has been established in the discussion preceding the theorem. The formula (7) for a_i follows immediately from (2) applied to $f(t) = g(t)/(t - x_i)$ and the observation that $f(x_i) = g'(x_i)$ and $f(x_j) = 0$ for $j \neq i$. ■

A Hero is Born

For any continuous function f on an open interval containing $[0, 1]$, define $f^{(0)} = f$ and

$$f^{(-k-1)}(t) = \int_0^t f^{(-k)}(s)\, ds,$$

for any natural number k. Recall that if k is positive and f is a k-times differentiable function, we write $f^{(k)}$ for the kth derivative of f. Note that $f^{(k+1)}$ is the derivative of $f^{(k)}$ for every integer k and that $f^{(k)}(0) = 0$ for all negative values of k. It is now a simple exercise in integration by parts to see that

$$\int_0^1 f(t)t^j dt = \sum_{i=1}^{j+1} (-1)^{i-1} \frac{j!}{(j+1-i)!} f^{(-i)}(1). \tag{8}$$

Proposition 2. *Let f be a continuous function on an open interval containing* $[0, 1]$.

(i) *We have that*

$$\int_0^1 f(t)t^j\, dt = 0, \quad \text{for } j = 0, \ldots, m-1,$$

if and only if $f^{(-j)}(1) = 0$ for $j = 1, \ldots, m$.

(ii) *If*

$$\int_0^1 f(t)t^j\, dt = 0, \quad \text{for } j = 0, \ldots, m-1,$$

then f has at least m distinct zeros in $(0, 1)$.

Proof. The first part follows easily from (8).

For the second part, assume that f has less than m zeros in $(0, 1)$. Note that

$$\int_0^1 f(t)p(t)\, dt = 0$$

for every polynomial p of degree less than m. By replacing f by $-f$ if necessary, we may assume that f is positive near 0. Then there are numbers

$$0 = u_0 < u_1 < \cdots < u_k < u_{k+1} = 1$$

such that $k < m$, f is non-negative on $[u_{2i}, u_{2i+1}]$, and f is non-positive on $[u_{2i+1}, u_{2i+2}]$ for $i = 0, 1, \ldots$. But then

$$f(t)(t - u_1) \ldots (t - u_k)(-1)^k$$

is non-negative on $[0, 1]$ and

$$\int_0^1 f(t)(t - u_1) \ldots (t - u_k)(-1)^k dt = 0.$$

This implies that $f = 0$, which is a contradiction. ■

Suppose now that $(x_0, a_0), \ldots, (x_s, a_s)$ is a strong n-integrating sequence. By (6) and Proposition 2 applied to

$$f(t) = t(t - x_1)\ldots(t - x_s),$$

we see that f has at least $n - s$ roots in $(0, 1)$. Thus $n - s \leq s$, that is, $2s \geq n$. On the other hand, part (iv) of Lemma 3 says that either $2s \leq n$ or $2s \leq n + 1$ and $x_s = 1$. When $n = 2k$ is even, this forces $s = k$. When $n = 2k + 1$ is odd, we must have $s = k + 1$ and $x_s = 1$. In other words, $s = \lfloor \frac{n+1}{2} \rfloor$ for all n and $x_s = 1$ when n is odd. Note that $n - s = k$. By part (i) of Proposition 2 we see that 1 is a root of $f^{(-k)}$ of multiplicity at least k when n is even, and of multiplicity at least $k + 1$ when n is odd (since $x_s = 1$). Also, 0 is a root of $f^{(-k)}$ of multiplicity at least $k + 1$. Hence, when n is even, $f^{(-k)}(t)$ is a polynomial of degree $2k + 1$ which is divisible by $t^{k+1}(t - 1)^k$. It follows that

$$f^{(-k)}(t) = b_n t^{k+1}(t - 1)^k,$$

where b_n is the leading coefficient of $f^{(-k)}(t)$. Similarly, when n is odd, $f^{(-k)}(t)$ is a polynomial of degree $2k + 2$ which is divisible by $t^{k+1}(t - 1)^{k+1}$. It follows that

$$f^{(-k)}(t) = b_n t^{k+1}(t - 1)^{k+1}.$$

Since $f(t)$ is monic of degree $s + 1 = n - k + 1$, the leading coefficient of $f^{(-k)}(t)$ is equal to

$$\begin{aligned} b_n &= \frac{1}{(n - k + 2)(n - k + 3)\ldots(n + 1)} \\ &= \frac{(n + 1 - k)!}{(n + 1)!} = \frac{(s + 1)!}{(n + 1)!}. \end{aligned}$$

Summarizing what we have just discovered, let $B_n(t)$ be the polynomials defined as:

$$B_n(t) = \frac{(s + 1)!}{(n + 1)!} t^{k+1}(t - 1)^s, \tag{9}$$

where $s = \lfloor \frac{n+1}{2} \rfloor$ and $k = n - s$. Let $g_n = B_n^{(k)}$ be the kth derivative of B_n.

Lemma 4. *Let $(x_0, a_0), \ldots, (x_s, a_s)$ be a strong n-integrating sequence. Then $s = \lfloor \frac{n+1}{2} \rfloor$ and*

$$g_n(t) = t(t - x_1)\ldots(t - x_s).$$

Thus, the numbers x_i are uniquely determined as the roots of g_n and

$$a_i = \frac{1}{g_n'(x_i)} \int_0^1 \frac{g_n(t)}{t - x_i}\, dt \tag{10}$$

for $i = 0, \ldots, s$. In particular, $a_0 = \dfrac{1}{(s + 1)(k + 1)}$ and $C_n = 1 - a_0$.

Proof. The first part of the lemma has already been established, and (10) follows from (7). It remains to evaluate

$$a_0 = \frac{1}{g_n'(0)} \int_0^1 \frac{g_n(t)}{t}\, dt.$$

Since $g_n' = B_n^{(k+1)}$, where $k = n - s$, the number $g_n'(0)$ is the constant term of $B_n^{(k+1)}$. It follows easily from (9) that

$$g_n'(0) = (-1)^s \frac{(s+1)!(k+1)!}{(n+1)!}.$$

Note that the polynomial $t^{-i} B_n^{(k-i)}$ vanishes at 0 and 1 for $i = 1, \ldots, k$. Using this and integration by parts k-times we get

$$\begin{aligned}\int_0^1 B_n^{(k)}(t) t^{-1}\, dt &= k! \int_0^1 B_n(t) t^{-(k+1)}\, dt \\ &= \frac{k!(s+1)!}{(n+1)!} \int_0^1 (t-1)^s\, dt \\ &= (-1)^s \frac{k!(s+1)!}{(s+1)(n+1)!}.\end{aligned}$$

Thus,

$$\begin{aligned}a_0 &= \frac{1}{g_n'(0)} \int_0^1 B_n^{(k)}(t) t^{-1}\, dt \\ &= (-1)^s \frac{(n+1)!}{(s+1)!(k+1)!} (-1)^s \frac{k!(s+1)!}{(s+1)(n+1)!} \\ &= \frac{1}{(s+1)(k+1)}.\end{aligned}$$

■

Note that the value of C_n in Lemma 4 agrees with what we announced in the introduction.

Finale

Lemma 4 tells us that for each n there is at most one strong n-integrating sequence and, if it exists, then the value of C_n announced in the Introduction is correct. In order to complete our argument we need to establish that the sequence $(x_0, a_0), \ldots, (x_s, a_s)$, where $x_0, \ldots, x_s$ are the roots of g_n and the numbers $a_0, \ldots, a_s$ are defined by (10), is indeed a strong n-integrating sequence.

Our first step is to show that g_n has $s+1$ roots $0 = x_0 < x_1 < \cdots < x_s \le 1$, that $x_s = 1$ when n is odd, that the numbers $a_0, \ldots, a_s$ are positive, and that the sequence $(x_0, a_0), \ldots, (x_s, a_s)$ is n-integrating. Recall that $s = \lfloor \frac{n+1}{2} \rfloor$ and $k = n - s$. It follows easily from the definition of g_n that $g_n(0) = 0$ for all n and $g_n(1) = 0$ for n odd. Moreover, for $1 \le j \le k$, the $(k-j)$th derivative $B_n^{(k-j)}$ of B_n, which is

the same as $g_n^{(-j)}$, vanishes at 0 and 1. It follows by part (i) of Proposition 2 that

$$\int_0^1 g_n(t)t^j\,dt = 0, \quad \text{for } j = 0, \ldots, k-1.$$

By part (ii) of Proposition 2, the polynomial g_n has at least k distinct roots in the interval $(0, 1)$. Since g_n is monic of degree $s+1$, we have that $g(0) = 0$, and $g(1) = 0$ for n odd, we conclude that

$$g_n(t) = (t - x_0)(t - x_1)\ldots(t - x_s),$$

where $0 = x_0 < x_1 < \cdots < x_s \le 1$ and $x_s = 1$ when n is odd. By Theorem 2, the sequence $(x_0, a_0), \ldots, (x_s, a_s)$ is n-integrating. When n is even, (2) applied to the polynomial $f(t) = g_n^2(t)/(t - x_i)^2$, which has degree n and is non-negative on $[0, 1]$, yields that a_i is positive. Similarly, when n is odd, (2) applied to the polynomial

$$f(t) = \frac{g_n^2(t)}{((1-t)(t - x_i)^2},$$

which has degree n and is non-negative on $[0, 1]$, shows that a_i is positive. Thus all the numbers $a_0, \ldots, a_s$ are positive.

It remains to show the existence of an extreme polynomial h for

$$(x_0, a_0), \ldots, (x_s, a_s).$$

Part (iii) of Lemma 3 will guide us to finding out what h has to be. Indeed, we know that $h(0) = 0$, the derivative of h vanishes at $x_1, \ldots, x_k$, and it vanishes at x_{k+1} when n is odd. Also,

$$h(x_1) = \cdots = h(x_s) = M$$

for some M (equal to $\sup_{t\in[0,1]} |h(t)|$). It follows that the polynomial $h - M$ is divisible by

$$(t - x_1)^2 \ldots (t - x_k)^2$$

when $n = 2k$ is even and it is divisible by

$$(t - x_1)^2 \ldots (t - x_k)^2(t - 1)$$

when $n = 2k + 1$ is odd. Since h has degree at most n, we see that we do not have much choice for h. Let $M_n = (x_1 \ldots x_k)^2$, and define the polynomials h_n as follows:

$$h_n(t) = \begin{cases} M_n - (t - x_1)^2 \ldots (t - x_k)^2, & \text{if } n = 2k, \\ M_n + (t - x_1)^2 \ldots (t - x_k)^2(t - 1), & \text{if } n = 2k + 1. \end{cases} \tag{11}$$

It is clear that $h_n(0) = 0$ and M_n is the largest value of h_n on $[0, 1]$. Since the sequence $(x_0, a_0), \ldots, (x_s, a_s)$ is n-integrating and the numbers $a_0, a_1, \ldots, a_s$ are positive, we have

$$\begin{aligned} \int_0^1 h_n(t)\,dt &= (a_1 + \cdots + a_s)M_n \\ &= (|a_1| + \cdots + |a_s|)M_n \ge C_n M_n. \end{aligned}$$

(recall that

$$C_n \le |a_1| + \cdots + |a_s|$$

by Lemma 1). On the other hand, if m_n is the smallest value of h_n on [0, 1], then

$$\int_0^1 h_n(t)dt \le C_n M_n + (1 - C_n)m_n$$

by Theorem 1. It follows that $m_n \ge 0$. Consequently, h_n is non-negative on [0, 1],

$$M_n = \sup_{t\in[0,1]} |h_n(t)|,$$

and therefore h_n is an extreme polynomial for $(x_0, a_0), \ldots, (x_s, a_s)$. This completes our proof of the following result, announced in the introduction:

Theorem 3. *The smallest number C_n such that*

$$\int_0^1 |f(t)|\, dt \le C_n \sup_{t\in[0,1]} |f(t)|$$

holds for all real polynomials f of degree at most n which have a root in [0, 1] *is given by the following formula:*

$$C_n = \begin{cases} 1 - \frac{1}{k(k+1)}, & \text{if } n = 2k - 1, \\ 1 - \frac{1}{(k+1)^2}, & \text{if } n = 2k. \end{cases}$$

We end this etude with the following exercise for the reader.

Exercise. Prove that if $n \ge 2$ and a non-zero polynomial f of degree at most n has a root in [0, 1] and satisfies

$$\int_0^1 |f(t)|\, dt = C_n \sup_{t\in[0,1]} |f(t)|$$

then either $f(t) = c \cdot h_n(t)$ or $f(t) = c \cdot h_n(1 - t)$ for some non-zero constant c.

Acknowledgments The author would like to express his deep gratitude to the anonymous referees for providing many corrections, comments, and suggestions which led to significant improvements of the exposition. This work was partially supported by a grant from the Simons Foundation (# 245855).

REFERENCES

[1] Friedberg, S. H., Insel, A. J., Spence, L. E. (2003). *Linear Algebra*, 4th ed. Upper Saddle River, NJ: Pearson.
[2] Klosinski, L. F., Alexanderson, G. L., Krusemeyer, M. (2017). *The Seventy-Seventh William Lowell Putnam Mathematical Competition*. *Amer. Math. Monthly* 124(8): 675–687.

Summary. We discuss a generalization of Problem A6 from the 2016 Putnam Competition, which asks to find the smallest constant C_n such that for any polynomial f of degree n which has a root in the interval [0, 1], the integral of $|f|$ on [0, 1] is bounded above by C_n times the largest value of $|f|$ on [0, 1]. Using several basic facts about integrals, polynomials, and linear algebra, we present a line of thought which naturally leads to the discovery of the answer and to several other interesting results.

MARCIN MAZUR (MR Author ID: 851571) received his masters degree in mathematics from Warsaw University (Poland) and his Ph.D. in mathematics from the University of Chicago. He is currently a professor of mathematics and the Chair of the Department of Mathematical Sciences at Binghamton University.

The *Who Wants to Be a Millionaire* Dilemma

MATTHEW DUREY
Massachusetts Institute of Technology
Cambridge, MA 02139
mdurey@mit.edu

In Season 2 of the game show *Who Wants to be a Millionaire*, contestant Bernie Cullen achieved the remarkable feat of reaching the million-dollar question having only used the *Phone a Friend* "lifeline." Being uncertain of the correct answer from the four options, Cullen decided to make use of the two remaining lifelines: *50:50*, in which two incorrect answers are randomly eliminated, and *Ask the Audience*, where each audience member votes on the remaining options. The question remained: In which order should he play the lifelines?

The host, Regis Philbin, advised Cullen to play the *50:50* lifeline first so that the audience would have half as many options from which to choose. In contrast, Cullen argued that the incorrect votes would be spread over the three other options, allowing the correct answer to take a greater ratio of the audience vote. Following his own strategy, Cullen first played the *Ask the Audience* lifeline, where one answer received a majority vote of 69%. Crucially, this answer remained an option after playing the *50:50* lifeline. Cullen followed the audience majority, correctly answered the question, and triumphantly won one million dollars!

For the reader unfamiliar with this game show, we provide a brief outline of the format. The contestant has to correctly answer a series of questions, which, in Season 2, were of increasing value (ranging from \$100 to \$1 million) and of increasing difficulty. If the contestant is stuck on any question, then they may use some, or all, of the three lifelines (*Ask the Audience*, *50:50*, and *Phone a Friend*) before answering. Each lifeline may only be played once per game. The game ends when either (1) the contestant gives an incorrect answer (at which point they leave with the value of the last milestone passed (\$1000 or \$32,000)), or (2) they choose to walk away before answering, earning exactly the value of the last correctly answered question, or (3) they answer all fifteen questions correctly and win \$1 million. If Cullen had answered the final question incorrectly then he would have taken away a measly(!) \$32,000, demonstrating his confidence in the audience vote.

In this note, we discuss the expected success of each strategy based on how well-informed, misinformed, or uninformed the audience is for a particular question. As in the game show, we label the four options A, B, C, and D. We denote the fraction of the audience that believes the answer to be A by a, where $0 \leq a \leq 1$, and similarly for B, C, and D. We denote the fraction of the audience members who have no idea what the answer is by n, with $0 \leq n \leq 1$. Notice that we have the constraint $a + b + c + d + n = 1$. For the sake of argument, we set A as the correct answer. We assume that when an audience member has no clue what the answer is (either because they did not know originally, or because their prior belief has been eliminated following the use of the *50:50* lifeline), they select an answer at random, uniformly over the remaining available options. All audience members must vote, and act independently of each other (so no copying one's neighbor!).

Math. Mag. **93** (2020) 374–377. doi:10.1080/0025570X.2020.1826848
MSC: Primary 60-01, Secondary 90-01

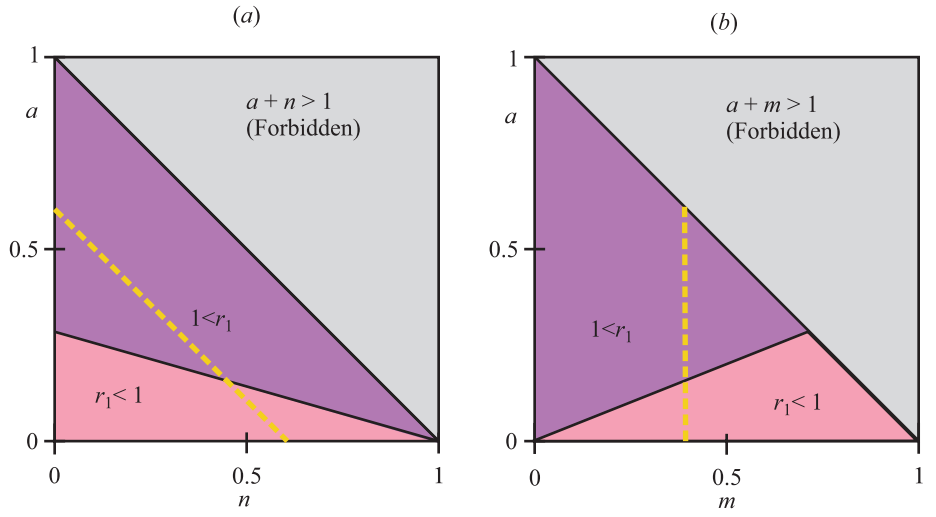


Figure 1 (a) Delineation in the (a, n)-plane of the success of Strategy 1, which is expected to be successful if $r_1 > 1$, and unsuccessful otherwise. The parameters a and n obey the constraint $a + n \leq 1$. The dotted line corresponds to the particular case $m = 0.4$, where $a + m + n = 1$. (b) The exact same data portrayed in the (a, m)-plane, subject to the constraint $a + m \leq 1$.

Strategy 1: *50:50,* Then *Ask the Audience*

To analyze this strategy, we first consider the case where the two remaining answers are A and B. A fraction $c + d$ of misinformed audience members find themselves to be wrong and instead vote for A or B with equal probability. In contrast, those that believe the answer is A or B remain convinced that they are correct. The expected fraction, $\mathbb{E}_{\text{corr|B}}$, of people to vote for the correct answer (A) given that B is retained following the use of the *50:50* lifeline is

$$\mathbb{E}_{\text{corr|B}} = a + \frac{1}{2}(c + d + n) = \frac{1 + a - b}{2},$$

while the expected fraction, $\mathbb{E}_{\text{inc|B}}$, that incorrectly vote for B is

$$\mathbb{E}_{\text{inc|B}} = b + \frac{1}{2}(c + d + n) = \frac{1 - a + b}{2},$$

where we have used that $a + b + c + d + n = 1$. The expected fraction of audience members to vote for the correct and incorrect answers given that C or D remains after playing the *50:50* lifeline may be computed similarly (where b in the above expressions is replaced by c or d, respectively).

Since each of the incorrect answers B, C, or D, may remain following the use of the *50:50* lifeline, each with probability $\frac{1}{3}$, the expected fraction of correct votes is

$$\mathbb{E}_{\text{corr}} = \frac{1}{3}\Big(\mathbb{E}_{\text{corr|B}} + \mathbb{E}_{\text{corr|C}} + \mathbb{E}_{\text{corr|D}}\Big),$$

and similarly for the incorrect votes, $\mathbb{E}_{\text{inc}}$. (This intuitive step may be formally justified by means of the law of total expectation [1, 2].) After some algebra, we obtain

$$\mathbb{E}_{\text{corr}} = \frac{1}{3} + \frac{2}{3}a + \frac{1}{6}n$$

and

$$\mathbb{E}_{\text{inc}} = 1 - \mathbb{E}_{\text{corr}} = \frac{2}{3} - \frac{2}{3}a - \frac{1}{6}n.$$

The ratio, r_1, of expected correct to incorrect votes is

$$r_1 = \frac{\mathbb{E}_{\text{corr}}}{\mathbb{E}_{\text{inc}}} = \frac{2 + 4a + n}{4 - 4a - n}.$$

For the contestant to expect success playing this strategy (corresponding to $r_1 > 1$), they require

$$2 + 4a + n > 4 - 4a - n,$$

which simplifies to $4a + n > 1$. In other words, there must be a sufficiently large fraction of well-informed, a, and uninformed, n, audience members, as presented in Figure 1(a). In terms of the fraction, m, of misinformed audience members, where

$$m = b + c + d \qquad \text{and} \qquad a + m + n = 1,$$

the requirement $4a + n > 1$ is equivalent to $a > \frac{1}{3}m$ (see Figure 1(b)): The number of well-informed voters must exceed the average fraction of misinformed voters.

Strategy 2: *Ask the Audience*, then *50:50*

Employing this alternative strategy, we expect a fraction $\mathbb{E}_{\text{A}} = a + \frac{1}{4}n$ of audience members to vote for A, a fraction $\mathbb{E}_{\text{B}} = b + \frac{1}{4}n$ to vote for B, and so forth. Hence, the expected fraction of correct votes is $\mathbb{E}_{\text{corr}} = \mathbb{E}_{\text{A}}$. The probability that any of the incorrect answers is selected to appear in the *50:50* is $\frac{1}{3}$, so the expected fraction of votes for the remaining incorrect option is

$$\mathbb{E}_{\text{inc}} = \frac{1}{3}\left(\mathbb{E}_{\text{B}} + \mathbb{E}_{\text{C}} + \mathbb{E}_{\text{D}}\right) = \frac{1 - a}{3} - \frac{n}{12}.$$

Following this second strategy, the ratio, r_2, of expected correct to incorrect votes is

$$r_2 = \frac{\mathbb{E}_{\text{corr}}}{\mathbb{E}_{\text{inc}}} = \frac{12a + 3n}{4 - 4a - n}.$$

For the contestant to expect success when applying this strategy, they require $r_2 > 1$, corresponding to $4a + n > 1$. Surprisingly, this condition for expected success is identical to that of Strategy 1! We conclude that, given a and n, either both strategies are expected to succeed, or both strategies are expected to fail, as depicted in Figure 2.

Which of the two strategies yield the greatest ratio of correct to incorrect answers, thus giving the contest the greatest confidence in the audience vote? A short calculation shows that the condition $r_2 > r_1$ is equivalent to $4a + n > 1$: When both strategies succeed, Strategy 2 will give the sharper signal, consistent with Cullen's assertion.

Conclusion

The foregoing analysis provides a pedagogical example of decision theory to a familiar and inviting problem, utilizing the well-known law of total expectation [**1**,**2**]. There are

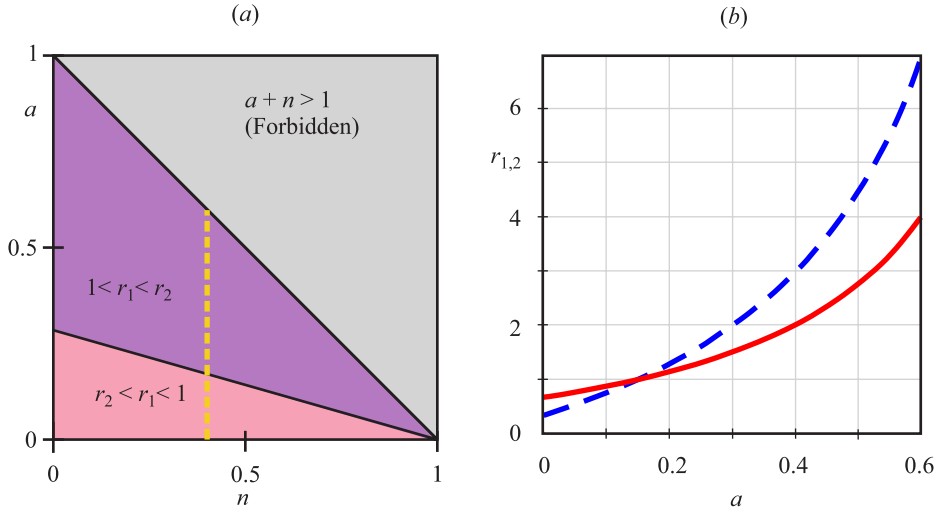


Figure 2 (a) Delineation in the (a, n)-plane of the success of each strategy. Strategy j is expected to be successful if $r_j > 1$. The parameters a and n obey the constraint $a + n \leq 1$. (b) The values of r_1 (solid curve) and r_2 (dashed curve) as a function of a for $n = 0.4$, corresponding to the dotted line in (a).

several extensions to this problem that one could consider, for example: the variance of each strategy, giving a measure of uncertainty; audience votes with spatial dependence, corresponding to uncertain voters copying their neighbors; replacement of the *50:50* lifeline by the *Double Dip* lifeline (which was used in later series of the game show), in which the contestant may give a second answer if their first answer is incorrect. Similar analyses may be performed for a wide class of problems, such as the behavior of voters in political campaigns when various candidates are eliminated at different stages of the campaign.

REFERENCES

[1] Grimmett, G., Stirzaker, D. (2001). *Probability and Random Processes*. Oxford: Oxford University Press.
[2] Grimmett, G., Welsh, D. (2014). *Probability: An Introduction*. Oxford: Oxford University Press.

Summary. We compare the expected effectiveness of two strategies arising from the game show *Who Wants to be a Millionaire*. If a contestant intends to use both the *50:50* and *Ask the Audience* lifelines on a single question, in which order should they play each lifeline? We analyze both of these strategies using elementary techniques from probability theory.

MATTHEW DUREY (ORCID 0000-0002-4232-1705) is an Instructor in Applied Mathematics at Massachusetts Institute of Technology. His research interests include fluid mechanics and dynamical systems.

Almost Equilateral Heronian Triangles

ROGER B. NELSEN
Lewis & Clark College
Portland, OR 97219
nelsen@lclark.edu

A *Heronian triangle* is one whose sides (a, b, c) and area K are integers. An *almost equilateral Heronian triangle* is a Heronian triangle whose sides are consecutive integers such as (3, 4, 5), with area 6, and (13, 14, 15), with area 84 [1]. Are there others? If so, how many?

To answer, we first recall that the area K of an arbitrary triangle with sides a, b, c is given by Heron's formula:

$$K = \sqrt{s(s-a)(s-b)(s-c)},$$

where s denotes the *semiperimeter* $s = (a+b+c)/2$. Letting $(b-1, b, b+1)$ denote the sides of an almost equilateral Heronian triangle yields $s = 3b/2$ and

$$K^2 = \left(\frac{3b}{2}\right)\left(\frac{b+2}{2}\right)\left(\frac{b}{2}\right)\left(\frac{b-2}{2}\right)$$

or

$$16K^2 = 3b^2(b^2-4).$$

Hence, b must be even. If h denotes the altitude to the even side b, then we also have $K = bh/2$. Thus,

$$3b^2(b^2-4) = 4b^2h^2,$$

or

$$3b^2 - (2h)^2 = 12.$$

We now prove the following theorem:

Theorem 1. *Infinitely many almost equilateral Heronian triangles exist.*

Proof. In Figure 1, we have an equilateral triangle with side $3b + 4h$, where b and h are integers. The triangle is divided into three equilateral triangles each in two shades of gray with side $2b + 2h$, overlapping pairwise in three equilateral triangles each in dark gray with side b, and a white equilateral triangle with side $2h$ in the center.

Using the inclusion-exclusion principle to count the small $\triangle$ and $\triangledown$ in the figure (noting that the total number of $\triangle$ and $\triangledown$ triangles in a larger equilateral triangle is the square of its side length) yields

$$(3b+4h)^2 = 3(2b+2h)^2 - 3b^2 + (2h)^2,$$

or equivalently,

$$3b^2 - (2h)^2 = 3(2b+2h)^2 - (3b+4h)^2.$$

Thus, if

$$3b^2 - (2h)^2 = 12,$$

Math. Mag. **93** (2020) 378–379. doi:10.1080/0025570X.2020.1817708
MSC: 51M04, 97C40

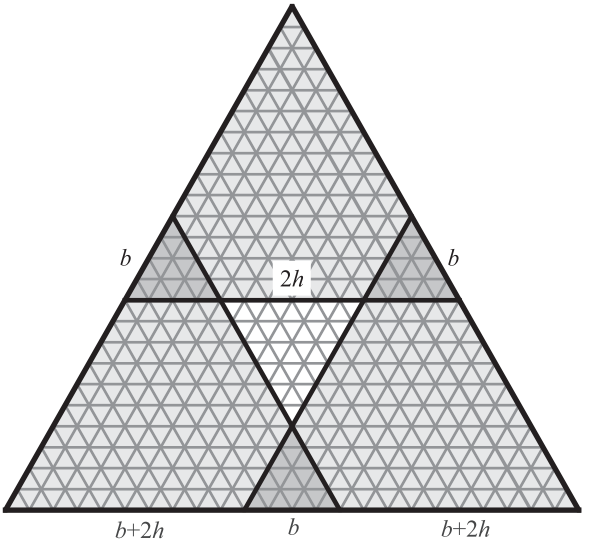


Figure 1 Dividing up a large equilateral triangle, of side length $3b + 4h$.

then

$$3(2b + 2h)^2 - (3b + 4h)^2 = 12.$$

Now let b denote the even side, and h the altitude to the even side, of an almost equilateral triangle satisfying

$$3b^2 - (2h)^2 = 12.$$

The above recursion then produces a larger almost equilateral Heronian triangle with even side $2b + 2h$ and altitude to the even side $(3b + 4h)/2$.

When $(b, h) = (4, 3)$, as in the $(3, 4, 5)$ right triangle, the recursion yields the following infinite sequence of almost equilateral Heronian triangles:

$$(3, 4, 5) \to (13, 14, 15) \to (51, 52, 53)$$
$$\to (193, 194, 195) \to (723, 724, 725) \ldots .$$

■

Acknowledgments The author wishes to thank an anonymous referee for helpful suggestions on an earlier draft of this note.

REFERENCE

[1] Beauregard, R. A., Suryanarayan, E. R. (1998). The Brahmagupta triangles. *College Math. J.* 29(1): 13–17. doi.org/10.1080/07468342.1998.11973907

Summary. We use Heron's formula and the inclusion–exclusion principle to show that infinitely many almost equilateral Heronian triangles exist.

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

When Two Wrongs Make a Right

LEONARD VAN WYK
James Madison University
Harrisonburg, VA 22807
vanwykla@jmu.edu

Those who have taught calculus know there is no shortage of abused differentiation rules. For example: treating 2 like x and concluding the derivative of the constant e^2 is e^2. Or, for the more severely confused, treating e like x and concluding the derivative of e^2 is $2e$.

But perhaps the mother of all misunderstood functions is $f(x) = x^x$, since there are two well-known differentiation rules just begging to be improperly applied: (1) the rule for functions with a constant exponent, such as $y = x^2$, resulting in a derivative of $x \cdot x^{x-1}$, and (2) the differentiation rule for functions with a constant base, such as $y = 2^x$, resulting in a derivative of $(\ln x)x^x$. Each of these abuses is fairly common. However, a colleague observed that the *sum* of these two incorrect solutions is actually the correct solution, $f'(x) = x^x + x^x \ln x$. One way to understand why this is not just a happy coincidence involves introducing a third variable and using partial differentiation.

Specifically, let $z = x^y$. Take the total differential:

$$\begin{aligned} dz &= \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy \\ &= yx^{y-1}dx + (\ln x)x^y dy. \end{aligned}$$

Setting $x = y$ yields

$$dz = x \cdot x^{x-1}dx + (\ln x)x^x dx,$$

or, equivalently,

$$\frac{dz}{dx} = x^x + (\ln x)x^x.$$

In fact, the same holds for all functions of the form $f(x) = g(x)^{h(x)}$, where g and h are differentiable (assuming the confused differentiator knows the chain rule). Assuming the exponent, $h(x)$, is constant yields the incorrect derivative of $h(x)g(x)^{h(x)-1}g'(x)$. Assuming the base, $g(x)$, is constant yields the incorrect derivative of $(\ln g(x))g(x)^{h(x)}h'(x)$. The sum of these two incorrect solutions is again the correct solution, and for the same reason as above.

So, apparently, it is possible for two wrongs to make a right after all.

Summary. We consider an interesting observation regarding derivatives of functions like $f(x) = x^x$. Specifically, that two mistaken approaches to taking the derivative, when added, lead to the correct result.

LEONARD VAN WYK (MR Author ID: 352132) is a professor of mathematics at James Madison University. His mathematical interests include geometric group theory, knot theory, and unexpected connections.

Math. Mag. **93** (2020) 380. doi:10.1080/0025570X.2020.1818526
MSC: 00-01

A Probabilistic Derivation of the Formula for the Sum of a Geometric Series

JEREMY F. ALM
Lamar University
Beaumont, TX 10009
jalm@lamar.edu

The usual derivation for the sum of a geometric series involves taking limits of both sides of the algebraic identity

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

In this note, we provide a derivation that uses almost no algebra and avoids explicit limits.

Suppose Alice and Bob play a rock-paper-scissors-like game in which, in any given round, Alice can win, Bob can win, or they can draw. The probability that Alice wins any given round is p, the probability that Bob wins is q, and the probability of a draw is r, where $p + q + r = 1$. Any two distinct rounds are independent. The first player to win a round wins the game. What is the probability that Alice wins the game?

On the one hand, there are infinitely many ways for Alice to win, as there can be any nonnegative integer number of draws before Alice wins a round. The probability that Alice wins the game in the first round is p, the probability that she wins in the second round is pr, the probability that she wins in the third round is pr^2, and so on. Since these events are mutually exclusive, we may sum them to get

$$p + pr + pr^2 + pr^3 + \cdots = \sum_{k=0}^{\infty} p \cdot r^k$$

On the other hand, we may compute directly the probability that Alice wins by observing that the number of draws before Alice wins a round is irrelevant! So suppose someone wins in the kth round. Then we know that in *this* round, there was no draw, so the probability that it was Alice who won is $p/(p + q)$. Hence,

$$\sum_{k=0}^{\infty} p \cdot r^k = \frac{p}{p+q} = \frac{p}{1-r},$$

as desired. This derivation is easily modified for the case of a truncated series, which is left as an exercise for the reader.

Summary. We derive the formula for the sum of a geometric series using elementary probability theory. Limits and algebra are avoided.

JEREMY F. ALM (MR Author ID 707759) is Professor and Chair of Mathematics and Director of the Quality Enhancement Plan at Lamar University. He occasionally gets to do mathematics. In order to remain sane, he pretends to be a musician; he plays electric guitar, harmonica, and ukulele rather poorly but with great enthusiasm. He hails from southern Wisconsin, but loves the seafood in his new home on the Gulf Coast.

Math. Mag. **93** (2020) 381. doi:10.1080/0025570X.2020.1818525
MSC: 00A08

Carl B. Allendoerfer Awards

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of articles of expository excellence published in MATHEMATICS MAGAZINE. The Award is named for Carl B. Allendoerfer, a distinguished mathematician at the University of Washington and president of the Mathematical Association of America, 1959–60.

Beth Malmskog and Kathryn Haymaker

"What (quilting) circles can be squared?" MATHEMATICS MAGAZINE, Volume 92, Number 3, June 2019, pages 173–186.

For evidence that mathematics—and the origin of mathematical journeys—can arise in unexpected places, one needs to look no farther than this article. The story begins with a question from a friend about a quilting circle: how can one arrange five rounds of quilt hand-offs among a group of five quilters so that every quilter hands off once to every other quilter? That way each quilter gets to meet every other one, rather than, say, always passing to the same person. The friend is not quite able to get this to work, and so appeals to Prof. Malmskog. As the authors write, "who could resist"?

It turns out the quilting hand-offs can be represented in the form of a Latin square. We assign each quilter a number, using the numbers $0, 1, \ldots, n-1$ for n quilters. The quilters are arranged into the left column of the Latin square. The quilters they hand-off-to appear in subsequent columns. A sample Latin square for four quilters is shown in Figure 1. Notice that the first row indicates quilter 0 hands off to quilter 1, who hands off to quilter 3, who hands off to quilter 2.

0	1	3	2
1	2	0	3
2	3	1	0
3	0	2	1

Figure 1 A Latin square for four quilters.

The authors' goal is to find not just any Latin square, but one in which all the hand-offs are different, so that each quilter hands off to each other quilter. The hand-offs from one quilter to another correspond to the sequences (i, j) in each row of the Latin

Math. Mag. **93** (2020) 382–385. doi:10.1080/0025570X.2020.1819057
MSC: Primary 01A70

square. For instance, in the first row of Figure 1, the sequences are (0, 1), (1, 3), and (3, 2). To ensure that every quilter hands off once to every other quilter, we require that each pair (i, j) appear one and only one time when reading across rows. Latin squares with this additional property are called *row-complete*, and the Latin square in Figure 1 is an example. In addition to their desirability for quilting circles, row-complete Latin squares play a role in the design of experiments in which treatments might have residual effects, such as taste-testing experiments.

Surprisingly, no row-complete Latin squares of order 5 exist, as the authors show using a case analysis. So there was a good reason the first author's friend couldn't get things to work! The question of determining whether there is a row-complete Latin square of given size n is a deep one, which remains unresolved in full generality. Constructions for even n date from the 1940s, and constructions for odd composite n date from the 1990s. The case of prime n, aside from $n = 3, 5, 7$, remains mysterious.

You may notice that in Figure 1, the sequence of successive differences of elements in each row, considered modulo 4, is always 1, 2, 3. Such row-complete Latin squares are called *rotational*, and it's possible to find them whenever n is even. Interestingly, the existence of such Latin squares is equivalent to the cyclic group of order n having a group-theoretic property called sequenceability, and a theorem from the 1960s shows that this holds only for even n.

The article closes by introducing "quiltdoku" problems: can a partially filled grid be completed to become a row-complete Latin square? A new fabric-themed puzzling craze could be upon us.

The authors let us accompany them on their journey with clarity, gentleness, and efficiency. By the end, we've visited graph theory, group theory, combinatorics, and even experimental design—a pleasing patchwork of mathematical ideas.

Response from the Authors

We are thrilled and honored to receive the Carl B. Allendoerfer Award! From the question that sparked our inquiry through our continuing investigations today, the work chronicled in this paper has been about building connection and community through mathematics. Even the original quilting question came to us through friendship and the surprising reach of a mathematics puzzle segment on a community radio show in Colorado—this great question was one of the perks of being known as the radio math lady within a tiny world. This award is especially wonderful to us in this context, because it lets us know that this we have connected with the larger world in a meaningful way. Thank you to the awards committee and to the MAA!

Beth Malmskog wishes more of her friends would turn their hobbies into math problems for her to work on. Dr. Malmskog received her Ph.D. from Colorado State University in 2011. Her research is in number theory and discrete mathematics, including computational questions and applications. She is now an assistant professor in the Department of Mathematics and Computer Science at Colorado College.

Katie Haymaker obtained a Ph.D. in mathematics from the University of Nebraska-Lincoln in 2014. She is an associate professor in the Department of Mathematics and Statistics at Villanova University. Dr. Haymaker's research focuses on coding theory and applied discrete mathematics. She is grateful that Dr. Malmskog included her in discussions of this fascinating problem, which has led to many exciting discoveries and re-discoveries.

J. Arias De Reyna, David Clark, and Noam D. Elkies

"A modern solution to the Gion Shrine problem." Mathematics Magazine Volume 92, Number 2, April 2019, pages 110–122.

This article shines a welcome light on a little-known nook of mathematical history: remarkable computationally-intensive geometry problems inscribed in wooden tablets and hung in 18th-century Buddhist temples and Shinto shrines. These formed a part of a cultural flourishing in Japan under the Tokugawa shogunate, in which kabuki theater, haiku poetry, ukiyo-e woodblock printing, and a unique style of mathematics called *wasan* enjoyed a surge of interest. *Wasan* focused on geometric problems, such as packings with circles and polygons. An especially appealing set of theorems would be inscribed on a tablet and hung in a shrine, "as an offering to the gods, a challenge to other worshippers, and an advertisement for the school producing the work," in the words of the authors. Being a closed society, the Japanese drew inspiration from Greek geometry and Chinese computation, and *wasan* problems often blended the two.

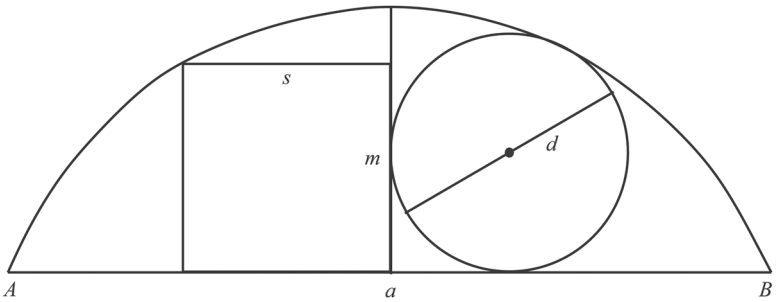


Figure 2 The Gion shrine problem.

One particularly famous tablet was found in Kyoto's Gion shrine. It consists of a square and circle inscribed in a circular arc as in Figure 2. The square has side length s, the circle has diameter d, and two other segments relating to the arc have lengths a and m. The challenge is to recover a, d, s, and m given only the quantities

$$p = a + d + s + m$$

and

$$q = (m/a) + (d/m) + (s/d),$$

using the geometrical relationships of the quantities involved.

The first recorded solution was due to Tsuda Nobuhisa in 1749, who gave the desired quantities in terms of the roots of a degree-1024 polynomial. Further 18th-century solutions reduced the degree of the necessary polynomial to 46, and then to 10. The latter uses only the Pythagorean theorem, linear algebra, and "a great deal of algebraic persistence," as the authors put it. The authors then give a modern solution that relies on trigonometric functions to describe a, d, s, and m in terms of a simpler degree-10 polynomial, with the added benefits of elucidating the existence and uniqueness of solutions.

The article then turns to a problem that would not have been on the minds of wasan practitioners, but would interest anyone who cares about Diophantine equations: is

there a solution to the Gion shine problem where a, d, m, and s are all rational numbers? This leads to the problem of finding all solutions to $y^2 = x^3x$ with both x and y rational. Going back at least to Fibonacci, this problem was solved by Fermat using his famous method of descent, which has reverberated through modern number theory.

The article concludes with remarks about some more recent developments in the study of Diophantine equations. Thus a voyage that started in centuries-ago Japan winds up near the present day. The article gives a vivid portrait of how widely separated parts of mathematical history can in fact be intertwined, and reminds us of the universality of mathematics in the human experience.

Response from the Authors

It is a pleasure and a surprise to be awarded the Allendoerfer prize this year. In addition to being an interesting and challenging exercise, the Gion Shrine problem gave us an opportunity to portray mathematics at the intersection of art, recreation, science, history, and culture. When we arrived at the solution, one of us (Arias de Reyna, who lived under a dictator in Spain and did not have access to math books) noticed how he would have enjoyed reading it as a teenager. Our paper is written, in large part, for this young person. We hope that our solution weaving together geometry, trigonometry, and algebra might be read and enjoyed by someone with a similar passion for mathematics. The last part of the paper on Diophantine equations was a relatively late but very welcome addition, revealing the Gion shrine problem to be another point of intersection, this time between traditional Japanese math and Western number theory.

J. Arias De Reyna learned mathematics from books starting at age thirteen; at the time, even books were difficult to get in dictator Franco's Spain. He has published a book about Carleson's proof on the convergence of Fourier series, defining the largest known rearrangement invariant space of functions with almost everywhere convergent series. He also obtained good bounds for the Riemann–Siegel expansion.

David Clark was trained as a quantum topologist, but has recently become interested in the history of Japanese mathematics. In 2017 he hosted an international conference on the topic in Ashland, Virginia. Clark regularly takes students to Japan to learn about *sangaku* tablets, and has written about his experiences in *Math Horizons*.

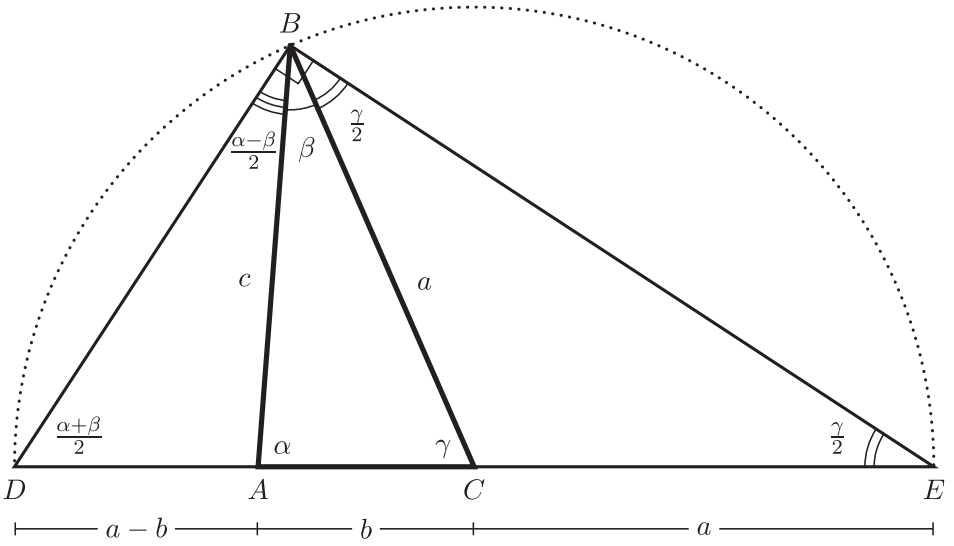
Noam D. Elkies is a number theorist, much of whose work concerns Diophantine geometry and computational number theory. He was granted tenure at Harvard at age 26, the youngest in the University's history. Outside of math, Elkies' main interests are music—mainly classical piano and composition—and chess, where he specializes in composing and solving problems.

The Mollweide Equations from the Law of Sines

REX H. WU
New York Presbyterian
Lower Manhattan Hospital
New York, NY 10038
rexhwu@yahoo.com

Mollweide's equations: In $\triangle ABC$, we have that

$$\frac{\sin\left(\frac{\alpha-\beta}{2}\right)}{\cos\left(\frac{\gamma}{2}\right)} = \frac{a-b}{c} \qquad \text{and} \qquad \frac{\cos\left(\frac{\alpha-\beta}{2}\right)}{\sin\left(\frac{\gamma}{2}\right)} = \frac{a+b}{c}.$$



Apply the law of sines to $\triangle ABD$ and $\triangle ABE$. Note that $\frac{\alpha+\beta}{2}$ and $\frac{\gamma}{2}$ are complementary, as are $\frac{\alpha-\beta}{2}$ and $\beta+\frac{\gamma}{2}$.

For another proof of the first equation, see DeKleine [1].

Summary. We provide a visual proof for the two Mollweide equations in one figure.

REX H. WU (MR Author ID: 1293646, ORCID 0000-0003-0970-3741) is interested in the visualization of mathematical identities and concepts. He wishes to thank the anonymous reviewers and the editor for their suggestions.

REFERENCE

[1] DeKleine, H. A. (1988). Proof Without Words: Mollweide's Equation. *Math. Mag.* 61(5): 281. DOI: 10.1080/0025570X.1988.11977390

Math. Mag. **93** (2020) 386. doi:10.1080/0025570X.2020.1817707
MSC: 97G60

Who Outlives Whom?

JAMES E. CIECKA
DePaul University
Chicago, IL 60604
jciecka@depaul.edu

A and B are life partners. A is exactly age x and B is exactly age y. Let ω be the youngest age at which everyone in the general population, regardless of gender, is dead (about age 120). The United States life table for A's gender tells us there are $a_x, a_{x+1}, \ldots, a_\omega$ people still living (from a radix of 100,000 newborns) at exact ages $x, x+1, \ldots, \omega$, with $a_\omega = a_{\omega+1} = \cdots = 0$ [1]. Similarly, the life table for B's gender gives us $b_y, b_{y+1}, \ldots, b_\omega$, living at exact ages $y, y+1, \ldots, \omega$, with $b_\omega = b_{\omega+1} = \cdots = 0$. What are the probabilities that A outlives B, that B outlives A, and that they both die in the same year?

The probability that A outlives B (i.e., that A dies after B) is given by

$$\frac{1}{a_x b_y} \sum_{t=0}^{\omega-x-1} a_{x+t+1} \left(b_{y+t} - b_{y+t+1}\right),$$

the probability that B outlives A is given by

$$\frac{1}{a_x b_y} \sum_{t=0}^{\omega-x-1} b_{y+t+1} \left(a_{x+t} - a_{x+t+1}\right),$$

and the probability that A and B die in the same year is

$$\frac{1}{a_x b_y} \sum_{t=0}^{\omega-x-1} \left(a_{x+t} - a_{x+t+1}\right)\left(b_{y+t} - b_{y+t+1}\right).$$

These three probabilities are represented visually by areas in Figure 1, and this enables us to prove the following proposition:

Proposition 1. *The three probabilities partition the rectangle in Figure 1. That is, none of the underlying death and survival events overlap, and the probabilities of these events sum to* 1.

Proof. Lightly shaded area divided by the area of the entire rectangle is the probability that A outlives B. Darkly shaded area divided by the area of the entire rectangle is the probability that B outlives A, and unshaded area divided by the area of the entire rectangle is the probability that both die in the same year. No areas overlap, and their sum divided by the area of the entire rectangle is 1. ■

Historical Remark. Using rectangular areas to represent life probabilities dates to Edmond Halley's work in 1693 [2]. In his seminal paper on life annuities, he drew a rectangle with sides that represented the number of people in a life table at two different ages. Halley performed multiplications that he thought could be "better explained by expounding these products by rectangular parallelograms." He considered two people of different ages and determined the probability that at least one of the two would survive in future years by using the ratios of areas. This formed the basis of the very first valuation of a joint life annuity (such as a pension) that paid as long as at least one person survived.

Math. Mag. **93** (2020) 387–388. doi:10.1080/0025570X.2020.1818530
MSC: Primary 60-01; 60-03

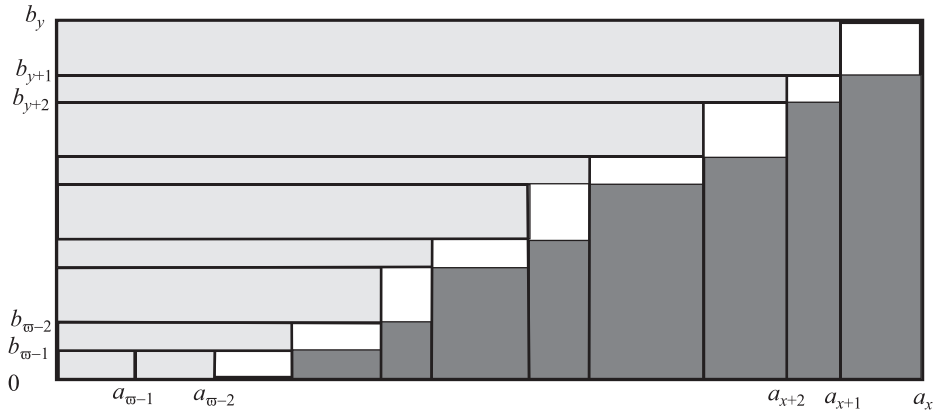


Figure 1 A visual proof of Proposition 1.

Notation Remark. The minimum upper index of summation that can be used in computing probabilities is $\omega - \max(x, y) - 1$. Using $\omega - x - 1$ results in correct probability calculations, but also produces extra terms, all of which are zero, when $x < y$. From Figure 1, B must be two years older than A because B reaches age ω when A is age $\omega - 2$. Therefore, $y = x + 2$, and the minimum upper index of summation becomes $\omega - y - 1 = \omega - (x + 2) - 1 = \omega - x - 3$.

Acknowledgments The author wishes to thank an anonymous referee who made several comments which led to a much improved note.

REFERENCES

[1] Arias, E., Heron, M., Xu, J. (2017). *National Vital Statistics Reports, United States Life Tables 2013*. Tables 2 and 3, U.S. Department of Health and Human Services, Center for Disease Control and Prevention. 66(3): 11–14.

[2] Halley, E. (1693). An estimate of the degrees of mortality of mankind, drawn from the curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. *Philos. Trans.* 17: 596–656.

Summary. What are the probabilities that A outlives B, that B outlives A, and that they both die in the same year? We provide a visual proof to show that the three probabilities sum to 1. Proving this proposition algebraically is cumbersome and not insightful. The visual proof is much more straightforward and clear.

JAMES E. CIECKA (MR Author ID 1239016) is a professor of economics at DePaul University and past chairman of the department. He is the co-executive editor of the *Journal of Forensic Economics*, which specializes in research applying economic theory and econometrics to litigation-related matters. He received his Ph.D. in economics from Purdue University.

PROBLEMS

LES REID, *Editor*
Missouri State University

EUGEN J. IONAŞCU, *Proposals Editor*
Columbus State University

RICHARD BELSHOFF, Missouri State University; *MAHYA GHANDEHARI*, University of Delaware; *EYVINDUR ARI PALSSON*, Virginia Tech; *GAIL RATCLIFF*, Eastern Carolina University; *ROGELIO VALDEZ*, Centro de Investigación en Ciencias, UAEM, Mexico; *Assistant Editors*

Proposals

To be considered for publication, solutions should be received by May 1, 2021.

2106. *Proposed by Timothy Hall, PQI Consulting, Cambridge, MA.*

For a fixed positive integer $n > 1$, characterize those integers k such that the Maclaurin series for

$$f_{n,k}(x) = \sqrt[n]{1 + kx}$$

(as a function of x) has integer coefficients.

2107. *Proposed by Seán M. Stewart, Bomaderry, Australia.*

Evaluate

$$\int_0^\infty \frac{\log\left(1 + x^6\right)}{1 + x^2}\, dx.$$

2108. *Proposed by Souvik Dey (graduate student), University of Kansas, Lawrence, KS.*

Let G be a torsion-free abelian group such that every proper subgroup of G is cyclic. Show that G is cyclic.

2109. *Proposed by George Stoica, Saint John, NB, Canada.*

Let S^1 denote the multiplicative group of complex numbers of absolute value 1. Characterize all continuous maps $f : S^1 \to S^1$ such that

$$f(z^m) = f(z)^m$$

for all $z \in S^1$.

Math. Mag. **93** (2020) 389–398. doi:10.1080/0025570X.2020.1825277.

We invite readers to submit original problems appealing to students and teachers of advanced undergraduate mathematics. Proposals must always be accompanied by a solution and any relevant bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. Submitted problems should not be under consideration for publication elsewhere.

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

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

2110. *Proposed by Rob Downes, Newark Academy, Livingston, NJ.*

Let $\triangle ABC$ be an isosceles triangle with $AB = AC = a$ and $BC = b$. Let X be a point in the interior of $\triangle ABC$ such that the inradii of $\triangle AXB$, $\triangle BXC$, and $\triangle CXA$ are all equal. Express that common radius r in terms of a and b.

Quickies

1105. *Proposed by Hideyuki Ohtsuka, Saitama, Japan.*

The Tribonacci numbers are defined recursively by $T_0 = 0$, $T_1 = 1$, $T_2 = 1$ and

$$T_n = T_{n-1} + T_{n-2} + T_{n-3}$$

for $n \geq 3$. Prove that

$$2\left(T_n^4 + T_{n+1}^4 + T_{n+2}^4 + T_{n+3}^4\right) = \left(T_n^2 + T_{n+1}^2 + T_{n+2}^2 + T_{n+3}^2\right)^2 + 8T_nT_{n+1}T_{n+2}T_{n+3}.$$

This is the analog of Candido's identity for the Fibonacci numbers:

$$2\left(F_n^4 + F_{n+1}^4 + F_{n+2}^4\right) = \left(F_n^2 + F_{n+1}^2 + F_{n+2}^2\right)^2.$$

1106. *Proposed by Rafael Jakimczuk, Universidad Nacional de Luján, Buenos Aires, Argentina.*

Let $\sigma(n)$ denote the sum of the positive divisors of n for $n \geq 1$. Prove that

$$n < \sigma(n) < n \log n + n.$$

Solutions

Associate complex numbers **December 2019**

2081. *Proposed by Ioan Băetu, Botoşani, Romania.*

Finitely many distinct complex numbers $z_1, z_2, \ldots, z_n$ are called *associate over* $\mathbb{Q}$ if they are the n roots of a degree-n irreducible polynomial in $\mathbb{Q}[X]$. Assume that $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ are collections of complex numbers that are associate over $\mathbb{Q}$. If $y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n$ are rational, prove that

$$y_1 - x_1 = y_2 - x_2 = \cdots = y_n - x_n.$$

Solution by Paul Budney, Sunderland, MA.
Denote the minimal polynomials (over $\mathbb{Q}$) of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ by

$$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0 \text{ and } g(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_0,$$

respectively. Let $q_i = x_i - y_i \in \mathbb{Q}$. Then y_i is a root of $f(z + q_i)$. Now $f(z)$ has rational coefficients and is monic, irreducible over $\mathbb{Q}$, and of degree n. Since $q_i \in \mathbb{Q}$, $f(z + q_i)$ has the same properties as $f(z)$. Therefore $f(z + q_i)$ must be the minimal polynomial for y_i over $\mathbb{Q}$. Since minimal polynomials are unique, $f(z + q_i) = g(z)$. Equating the coefficients of z^{n-1}, we have

$$nq_i + a_{n-1} = b_{n-1} \text{ or } q_i = \frac{1}{n}(b_{n-1} - a_{n-1}).$$

Since the expression for q_i does not involve i, $q_i = x_i - y_i$ is a constant.

Also solved by Anthony Bevelacqua, Elton Bojaxhiu (Germany) & Enkel Hysnelaj (Australia), Robert Calcaterra, Eugene Herman, Dmitri Fleischman, Elias Lampakis (Greece), John H. Smith, Edward White & Roberta White, and the proposer.

Maximum and minimum eigenvalues of a matrix — December 2019

2082. *Proposed by Yves Nievergelt, Department of Mathematics, Eastern Washington University, Cheney, WA.*

Let

$$L = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ z_1 & \dots & z_n & 1 \end{pmatrix}$$

be any $(n + 1) \times (n + 1)$ real matrix differing from the identity matrix only on the nondiagonal entries $\mathbf{z} = (z_1, \dots, z_n)$ of its last row. Let L^T be the transpose of L. Find the largest and smallest eigenvalues of the matrices $L^T L$ and LL^T in terms of $\mathbf{z}$.

Solution by Michael Vowe, Therwil, Switzerland.
It is well known that AB and BA have the same characteristic polynomial and hence the same eigenvalues. Therefore it suffices to prove the result for LL^T. Let $Z = \sum_{k=1}^{n} z_k^2$. Then

$$LL^T - \lambda I = \begin{pmatrix} 1-\lambda & 0 & \cdots & 0 & z_1 \\ 0 & 1-\lambda & & 0 & z_2 \\ \vdots & & \ddots & \vdots & \\ 0 & 0 & & 1-\lambda & z_n \\ z_1 & z_2 & \cdots & z_n & 1+Z-\lambda \end{pmatrix}.$$

Performing the row operations $R_{n+1} - z_i R_i \to R_{n+1}, i = 1, \dots, n$, we obtain the matrix

$$M(z_1, \dots, z_n) = \begin{pmatrix} 1-\lambda & 0 & \cdots & 0 & z_1 \\ 0 & 1-\lambda & & 0 & z_2 \\ \vdots & & \ddots & \vdots & \\ 0 & 0 & & 1-\lambda & z_n \\ \lambda z_1 & \lambda z_2 & \cdots & \lambda z_n & 1-\lambda \end{pmatrix}.$$

We claim that the characteristic polynomial for LL^T is

$$\det(M(z_1, \ldots, z_n)) = (1-\lambda)^{n-1}\left(\lambda^2 - (2+Z)\lambda + 1\right), n \geq 1.$$

We will proceed by induction on n. The case $n = 1$ is immediate.

Expanding along the first row of $M(z_1, \ldots, z_n)$, we find that

$$\begin{aligned}
\det(M(z_1, \ldots, z_n)) &= (1-\lambda)\det(M(z_2, \ldots, z_n)) \\
&\quad + (-1)^n z_1 \det\begin{pmatrix} 0 & 1-\lambda & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 1-\lambda \\ \lambda z_1 & \lambda z_2 & \cdots & \lambda z_{n-1} & \lambda z_n \end{pmatrix} \\
&= (1-\lambda)\det(M(z_2, \ldots, z_n)) \\
&\quad + (-1)^n z_1\left((-1)^{n-1}\lambda z_1 (1-\lambda)^{n-1}\right) \\
&\quad \text{(expanding along the first column of the second matrix)} \\
&= (1-\lambda)(1-\lambda)^{n-2}\left(\lambda^2 - \left(2 + \sum_{k=2}^{n} z_k^2\right)\lambda + 1\right) \\
&\quad - \lambda z_1^2 (1-\lambda)^{n-1} \quad \text{(by induction)} \\
&= (1-\lambda)^{n-1}\left(\lambda^2 - \left(2 + \sum_{k=1}^{n} z_k^2\right) + 1\right)
\end{aligned}$$

and the proof of the claim is complete.

Therefore the $n+1$ eigenvalues are $\lambda_1, \ldots, \lambda_{n-1} = 1$ and

$$\lambda_n = \frac{1}{2}\left(2 + Z - \sqrt{Z^2+4Z}\right) \text{ and } \lambda_{n+1} = \frac{1}{2}\left(2 + Z + \sqrt{Z^2+4Z}\right).$$

Since $Z > 0$ and $Z < \sqrt{Z^2+4Z}$, $\lambda_n < 1$ is the smallest eigenvalue and $\lambda_{n+1} > 1$ is the largest.

Also solved by Robert A. Agnew, Hafez Al-Assad (Syria). Michel Bataille (France) Elton Bojaxhiu (Germany) & Enkel Hysnelaj (Australia), Brian Bradie, Dmitry Fleischman, Eugene A. Herman, John C. Kieffer, Elisa Lampakis (Greece), Albert Natian, Northwestern University Problem Solving Group, Albert Stadler (Switzerland), and the proposer.

A series with recursively defined terms **December 2019**

2083. *Proposed by Paul Bracken, University of Texas Rio Grande Valley, Edinburg, TX.*

Let a_1 be a positive real number. Define the sequence $\{a_n\}$ recursively by $a_{n+1} = n^2/a_n$ for $n = 1, 2, \ldots$. Evaluate

$$\lim_{n\to\infty} \frac{1}{\ln n} \sum_{k=1}^{n} \frac{1}{a_k}.$$

Solution by Michael Goldenberg, The Ingenuity Project, Baltimore Polytechnic Institute, Baltimore, MD, and Mark Kaplan, Towson University, Baltimore, MD.

Evaluating the first few terms of the sequence, we find that

$$\begin{aligned}
\frac{1}{a_2} &= \frac{a_1}{1^2},\\
\frac{1}{a_3} &= \frac{1^2}{2^2 a_1},\\
\frac{1}{a_4} &= \frac{2^2 a_1}{3^2 \cdot 1^2},\\
\frac{1}{a_5} &= \frac{3^2 \cdot 1^2}{4^2 \cdot 2^2 a_1},\\
\frac{1}{a_6} &= \frac{4^2 \cdot 2^2 a_1}{5^2 \cdot 3^2 \cdot 1^2}.
\end{aligned}$$

From this information, we conjecture that

$$\frac{1}{a_{2k}} = \frac{((2k-2)!!)^2 a_1}{((2k-1)!!)^2} \quad \text{and}$$
$$\frac{1}{a_{2k+1}} = \frac{((2k-1)!!)^2}{((2k)!!)^2 a_1},$$

which is readily verified by induction.

Stirling's formula states that

$$n! = \sqrt{2\pi}\, n^{n+1/2} e^{-n} \left(1 + O\left(\frac{1}{n}\right)\right).$$

We have

$$\begin{aligned}
(2k)!! &= 2^k k!\\
&= \sqrt{\pi}\, 2^{k+1/2} k^{k+1/2} e^{-k} \left(1 + O\left(\frac{1}{k}\right)\right)
\end{aligned}$$

and

$$\begin{aligned}
(2k-1)!! &= \frac{(2k)!}{2^k k!}\\
&= \frac{\sqrt{2\pi}(2k)^{2k+1/2} e^{-2k} \left(1 + O\left(\frac{1}{k}\right)\right)}{2^k \sqrt{2\pi}\, k^{k+1/2} e^{-k} \left(1 + O\left(\frac{1}{k}\right)\right)}\\
&= 2^{k+1/2} k^k e^{-k} \left(1 + O\left(\frac{1}{k}\right)\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\frac{1}{a_{2k}} &= \frac{((2k-2)!!)^2 a_1}{((2k-1)!!)^2}\\
&= \frac{\pi\, 2^{2k-1} k^{2k-1} e^{-2k} \left(1 + O(1/k)\right)}{2^{2k+1} k^{2k} e^{-2k} \left(1 + O(1/k)\right)} a_1\\
&= \frac{\pi a_1}{4k} + O\left(\frac{1}{k^2}\right)
\end{aligned}$$

and

$$\begin{aligned}\frac{1}{a_{2k+1}} &= \frac{((2k-1)!!)^2}{((2k)!!)^2 a_1} \\ &= \frac{2^{2k+1}k^{2k}e^{-2k}\,(1+O(1/k))}{\pi\ 2^{2k+1}k^{2k+1}e^{-2k}a_1\,(1+O(1/k))} \\ &= \frac{1}{\pi k a_1} + O\left(\frac{1}{k^2}\right).\end{aligned}$$

Since $\lim_{n\to\infty} 1/a_n = 0$, the limit in question is the same whether n is restricted to be even or to be odd. Therefore

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{\ln n}\sum_{k=1}^{n}\frac{1}{a_k} &= \lim_{m\to\infty}\frac{1}{\ln(2m+1)}\sum_{k=1}^{2m+1}\frac{1}{a_k} \\ &= \lim_{m\to\infty}\frac{\frac{1}{a_1}+\sum_{k=1}^{m}\frac{1}{k}\left(\frac{\pi a_1}{4}+\frac{1}{\pi a_1}\right)+O(1)}{\ln(2m+1)} \\ &= \lim_{m\to\infty}\frac{\left(\frac{\pi a_1}{4}+\frac{1}{\pi a_1}\right)\sum_{k=1}^{m}\frac{1}{k}+O(1)}{\ln(2m+1)} \\ &= \lim_{m\to\infty}\frac{\left(\frac{\pi a_1}{4}+\frac{1}{\pi a_1}\right)\ln m+O(1)}{\ln(2m+1)} \\ &= \frac{\pi a_1}{4}+\frac{1}{\pi a_1}.\end{aligned}$$

Also solved by Robert A. Agnew, Michel Bataille (France), Brian Bradie, Dmitry Fleischman, Russell Gordon, Eugene A. Herman, Elias Lampakis (Greece), Albert Natian, Chikkanna Selvaraj, Michael Vowe (Switzerland), and the proposer. There were two incomplete or incorrect solutions.

All roads lead to Rome **December 2019**

2084. *Proposed by Andrei Ionescu, Lucretiu Patrascanu High School, Romania.*

A *random n-tournament* is a complete directed graph on n vertices in which the direction of each edge is chosen uniformly and independently at random. A vertex of a random tournament is called a *Rome* if it can be reached from every other vertex: *"all roads lead to Rome."* Let R be the number of Romes, and let $\mathbb{E}_n[R]$ be the expected number of Romes in a random n-tournament. Find

$$\lim_{n\to\infty}\frac{\mathbb{E}_n[R]}{n}.$$

Solution by Edward Schmeichel, San José State University, San José, CA.
We will prove that

$$\lim_{n\to\infty}\frac{\mathbb{E}_n(R)}{n} = 1 \tag{1}$$

via a series of reductions.

Let $p_{n,k}$ denote the probability that a random n-tournament satisfies $R = k$ for $0 \le k \le n$. We have

$$1 \ge \frac{\mathbb{E}_n(R)}{n} = \frac{1}{n}\sum_{k=0}^{n} k\ p_{n,k} \ge \frac{1}{n} n\ p_{n,n} = p_{n,n},$$

so (1) would be implied by

$$\lim_{n\to\infty} p_{n,n} = 1. \tag{2}$$

An n-tournament T with $R < n$ contains vertices x and y such that y is unreachable from x. If $X \subset V(T)$ denotes the set of vertices reachable from x, then every edge between X and $\overline{X} = V(T) - X$ must be directed toward X, otherwise some vertex in $\overline{X}$ would be reachable from x. If $|X| = k$, $1 \le k \le n-1$, there are $\binom{n}{k}$ possible choices for X, each with probability $1/2^{k(n-k)}$ that the $k(n-k)$ edges between X and $\overline{X}$ are all directed toward X. Thus the probability $1 - p_{n,n}$ that a random n-tournament has $R < n$ satisfies the inequality

$$1 - p_{n,n} \le \sum_{k=1}^{n-1} \frac{\binom{n}{k}}{2^{k(n-k)}}.$$

Therefore (2) would be implied by

$$\lim_{n\to\infty} \sum_{k=1}^{n-1} \frac{\binom{n}{k}}{2^{k(n-k)}} = 0. \tag{3}$$

Finally, (3) would be implied by

$$\frac{\binom{n}{k}}{2^{k(n-k)}} \le \frac{n}{2^{n-1}} \quad \text{for } 1 \le k \le n-1, \tag{4}$$

since (4) implies that

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

and $\lim_{n\to\infty} n(n-1)/2^{n-1} = 0$ then gives (3).

It remains to prove (4). First note that

$$(n+1-k)2^k \ge 2n, \quad \text{for } 1 \le k \le n, \tag{5}$$

since the inequality is an equality when $k = 1$ and the left-hand side is a nondecreasing function of k for $1 \le k \le n$.

Clearly (4) holds for $n = 2$. We proceed by induction on n.

$$\begin{aligned}\frac{\binom{n+1}{k}}{2^{k(n+1-k)}} &= \frac{n+1}{(n+1-k)2^k} \cdot \frac{\binom{n}{k}}{2^{k(n-k)}} \\ &\le \frac{n+1}{(n+1-k)2^k} \cdot \frac{n}{2^{n-1}} \quad \text{(by the inductive hypothesis)}\end{aligned}$$

$$\leq \frac{n+1}{2n} \cdot \frac{n}{2^{n-1}} \text{ (by (5))}$$
$$= \frac{n+1}{2^n},$$

proving (4).

Editor's Note. Several solvers noted that the result follows from the fact that every vertex in a strongly connected graph is a Rome and that the fraction of tournaments that are strongly connected goes to 1 as $n \to \infty$.

Also solved by Elton Bojaxhiu (Germany) & Enkel Hysnelaj (Australia), Kent E. Morrison, José H. Nieto (Venezuela), Enrique Treviño, Mark Wildon (UK), and the proposer.

An exponential Diophantine equation **December 2019**

2085. *Proposed by Florin Stanescu, Şerban Cioculescu School, Găeşti, Romania.*

Solve the equation

$$3^x \cdot 4^y + 5^z = 7^w$$

in nonnegative integers w, x, y, z.

Solution by John Christopher, California State University, Sacramento, Sacramento, CA.

The only solution is $(x, y, z, w) = (1, 2, 0, 2)$.

If $w = 0$, 1, or 2, it is straightforward to verify that the given solution is the only solution. Now consider $w \geq 3$.

Case 1: If w is odd, then

$$7^w \equiv 3 \pmod 4, \quad \text{but } 3^x 4^y + 5^z \equiv 1 \pmod 4.$$

Case 2: If w is even and z is odd, then

$$7^w \equiv 1 \pmod 3, \quad \text{but } 3^x 4^y + 5^z \equiv 2 \pmod 3.$$

Case 3: If w and z are both even, let $w = 2r$ and $z = 2s$. Then

$$3^x 4^y = (7^r - 5^s)(7^r + 5^s). \tag{1}$$

Case 3a: If r and s are both odd, then

$$7^r - 5^s \equiv 2 \pmod 3 \text{ and } 7^r - 5^s \equiv 2 \pmod 4,$$

which implies that $7^r - 5^s = 2q$ for some odd integer q that is not a multiple of 3, therefore (1) cannot hold.

Case 3b: If r is even and s is odd, then

$$7^r - 5^s \equiv 2 \pmod 3 \text{ and } 7^r - 5^s \equiv 4 \pmod 8,$$

which implies that $7^r - 5^s = 4q$ for some odd integer q that is not a multiple of 3, therefore (1) cannot hold.

Case 3c: If r and s are both even, then

$$7^r + 5^s \equiv 2 \pmod 3 \text{ and } 7^r + 5^s \equiv 2 \pmod 4,$$

which implies that $7^r + 5^s = 2q$ for some odd integer q that is not a multiple of 3, so again (1) cannot hold.

Case 3d: If r is odd and s is even, then

$$7^r - 5^s \equiv 0 \pmod 3 \text{ and } 7^r - 5^s \equiv 2 \pmod 4$$

and

$$7^r + 5^s \equiv 2 \pmod 3 \text{ and } 7^r + 5^s \equiv 2 \pmod 4.$$

Therefore in order for (1) to hold, we must have

$$7^r - 5^s = 2 \cdot 3^x \tag{2}$$

and

$$7^r + 5^s = 2^{2k+1}. \tag{3}$$

In (3), $k = 0$ forces $r = s = 0$. But r is odd, so we must have $k > 0$. Adding and subtracting (2) and (3) and dividing by 2 yields

$$7^r = 2^{2k} + 3^x \tag{4}$$

and

$$5^s = 2^{2k} - 3^x. \tag{5}$$

Now (4) implies that $3 \equiv 3^x \pmod 4$ so x is odd. Therefore

$$2^{2k} \equiv \pm 1 \pmod 5 \text{ and } 3^x \equiv \pm 2 \pmod 5,$$

so

$$2^{2k} - 3^x \not\equiv 0 \pmod 5.$$

If $s > 0$, the left-hand side of (5) is a multiple of 5, but the right-hand side is not, so this is impossible.

Finally, consider the left-hand side of (3) when $s = 0$. We have

$$7^r + 5^s = 7^r + 1 \equiv 8 \pmod{16},$$

which implies that

$$7^r + 5^s = 8q$$

with q odd. If $r > 1$, then $q > 1$ but in that case $(7^r + 5^s)/8$ is not a power of 2, so (3) cannot hold.

We have eliminated all the cases when $w \geq 3$, and the result follows.

Also solved by Hafez Al-Assad (student) (Syria), Elton Bojaxhiu (Germany) & Enkel Hysnelaj (Australia), Gregory Dresden, Westmont College Problem Solving Group, Albert Stadler (Switzerland), Vasile Teodorovici (Canada), Melanie Tian (student), Edward White & Roberta White, and the proposer. There were two incomplete or incorrect solutions.

Answers

Solutions to the Quickies from page 390.

A1105. Let $a = T_n$, $b = T_{n+1}$, and $c = T_{n+2}$. Then $T_{n+3} = a + b + c$ and the result follows immediately from the readily verified fact that

$$2\left(a^4 + b^4 + c^4 + (a+b+c)^4\right) = \left(a^2 + b^2 + c^2 + (a+b+c)^2\right)^2 + 8abc(a+b+c).$$

Note that the Tribonacci identity holds regardless of the initial conditions.

A1106. Suppose $d_1, d_2, \ldots, d_k$ are the positive divisors of n. Then

$$\begin{aligned}\sigma(n) &= d_1 + d_2 + \cdots + d_k \\ &\geq 1 + n \\ &> n.\end{aligned}$$

Note that since $1/x$ is a decreasing function for $x > 0$,

$$\int_i^{i+1} \frac{dx}{x} > \int_i^{i+1} \frac{dx}{i+1} = \frac{1}{i+1},$$

so

$$\int_1^n \frac{dx}{x} > \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

Therefore

$$\begin{aligned}\sigma(n) &= d_1 + d_2 + \cdots + d_k \\ &= n\left(\frac{1}{d_1} + \frac{1}{d_2} + \cdots \frac{1}{d_k}\right) \\ &\leq n\left(\frac{1}{1} + \frac{1}{2} + \cdots \frac{1}{n}\right) \\ &< n\left(1 + \int_1^n \frac{dx}{x}\right) \\ &= n\,(1 + \log n) \\ &= n\log n + n.\end{aligned}$$

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles, books, and other materials are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of mathematics literature. Readers are invited to suggest items for review to the editors.

Ayala-Rincón, Mauricio, and Gabriel Ferreira Silva, Why we need structured proofs in mathematics, cicm-conference.org/2020/NFM/paper_4_Ayala_Silva.pdf.

Lamport, Leslie, How to write a 21st century proof, *Journal of Fixed Point Theory and Applications* 11 (2012) 43–63.

_____, How to write a proof, in *Global Analysis in Modern Mathematics*, edited by Karen K. Uhlenbeck, 311–321; Publish or Perish, 1993. Reprinted in *American Mathematical Monthly* 102 (1995) (7) 600–608, citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.439.7603&rep=rep1&type=pdf.

Once upon a time, high school geometry in the U.S. featured proofs written in two columns, statement and reason. Despite the value of that format for fostering reasoning, the format and proofs themselves in geometry went away, because proofs are "too hard" for today's U.S. students. Structured proofs, a term originated by Lamport, are the analogue for the rest of mathematics. He and other authors tout them as a way of reducing the opportunity to arrive at a false result. They cite an erroneous proof early in John Kelley's famous *General Topology*; Ayala-Rincón and Ferreira Silva point out the error and provide a correct structured proof. Lamport (the author of LaTeX) provides simpler examples by dissecting a proof of the calculus corollary $f'(x) > 0$ on $I \Rightarrow f$ is increasing on I, and, in his earlier paper, that the square root of 2 is irrational. Lamport provides a LaTeX package for conveying a structured proof in textual format. Today, hypertext provides a natural environment for the hierarchical layers of the tree of a structured proof. Structured proof is an idea too good for the profession to continue to ignore.

Thompson, Abigail, A word from. . . , *Notices of the AMS* 66 (11) (December 2019) 1778–1779.

Letters to the editor, *Notices of the AMS* 67 (4) (April 2020) 463–466. ams.org/journals/notices/202004/rnoti-p463.pdf.

Responses to "A word from … Abigail Thompson," *Notices of the AMS* (online only) ams.org/journals/notices/202001/rnoti-o1.pdf.

Thompson, Abigail, Does diversity trump ability? An example of the misuse of mathematics in the social sciences, *Notices of the AMS* 61 (9) (October 2014) 1024–1030.

Letters to the editor, *Notices of the AMS* 62 (1) (January 2015) 9–10. ams.org/journals/notices/201501/rnoti-p9.pdf.

Abigail Thompson, chair of mathematics at UC-Davis and a vice-president of the AMS, objects to a policy at the University of California mandating that faculty applying for positions submit a "contributions to diversity" statement, with a low score on "track record for advancing diversity" eliminating candidates. Thompson calls this a political test and harks back to the University of California tradition of requiring a loyalty oath in the 1950s [as did the National Science Foundation Fellowships through the 1960s]. Hundreds of mathematicians responded to the recent piece. Some who disagreed with her also denied her right to express such sentiment (in the *Notices* or anywhere) or were strongly critical of the AMS for publishing it. Others rejected any attempt to silence her or intimidate the AMS. A few objected to the UC policy or were sympathetic to the argument of a historical parallel. The earlier paper of hers in the *Notices* disputes an alleged theorem from social scientists that "diversity trumps ability."

Math. Mag. **93** (2020) 399–400. doi:10.1080/0025570X.2020.1819055

Cheng, Eugenia, *x + y: A Mathematician's Manifesto for Rethinking Gender*, Basic Books, 2020; x+272 pp, $28, $16.99(ebook). ISBN 978-1-5416-4650-6, 978-1-5416-4651-3.

In this remarkably sensitive, thought-provoking, and largely sensible approach to the divisiveness around gender equality, author Cheng contrasts individualistic thinking with community-minded thinking, denominating the former "ingressive" and the latter "congressive." She uses a mathematical approach, based on her background in category theory, to frame a theory that looks only at how people relate rather than at biological descriptions and gender preconceptions (the book is filled with diagrams of the ideas). "[W]e can treat men and women the same if they relate to others in the same way." She claims that it is a "mistaken idea that men and women are measurably different" and that the roles that different character traits play need not be genderized; "we ask what qualities of humans we are genuinely going to value in society." She emphasizes valuing people "for what they bring" and character types "that we truly think are valuable." But what is character? and what about those with "poor" character traits or few skills? or who, such as some elderly, are completely dependent? In the classroom, congressive behavior "means collaboration, contribution to the group, depth rather than speed, and curiosity rather than knowledge"; "math could be more congressive by being about exploration and processes ... more about ways of thinking than about knowledge."

Devlin, Keith, Of course $2 + 2 = 4$ is cultural. That doesn't mean the sum could be anything else., mathvalues.org/masterblog/of-course-2-2-4-is-cultural-that-doesnt-mean-the-sum-could-be-anything-else.

As can be observed from the reviews above, mathematics is not removed from current cultural struggles over science, race, gender, and tribe. Devlin responds to a claim on Twitter that "The idea of $2 + 2$ being 4 is cultural and because of western imperialism/colonization, we think of it as the only way of knowing." Devlin asserts that mathematicians have "bought a package," a cultural creation based on reification, generalization, and abstraction from empirical facts. "The reason we mathematicians think that mathematics does not depend on culture is that *we are in that culture*" but "it's not by any stretch of the imagination Western Culture." He goes on to pose questions about mathematical cultures, claims that "mathematics is cultural—all of it," and—facetiously pursuing the set roles in a culture war—casts out his opponents, concluding that anyone who disagrees with him occupies a different mathematical culture.

Klarreich, Erica, Landmark math proof clears hurdle in top Erdős conjecture, quantamagazine.org/landmark-math-proof-clears-hurdle-in-top-erdos-conjecture-20200803/.

Erdős asked for conditions for an infinite sequence of positive integers to contain infinitely many arithmetic progressions of every length. His conjecture: It's sufficient for the sum of reciprocals of the sequence to diverge. Thomas Bloom (Cambridge University) and Olof Sisask (Stockholm University) have proved that Erdős's condition suffices for triples (progressions of length 3) (arxiv.org/abs/2007.03528). Earlier, in the 1930s, Johannes van der Corput had proved that the primes (the sum of whose reciprocals do diverge) contain infinitely many progressions of length 3. The new result shows that a more general explanation holds than the special structure of the primes. Not mentioned in the article is the related twin-prime conjecture, that the primes contain infinitely many prime pairs differing by 2.

Ellenberg, Jordan S., Geometry, inference, complexity, and democracy, *Bulletin of the American Mathematical Society* (to appear), arxiv:2006.10879v1.

By the time you read this, the 2020 U.S. elections will be over and the battle over the decennial redistricting will have begun. Ellenberg investigates what contribution mathematics can make to what it means for an apportionment of seats to be "fair," while taking into account the historical and legal constraints imposed in the various states. He explains the efficiency gap metric and shows how it is at variance with proportional representation. Then he describes ensemble sampling, the idea of comparing a particular apportionment to a sample of potential apportionments. Because of the size of the universe of potential apportionments, the sampling is best accomplished by a random walk. He ends by endorsing an analogue of the "I cut, you choose" method of fair division, in which opposing parties take turns freezing one legislative district from a map created by the other party.

Acknowledgments

In addition to our Associate Editors, the following referees have assisted the MAGAZINE *during the past year (May 2019 to May 2020). We thank them for their time and care.*

Lowell Abrams, *The George Washington University*
Claudio Alsina, *Universitat Politècnica de Catalunya*
Stephen Andrilli, *La Salle University*
Barry Balof, *Whitman College*
Edward Barbeau, Jr., *University of Toronto*
Julia Barnes, *Western Carolina University*
Ethan Berkove, *Lafayette College*
Saul Blanco, *Indiana University*
Harold Boas, *Texas A&M University*
Ethan Bolker, *University of Massachusetts, Boston*
Miklós Bóna, *University of Florida*
Carmine Boniello, *Università degli Studi di Salerno*
Andrew Bremner, *Arizona State University*
Jack Calcut, *Oberlin College*
Jason Callahan, *Saint Edward's University*
Thomas R. Cameron, *Davidson College*
Matt Davis, *Muskingum University*
Joshua Ducey, *James Madison University*
Tom Edgar, *Pacific Lutheran University*
David L. Finn, *Rose Hulman Institute of Technology*
Robert Foote, *Wabash College*
Joseph A. Gallian, *University of Minnesota, Duluth*
Jeremy Gibbons, *University of Oxford*
Russell A. Gordon, *Whitman College*
Leon Hall, *Missouri University of Science and Technology*
Rich Harris, *University of Bristol*
Brian Hopkins, *Saint Peter's University*
Danrun Huang, *Saint Cloud State University*
Thomas Hull, *Western New England College*
Hans Humenberger, *Universitat Wien*
Omid Ali Shahny Karamzadeh, *Institute for Research in Fundamental Sciences*
Steven Krantz, *Washington University in Saint Louis*
John Lorch, *Ball State University*
Martin Lukarevski, *Goce Delchev University*
Rick Mabry, *Louisiana State University, Shreveport*
Dan Ştefan Marinescu, *Iancu de Hunedoara National College*
Jane McDougall, *Colorado College*
Steven J. Miller, *Williams College*
David Nacin, *William Patterson University*
Roger Nelsen, *Lewis & Clark College*
Niels Overgaard, *Lunds Universitet*
Michael Penn, *Randolph College*
Jennifer Quinn, *University of Washington, Tacoma*
David Richeson, *Dickinson College*
Erel Segal-Halevi, *Ariel University*
Brittany Shelton, *Albright College*
David Singer, *Case Western Reserve University*
Hellmuth Stachel, *Technische Universität Wien*
Manon Stipulanti, *Hofstra University*
George Stoica, *Saint John, Canada*
Jeff Suzuki, *Brooklyn College*
Christopher N. Swanson, *Ashland University*
Donald Teets, *South Dakota School of Mines and Technology*
Eve Torrence, *Randolph-Macon College*
David Treeby, *Monash University, Australia*
Daniel Velleman, *Amherst College*
Raymond Viglione, *Kean University*
Stan Wagon, *Macalester College*
Kenneth S. Williams, *Carleton University*
Japheth Wood, *Bard College*

