LETTER FROM THE EDITOR

This is the first issue of the MAGAZINE with the MAA's new publishing partner Taylor & Francis. We look forward to working with Taylor & Francis to bring expository mathematics to you in print or electronically. This issue begins with Robert Thomas' appreciation of Theodosius of Bithynia's first book of spherics. Thomas' appreciation goes beyond historical evidence to reconstruct an ideal version of the book.

In the next article, Frederic Mynard reviews the classical example of stereographic projection to explain that the sphere can be viewed as the plane minus one point. This example of a one-point compactification motivates his observation that the pinched sphere is the one-point compactification of the punctured plane. Without using homotopy theory, he distinguishes topologically the plane from the punctured plane.

Each issue of the MAGAZINE in 2017 included a Pinemi puzzle by Lai Van Duc Thinh. In this issue, he provides the first Partiti puzzle; each issue of 2018 will include a Partiti puzzle. Andrés Caicedo and Brittany Shelton introduce the puzzle and provide some comments about partitions.

Many know that Benjamin Franklin created magic squares. But he failed to describe his method of construction. Ronald P. Nordgren analyzes one method and discusses others based on Euler's composition method.

The inspiration for an article can come from disparate sources. In "The Metric metric on S_4," Bret Jordan Benesh was inspired by a music video by the band Metric. He uses a concept from geometric group theory to define the Metric metric. For their article on determining all solutions to a functional equation, Mihály Bessenyei and Gréta Szabó were inspired by a problem from a contest problem book. Bessenyei and Szabó build up all solutions by first considering differentiable solutions and eventually use homomorphisms on groupoids to solve the problem.

In the next article, Shahnawaz Ahmed and Elias G. Saleeby provide an alternative proof for a formula for the volumes of generalized super-ellipsoids that was obtained by Dirichlet. They also use a geometric Monte Carlo method to estimate hyper-volumes, as well as consider volumes of revolution in higher dimensions.

In his article "Designing for minimum elongation," Niels Christian Overgaard revisits the calculus of variations problem of finding the shape of a vertically hanging rope so as to minimize the elongation of the rope due to the rope's own weight and the load of any object attached to the rope. He recalls an earlier solution, which he shows to be optimal, and gives three new complete solutions to the problem.

In their article, Arthur Befumo and Jonathan Lenchner begin by recalling Solomon Golomb's tromino theorem that proves that $2^n \times 2^n$ chess board can be covered by an L-shaped tromino. They extend Golomb's result to higher dimensions and consider covering chess boards with the straight tromino.

In between the articles are three proofs without words by Andrzej Piotrowski, Ángel Plaza, and Brian Hopkins. The issue concludes with the Problems and Reviews.

Finally, a warm farewell to Julie Beier. She no longer has the time to serve on the editorial board now that she has left academia. I will miss her sound advice. Good luck Julie!

Michael A. Jones, Editor

An Appreciation of the First Book of Spherics

R. S. D. THOMAS
St. John's College and University of Manitoba
Winnipeg, Manitoba R3T 2N2 Canada
robert.thomas@umanitoba.ca

The first three books of Euclid's *Elements* provide the charming classical introduction to deductive mathematics with the geometry of the blackboard. A blackboard with straightedge and a pair of compasses is old-fashioned, but the mathematics is also old-fashioned and none the worse for that, unlike two-thousand-year-old physics or chemistry. This paper introduces the equally old but less known geometry of the spherical blackboard and compasses.[1]

What can be done with compasses on a spherical blackboard? A surprising lot that this introduction will not discuss. The analog on the sphere of a straight line in the plane is a great circle, and to draw a great circle on the sphere you need the right setting for the compasses (Propositions 16 and 17), and that depends on the size of the sphere, surprisingly determinable (Proposition 19). With great circles playing the role of lines, you can set about moving the *Elements* to the sphere.

Introduction

The extant treatise on spherics, by Theodosios of Bithynia[2] (*fl.* 100 BCE ±100), is a work in three books as Euclid's *Elements* is a work in 13 books. As in the *Elements* the beginning is of special interest. I shall have almost nothing to say about the two more advanced books except that they depend on the first and are important in Euclid's astronomy book *Phenomena* [**3**]. If an ideal version of a classical work deteriorates over time because of handwritten transmission, then we could think of the "first book" in its present version[3] of the text as being an inferior copy of the original—not too unlike photocopies, of photocopies, of photocopies, … of an original. At some time between the temporally first book on the subject and the present text (itself copied for several hundred years), there may have been a version with all of the virtues possible, a "book" in somewhat the sense of Erdős.[4]

Math. Mag. **91** (2018) 3–15. doi:10.1080/0025570X.2017.1404798
MSC: Primary 01A20, Secondary 52-03
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.
[1] For a picture of the spherical blackboard at the up-to-date research Institute Mittag-Leffler in Djursholm, Sweden, see page three of www.ams.org/publications/journals/notices/201606/rnoti-p604.pdf.
[2] Not "of Tripoli,' as he is misidentified in [**8**], [**11**], and for practical purposes in [**6**]. Bithynia is an ancient piece of modern Turkey.
[3] What I refer to as the present version does not exist in any ancient form. It is an expert's opinion of what text lies behind the best of the Greek manuscripts carefully collected and compared, a so-called "critical edition," of which there are two, [**6**] and [**4**].
[4] Paul Erdős used to say that ideal proofs were in a book of just such proofs maintained by God. A collection of actual proofs inspired by this idea is [**1**].

This paper attempts to reconstruct that better version because it will be easier to appreciate than the present text, which has had some bad press. In particular, Thomas Heath [5] says that the work as a whole is applied and not interesting. *How much* of this reconstruction is historical may be a difficult historical question, but I intend to suggest how good the book might have been by showing what is still there if one looks for it carefully. I am not making anything up except the obvious final Corollary 5.

Consideration of this book is a worthwhile exercise for a couple of reasons. It is simple enough to be studied in high school except that is uses three-dimensional geometry. While deductive, it is apparently preaxiomatic, since there are no postulates either in the text or referred to. The game was afoot, but we do not know what the rules were. So it is a window on Greek geometry before the axiomatization accomplished by Euclid. What it uses are obvious facts about circles, things one can do with circles, and simple three-dimensional geometry never involving spheres to show results about circles on spheres. There seems to be no restriction on the geometrical knowledge that can be called upon except that no knowledge about the topic of the book is assumed. This is a straightforward methodological principle to which I need to draw attention. One naturally thinks of Euclidean theorems as encapsulating results used, but specific reference to Euclidean theorems is anachronistic since the (original) book was almost certainly written before the *Elements*. The present text is later than the *Elements*.

Since this is not a work of history, I am going to avoid historical questions. This exploration may be useful to historians in considering spherics, since their proper adherence to the degenerate textual evidence is not conducive to exploring possibilities. To have possibilities set out may encourage a reevaluation of the work. That will be properly historical work. My reconstruction is partly based on the work of Sidoli and Saito [7], which penetrates into what the work is trying to accomplish, not obvious in the text. They draw attention to the apparent practical goal of the book and how that differs from the way the book begins. (See also [9].)

It is not difficult to set out in general terms what the book accomplishes. It begins with the result that a plane through three points on a sphere cuts it in a circle. And it ends by showing how, with a pair of compasses that can transfer distances (unlike the collapsing compasses one can limit oneself to in Euclid's first book)—and a straightedge when working in the plane—it is possible to draw that circle on the surface of the sphere. This is perhaps not much, but it is not trivial and it gives the book unity and an appeal especially to the ancient Greeks. Along the way it shows how to determine the diameter of a sphere, a construction that is required in the final constructions (of each Platonic solid inscribed in a sphere) in book thirteen of the *Elements* and not something one can easily think how to do with straightedge and compasses. The document, if not a gem (polyhedral), is at least a pearl (spherical).

I shall set out the 21 propositions of the book, giving some indication of most proofs and simplifying some. The novelties that I am introducing are limited to interpretation and Corollaries 3–5 of the results that are in the version of Theodosios. My suggestion is that these few additions may have been in some former version of the book; even if they were not, they make it sounder and more interesting without deviant novelty.

The sections of the paper discuss definitions, relations among the defined terms (part 1), planes and the sphere, relations (part 2), great circles, the configuration of central importance consisting of a great circle perpendicularly bisecting another circle, the polar radius of a great circle, and then the final construction of the diameter of the sphere, great circles, and the pole of a given circle.

Definitions

The book begins with five definitions, of which I give literal translations.[5]

Definition 1. A *sphere* is a solid figure contained by a single surface, all lines to which, falling from a single point that lies within the figure, are equal to one another.

Definition 2. *Center of the sphere* is the point.

Definition 3. *Axis of the sphere* is a line passing through the center and bounded in each direction by the surface of the sphere, around which fixed line the sphere rotates.[6]

Definition 4. The *poles of the sphere* are endpoints of the axis.

Definition 5. *Pole of a circle in a sphere* names a point on the surface of the sphere all lines from which, falling on the circumference of the circle, are equal to one another.

In this context, a circle is a circular disk not just its circumference, and a circle's being *in* a sphere means not just being in it somewhere but having its circumference on the surface of the sphere. Accordingly, when one is said to *draw* a circle that means to draw its circumference on the surface of the sphere. Doing that requires the pole of the circle and compasses set to a distance for which Greek has no term. I shall refer to the straight-line radius for the compasses as the *polar radius* of the circle.

A necessary and sufficient condition that functions as a definition for a line to be perpendicular to a plane is that it is perpendicular to every line in the plane through its point of intersection with the plane.

Spheres and circles have centers, and a circle (on the sphere understood) has two poles because the point antipodal to the pole one would use to draw it is also equidistant from all of its points. Sorting out the relations among these four points for a circle is the subject of the next section and Relations among centers and poles (part 2).

The two initial propositions

My interpretation of the work as a whole is that it accomplishes just before its end its *practical* goal of being able to draw a great circle through any two points on the sphere (or likewise to extend a great-circle fragment to the whole circle). And also at its end its satisfying *theoretical* goal of being able to draw the circle through three points on the sphere, the circle that Proposition 1 assures us is the intersection of the sphere with the plane through the three points. That a plane is determined by three noncollinear points or two intersecting lines is *Elements* XI.2. This is an example of the sort of outside information called upon and usually not stated.

Proposition 1. *The plane through three points A, B, and C on the surface of a sphere cuts the surface of the sphere in the circumference of a circle.*

Proof. The perpendicular from the center of the sphere to the plane is a common side of right triangles, each with a radius of the sphere as hypotenuse. Since these two sides of the right triangles are equal, the third sides are all equal, and they are the equal radii of what must therefore lie on a circle with the foot of the perpendicular as its center. Easier still if the plane passes through the center of the sphere. ■

[5] A full literal translation of a critical edition of Book one into English is available on request. An English translation of a Latin translation from Greek is [**8**]; French translations are [**11**] and in [**4**].

[6] This and the next definition are of significance for the astronomical relevance of the more advanced books, but play no part in the work itself.

Corollary 1. *If a circle is in a sphere, the perpendicular from the center of the sphere to it falls at its center.*

The text makes the next proposition a problem, that is, a construction, but it is no more a construction to find the center of the sphere than Proposition 1 is a construction to find the center of the circle. There is a historical problem here, but for learning purposes this proposition is better considered a theorem, leaving constructions with compasses (and straightedge in the plane) to the final four propositions.

Proposition 2. *To find the center of a given sphere.*

Proof. The construction is to erect the perpendicular at the center of a circle produced by cutting the sphere with a plane, produce it both ways to become a diameter, and bisect it. Proof by contradiction. ■

Corollary 2. *If a circle is in a sphere and a perpendicular to its plane is erected at its center, the center of the sphere is on the perpendicular.*

What look like constructions in the proofs of Propositions 1 and 2, since there are no instruments to perform them, are better thought of as thought experiments, in which one can readily drop a perpendicular from the center of the sphere to the intersecting plane (in 1) and erect the perpendicular at the center of the circle (in 2) and bisect it when it is extended to be a diameter. While there is no way to *do* these things, one has no doubt that there is, internal to the sphere, a line perpendicular to the plane and through the center of the circle or sphere and that it has a mid-point. If the powers drawn on here were real instead of experimental thoughts, one could accomplish the goals of the work without doing much. For instance, one could simply measure the length of the diameter in Proposition 2 rather than go to the trouble of figuring out how long it is with a real construction in Proposition 19. The aims would become trivial. Drawing attention to this distinction between notional constructions and real constructions is the great contribution of Sidoli and Saito [**7**] to the understanding of this work. While one sort of construction is used to get the book going and in subsequent proofs, the aim of the book is constructions of an altogether different and more restricted kind.

Planes and the sphere

It is anachronistic to state results in terms of specifics, *e.g.*, "circle *ABC*," rather than the generality of the Euclidean style, "a circle." The general prose enunciations in the *Spherics* are hard to understand and not always complete or accurate. My excuses are clarity and also that we do not know that pre-Euclidean geometry was written Euclid's way.

The diagrams are my constructions to help with visualization. The manuscript diagrams (sometimes missing altogether) are not easy to decipher; studying them would be a work in itself. No plane diagram of a spatial configuration is entirely easy to read; these are relatively accurate representations made with Mathematica.

This section and the next, except for Proposition 6, can be omitted on first (or any) reading, and the proof sketches should probably be omitted by anyone not interested in the book for its own sake. Propositions 3–5 deal with planes that are tangent to the sphere rather than cutting it. One *does* need to know about great circles (Propositions 6, 11, and 12).

Proposition 3. *A sphere touches a plane in only one point.*

Proof. Proof by contradiction. ■

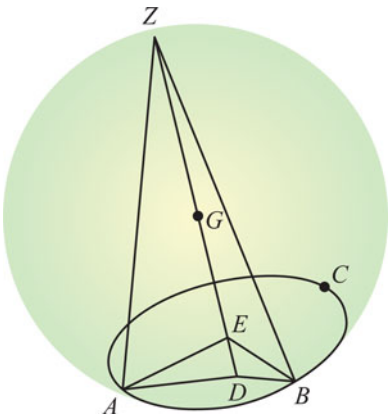


Figure 1 Diagram for Propositions 7 to 10. E is the center of circle ABC, G is the center of the sphere, and D and Z are the poles of the circle on the surface of the sphere.

Proposition 4. *Let a plane* Π *touch (but not cut) a sphere with center* B *at a point* A. *Then the line joining the point of contact* A *to the center* B *is perpendicular to* Π.

Proof. Each plane containing the line AB cuts the sphere in a great circle and cuts the plane Π in a straight line perpendicular to radius AB. Accordingly AB is perpendicular to Π. ■

Proposition 5. *If a sphere touches a plane that does not cut it, then the center of the sphere is on a perpendicular erected into the sphere at the point of contact.*

Proof. Proof by contradiction using Proposition 4. ■

The next theorem is about planes too, for all circles in the sphere have planes.

Proposition 6. *Circles through the center of a sphere are great circles.*[7] *Other circles in a sphere are equal to each other if equidistant from the center of the sphere, and the farther away from the center the smaller the circles.*

Proof. Consideration of right-angled triangles with the radius of the sphere as hypotenuse and the perpendicular to the center of the circle as one side and the radius of the circle as the other side shows how, since the length of the hypotenuse is fixed, the other sides are related. ■

Relations among centers and poles (part 2)

Propositions 7–10 make statements about the configuration shown in Figure 1, which shows the center of the sphere (there G) collinear with the center of the circle ABC (there E) and the pole (there D); if DEG is extended, then it reaches the surface again at the other pole of the circle, Z.

Proposition 7. *If a circle* Γ *is in a sphere* Σ, *a straight line through the center of* Σ *and the center of* Γ *is perpendicular to* Γ.

[7] The Greek word is just "greatest," which is descriptive (largest possible). Since English has a technical term, I use it. Circles not great are, in this sense, *small*.

Proof. Consideration of what turn out to be congruent triangles (Figure 1) with GE joining the centers as common side, GA or GB as second side and circle radius EA or EB as the other side show that the angles at E are all equal and so are right angles. ■

Proposition 8. *If a perpendicular is dropped from the center of a sphere to a circle in the sphere and extended in both directions, it meets the sphere at the poles of the circle.*

Proof. From the center of the sphere G let a perpendicular be dropped to the circle at E (Figure 1), its center, and the points at which the perpendicular extended meets the sphere be D and Z. Consideration of right-angled triangles DAE and DBE, for any two points A and B on the circle, with radii EA and EB equal and side ED common, shows that sides AD and BD are equal and so are polar radii of the circle. So D is a pole of the circle. Similarly Z is a pole of the circle. ■

Proposition 9. *If a perpendicular is dropped to a circle in a sphere from one of its poles D, it falls on the center of the circle, and extended it meets the sphere at the other pole of the circle Z.*

Proof. Let the foot of the perpendicular to the circle be E (Figure 1). Consideration of right-angled triangles DAE and DBE, for any two points A and B on the circle, with polar radii AD and BD equal and side ED common, shows that sides EA and EB are equal and so are radii of the circle. The foot of the perpendicular E is the center of the circle. Then consideration of right-angled triangles EAZ and EBZ, with the (common) portion of the diameter of the sphere EZ as one side and the radii of the circle EA and EB as the other side, shows that the hypotenuses AZ and BZ are equal and so are also polar radii of the circle. The opposite end Z of the diameter is also a pole of the circle. ■

Proposition 10. *If a circle is in a sphere, the line joining its poles D and Z is perpendicular to the circle and passes through the centers of the circle and of the sphere.*

Proof. Without loss of generality if the circle is small, let D be closer to the circle than Z (Figure 1), and let the line DZ pass through the plane of the circle at E. Consideration of triangles DAZ and DBZ, for any two points A and B on the circle, with polar radii DA and DB equal and polar radii AZ and BZ equal and side DZ common shows that the angles ADZ and BDZ are equal. Consideration of triangles with polar radii DA and DB equal, side DE common, and the equal angles at D between them shows that they are congruent. The special case of B opposite to A shows that the equal angles BED and AED are right. DE is perpendicular to the circle. Accordingly, E is the center of the circle by Proposition 9. DZ is perpendicular to the circle and so passes through the center of the sphere G by Corollary 2. ■

Great circles

Proposition 11. *In a sphere, two great circles bisect each other.*

Proof. If the center of the sphere is joined to the points at which the circumferences of the circles intersect, the resulting lines lie in the plane of each circle, therefore in the intersection of their planes, and so are a single straight line, which is a diameter of the sphere and of both circles, which therefore bisect each other. ■

Proposition 12. *In a sphere, circles that bisect each other are great circles.*
(Converse of 11)

Proof. Perpendiculars erected at the centers of the bisecting circles contain the center of the sphere, but they intersect at the center of the circles, which is common. Their center is the center of the sphere. ■

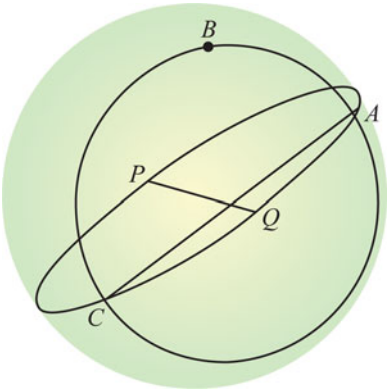


Figure 2 Diagram for Propositions 13 to 15. P and Q are the poles of circle ABC.

The three middle theorems on the key configuration

The cycle of Theorems 13–15 concerns the same configuration in a sphere, a small circle ABC and a great circle $PAQC$ through its poles P and Q (Figure 2). They show that three conditions are equivalent, as we would say, by showing that each implies the other two:

1. the great circle bisects ABC with line AC,
2. the great circle cuts ABC perpendicularly, and
3. the great circle passes through the poles of ABC.

The propositions say what knowledge gives what knowledge. 13: (2) implies (1) and (3). 14 : (1) implies (2) and so (3). 15: (3) implies (2) and so (1).

Note that when circle ABC is great, condition (1) is automatic and so cannot imply (2) or (3), but (2) and (3) are equivalent by Proposition 10.

Proposition 13. *If a great circle is perpendicular to small circle ABC, then the great circle bisects ABC and passes through its poles.*

Proof. For a small circle ABC, because the line PQ joining its poles runs perpendicularly through its center and the center of the sphere in the plane of $PAQC$, AC is a diameter of ABC. ■

Proposition 14. *If a great circle bisects small circle ABC, then it passes through the poles of ABC and is perpendicular to it.*

Proof. The line joining the center of the sphere and of $PAQC$ to the mid-point of diameter AC of circle ABC is perpendicular to ABC by Proposition 7. Because it is also in $PAQC$, containing both the center of the sphere and of AC, $PAQC$ is perpendicular to ABC. Passing through the poles is a consequence of the perpendicularity by Proposition 13. ■

Proposition 15. *If a great circle passes through the poles of ABC, then it bisects ABC perpendicularly.*

Proof. The line PQ between the poles of ABC is perpendicular to ABC, and so $PAQC$ must be perpendicular, and therefore bisects ABC by Proposition 13. ■

The polar radii of great circles

One can see in Figure 2 that there is room for a second great circle perpendicular to *PAQC* bisecting circle *ABC* perpendicularly. Such great circles bisect each other by Proposition 11. If circle *ABC* is also great, then a symmetric configuration is created in which each circle contains the poles of both the others. Joining poles to adjacent poles produces an octahedron (Figure 3). This configuration motivates and illustrates Propositions 16 and 17. It also motivates the final stage of Euclid's *Elements* in Book 13. The octahedron is the easiest of the Platonic solids to inscribe in a given sphere. Euclid does not draw them inside a sphere but, on the basis of the extracted diameter of the sphere, constructs the solid to be the right size so that the equal distances of all its vertices from its center is the radius of the given sphere. An equal sphere would fit around it perfectly.

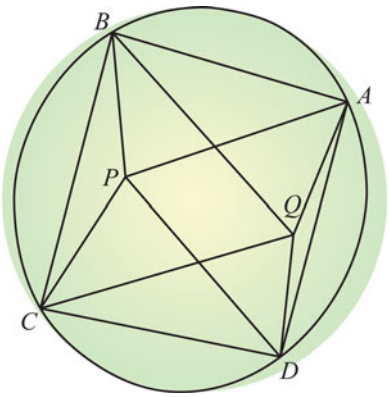


Figure 3 Octahedron inscribed in the sphere with, as edges, 8 polar radii of great circle *ABCD* from *P* and *Q* and a square inscribed in circle *ABCD* with sides equal to the polar radius.

To draw a great circle, it is necessary to know its polar radius. These theorems give a necessary and sufficient condition for a circle to be great in terms of its polar radius.

Proposition 16. *The polar radius of a great circle ABCD in a sphere is equal to the side of the square inscribed in the (or any) great circle.*

Proof. Let *AC* and *BD* be perpendicular diameters of circle *ABCD* meeting at *E* and let *A* be joined to pole *P* of *ABCD* and to *B*. Then *AB* is a side of a square inscribed in a great circle. Right triangles *AEP* and *AEB* are congruent, and so *AP* equals *AB* as required. ■

Proposition 17. *If the polar radius of a circle ABC in a sphere is equal to the side of a square inscribed in a great circle, then ABC is a great circle.*
(Converse of 16)

Proof. Great circle *PAQC* (Figure 2) has polar radii *PA* and *PC* (Figure 3) equal to the side of square *PAQC* inscribed in a great circle. Since *AC* is a diameter of great circle *PAQC* and so also of circle *ABC*, circle *ABC* must therefore be a great circle. ■

Final problems for straightedge and compasses

These four problems complete the work in four steps:

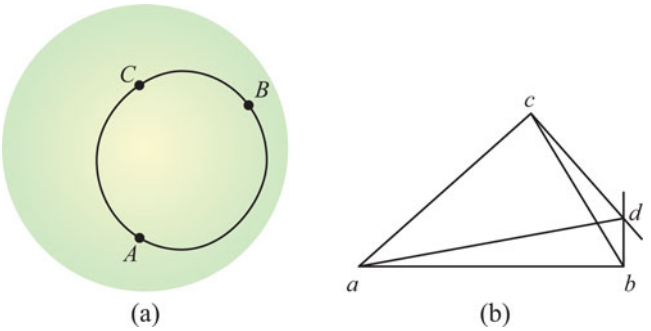


Figure 4 Proposition 18. (a) *A*, *B*, and *C* on the circle. (b) Triangle *abc* congruent to triangle *ABC* with perpendiculars at *b* and *c* meeting at *d*.

- Finding the diameter of a circle in the sphere, given the circle or (we can see) three points on it (not in the text).
- Finding the diameter of the sphere from the outside.
- Finding the polar radius of great circles, which allows drawing a great circle through any two points.
- Finding the pole of a given circle, which is necessary in later books. We can see, though (not in the text), that three points on the circle suffice, meaning that the pole can be found without the circle, which allows the circle through three points to be drawn.

The first of the problems involves the insight that allows the problem to be transferred to the plane by transferring distances with compasses. As a planar problem it is subject to the usual techniques with straightedge and compasses.

Proposition 18. *Given A, B, and C, points on the circumference of a circle in a sphere; to construct a line equal to the diameter of circle ABC.*

Proof. Let triangle *abc* be constructed in a plane (Figure 4(b)) so that *ab* is equal to chord *AB*, *ac* is equal to chord *AC*, and *bc* is equal to chord *BC* (Figure 4(a)). Also, let perpendiculars to *ab* and *ac* be drawn toward each other at *b* and *c* intersecting at *d*. Let *ad* be joined.

Since *ad* subtends a right angle at *b*, it is the diameter of a semicircle through *b*. Similarly, *ad* is the diameter of a semicircle through *c*. But having the same diameter, these are halves of the only circle with diameter *ad*, the undrawn circumcircle of triangle *abc*, which is equal to the circle *ABC* in the sphere. Since *ad* is a diameter of circle *abc*, it is equal to the diameter of circle *ABC*. ■

Corollary 3. *Given A, B, and C, points on the surface of a sphere; to construct a line equal to the diameter of the circle through A, B, and C.*

Proof. The circle *ABC*, which exists on account of Proposition 1, was not used in the construction. ■

Consider the two configurations, the triangle *ABC* in the sphere and the triangle *abc* in the plane. It is obvious that the isometry that took *ABC* to *abc* can be extended to the whole of the circumference of the circle *ABC* by a process that can be called triangulation. Any point *X* on the circumference is the vertex of an oriented triangle *XAB* (or *XBC* or *XCA*), and that triangle can be transferred isometrically to an oriented triangle *xab* (etc.), where *x* will be the point, located by intersecting circular arcs, on

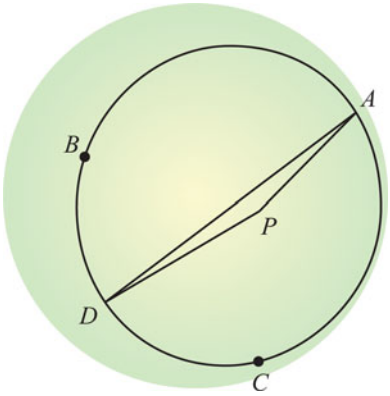


Figure 5 Diagram for Proposition 19. *P* is the pole of circle *ABC,* and *D* is diametrically opposite to *A*.

the undrawn circumference of the circumcircle of triangle *abc*. The circle does not need to be drawn; the isometry between the points on the circumferences of the two circles does not depend on the drawing of either circle, although a circle can help in identifying such points, of course. If the circle *ABC* is not given, we do not yet know how to draw it.

Moreover, using points already identified as points on the circumcircle of triangle *abc*, the same triangulation method acting on further points, can be used for the inverse of the original isometry. In particular, the point *d* can be mapped back to the point *D*, diametrically opposite to *A*, in three different ways, using oriented triangle *abd*, *acd*, or *bcd*. This renders the presence of the circle *ABC* superfluous not only in the original construction of *ad* but in the determination of *D*. Triangle *bcd* is most convenient and, when *ABC* is a great circle, the only triangle one can use (*AD* is then degenerate as a polar radius). So let *D* be determined by the inverse transformation applied to triangle *bcd*. Circles with poles *B* and *C* and polar radii *bd* and *cd* cross at *D*.

That this author should be responsible for the clever transfer of a problem of three-dimensional geometry inaccessibly inside a sphere to an easy problem in a plane is not as surprising as it might be if something similar were not also done in the fifth book of the *Elements*, in which problems of ratios of magnitudes are transferred to problems having to do with the lengths of straight lines. This sort of transfer we associate with Descartes, but his co-ordinate geometry is only the most vast and thoroughgoing such transfer up to his time and far beyond. That of Book five of the *Elements* seems to be the first important such transfer. If such transfers are not reversible, then they are of less interest, but they can still be important (invariants, *e.g.*, knot polynomials).

Corollary 4. *Given points A, B, and C on the surface of a sphere, to construct the point D on the circle through A, B, and C diametrically opposite to A.*

Proof. The circle exists by Proposition 1. Map *d* back to the sphere at *D* using triangle *bcd*. ■

These corollaries are not stated in the text of Theodosios, but they are both implicitly called upon for the proof of Proposition 19. They are the only way on offer to find the point *D* and the line *pr* below.

Proposition 19. *Given a sphere, to construct a line equal to the diameter of the sphere.*

Proof. Take any two points *P* and *A* on the surface of the sphere. With pole *P* and polar radius *PA* draw small circle *ABC* (Figure 5). With Corollary 4, determine the point on

the circle D diametrically opposite to A using the three points A, B, and C. The planes containing the poles of the circle cut the sphere in great circles, one of which contains triangle APD, bisecting the circle by Proposition 15. With Corollary 3, construct a line pr equal to the diameter of the great circle APD using points A, P, and D. Line pr is then equal to the diameter of the sphere. ■

It is not clear in the last book of the *Elements*, where a sphere is given and it is required to inscribe each Platonic solid in it, whether Proposition 19 is needed or not. But the diameter of the given sphere must be determined. No method is prescribed, which argues for this method. It has been argued to the contrary that, since Euclid defines a sphere as a solid of revolution of a semicircle, the diameter of the sphere can be obtained from the semicircle. I reject this argument because the diameter of a sphere so defined is no more obvious than the radius of a sphere defined as in the *Spherics*. How a sphere is defined—with equivalent definitions—makes no difference to what it looks like or how its diameter can be extracted. The question is how to play by the rules where the rules have not been specified. In this instance, they are not specified by Euclid either.

Proposition 20. *Given two points A and B on the surface of a sphere, to draw a great circle through A and B.*

Proof. If AB is a diameter of the sphere, then there are many great circles through A and B; we return to this case. Otherwise construct a line de in a plane equal to the diameter of the sphere. Construct the right bisector of de at f and extend it so that fg can be cut off equal to fe. Then eg is the side of a square inscribed in a great circle in the sphere and by Proposition 17 is the polar radius of great circles on the surface of the sphere. With poles A and B and polar radius equal to eg, draw great circles on the surface of the sphere to intersect at C and D if possible. If AB is a diameter of the sphere, only one circle will result. In that case, with any point on that single great circle as pole any great circle drawn will pass through A and B, and the construction is complete but not unique. If AB is not a diameter of the sphere, then a circle drawn through A with pole C or D will pass through B and will be great because its polar radius is equal to eg. ■

This construction has an easy extension and one more difficult. It is easy to draw a great circle touching a given circle at a given point, since the great circle's center must be on the great circle through the given point and the pole of the given circle. This construction is given in the second book (Proposition 14) when enough theory has been developed to allow the more difficult construction to be proved correct, to draw a great circle touching a given circle and through a feasible given point (Proposition 15).[8]

Proposition 21. *Given a circle ABC in a sphere, to find a pole of the circle.*

Proof. If ABC is a great circle, then great circles drawn with poles on it will cross at its poles P and Q. This would be indicated by the great circle through A and B passing through C.

If ABC is not a great circle, then find the points Z and K diametrically opposite to A and B, respectively, in the circle. The great circles through A and Z and through B and K will bisect ABC perpendicularly and so intersect on the surface of the sphere at the pole P, illustrated in Figure 6, and the other pole. ■

Corollary 5. *Given A, B, and C, points on the surface of a sphere, to draw the circle through the points A, B, and C.*

[8] An overview of the whole work in [**10**].

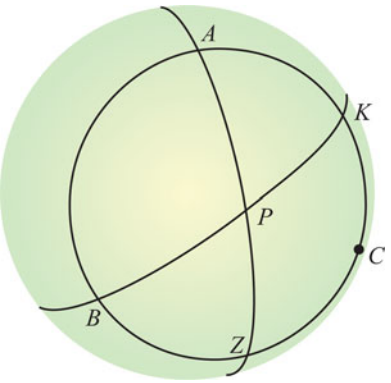


Figure 6 Diagram for Proposition 21. *A*, *B*, and *C* are the given points. *Z* and *K* are diametrically opposite to *A* and *B*. The great circles *AZ* and *BK* intersect at *P*, the pole of the circle.

Proof. By Proposition 1 there is a circle through points *A*, *B*, and *C*. Points *A*, *B*, and *C* suffice to find a pole *P* of the circle. With pole *P* and polar radius *PA*, the circle can be drawn. ■

The construction to draw this circle rounds out the claim at the beginning of the book that there is a circle through the points *A*, *B*, and *C*. Not only does it exist in principle, but it can be constructed. I cannot think that a mathematician could be responsible for what we find in the text and not also take this conclusive step! While it is missing, the book does seem to have lost its ending.

REFERENCES

[1] Aigner, M., Ziegler, G. M. (2010). *Proofs from the Book*. 4th ed. New York: Springer.
[2] Berggren, J. L. (1991). The relation of Greek spherics to early Greek astronomy. In: Bowen, A., ed. *Science and Philosophy in Classical Greece*. New York: Garland, pp. 227–248.
[3] Berggren, J. L., Thomas, R. S. D. (2006). *Euclid's Phænomena: A Translation and Study of a Hellenistic Treatise in Spherical Astronomy*. History of Mathematics Sources, Vol. 29, 2nd ed. Providence: American Mathematical Society and London Mathematical Society.
[4] Czinczenheim, C. (2000). *Édition, traduction et commentaire des Sphériques de Théodose.* Lille: Atelier national de reproduction des thèses. (Thèse de docteur de l'Université Paris IV.)
[5] Heath, T. L. (1921). *A History of Greek Mathematics*. (1981 reprint of Oxford University Press edition.) New York: Dover.
[6] Heiberg, J. L. (1927). Theodosius Tripolites [word deleted in *corrigenda*] Sphaerica, *Abhandlungen der Gesellschaft der Wissenschaften zu Göttingen, Philologisch-historische Klasse* (N.S.) 19(3):i–xvi and 1–199.
[7] Sidoli, N., Saito, K. (2009). The role of geometrical construction in Theodosius's *Spherics*. *Arch. History Exact Sci.* 63:581–609.
[8] Stone, E., trans. (1721). *Clavius's Commentary on the Sphericks of Theodosius Tripolitae: or, Spherical Elements, Necessary in all Parts of Mathematicks, Wherein the Nature of the Sphere is Considered.* London: Senex, Taylor, and Sisson.[9]
[9] Thomas, R. S. D. (2013). Acts of geometrical construction in the *Spherics* of Theodosios. In: Sidoli, N., Van Brummelen, G., eds. *From Alexandria, Through Baghdad: Surveys and Studies in the Ancient Greek and Medieval Islamic Mathematical Sciences in Honor of J.L. Berggren*. Berlin: Springer, pp. 227–237.
[10] Thomas, R.S.D. (2018). The definitions and theorems of The Spherics of Theodosios. In: Zack, M., Schlimm, D., eds. *Research in History and Philosophy of Mathematics - The CSHPM 2017 Annual Meeting in Toronto, Ontario*. Basel: Birkhäuser, to appear.
[11] Ver Eecke, P. (1959). *Les Sphériques the Théodose de Tripoli*, 2nd ed. Paris: Blanchard.

[9] Available at http://www.archive.org.

Summary. This paper offers an understanding of the contents of the first book of spherics, whoever wrote it—a rational reconstruction of what it may once have said that goes beyond the historical evidence, which is the second-century BCE *Spherics* of Theodosios. The reconstruction is done by bringing to the fore what is glossed over and adding the missing conclusion.

ROBERT THOMAS (MR Author ID: 197523, ORCID 0000-0003-4697-4209) is retired from 41 years at the University of Manitoba successively in the departments of computer science (8), applied mathematics (20), and mathematics (13), much of that time teaching engineering students, whom he enjoys, rather than geometry, which he enjoys. In the late 80s, he translated the *Phenomena* of Euclid, which applies the mathematics of the *Spherics*. Translating the latter was an obvious extension; finding it so interesting was a bonus. He also edits *Philosophia Mathematica*.[10]

[10] https://academic.oup.com/philmat/.

Solution to Partiti Puzzle

17 269	3 3	10 127	19 568	3 3	22 1489
8 17	17 458	9 9	4 4	2 2	18 567
9 36	2 2	16 367	1 1	12 39	13 148
17 458	10 19	4 4	8 8	18 567	2 2
15 267	3 3	5 5	2 2	5 14	12 39
23 1589	4 4	16 178	18 369	20 578	8 26

Math. Mag. **91** (2018) 15. doi:10.1080/0025570X.2018.1404835

Distinguishing the Plane from the Punctured Plane Without Homotopy

FRÉDÉRIC MYNARD
New Jersey City University
Jersey City, NJ 07305
fmynard@njcu.edu

What is topology about?

This is a note about topology, meant to illustrate a specific technique. If you do not know what topology is about, think of it this way: if you look at two congruent triangles in the plane, you consider them to be the same, because one is the image of the other under a rigid transformation. Hence, you look at objects up to rigid transformations. In topology, you look at objects up to a much larger class of transformations, called *homeomorphisms*, that is, bijections that are continuous with continuous inverse. If you are considering objects living in Euclidean space, think of these as the kind of deformations undergone by a piece of clay on the potter's wheel but without tearing, punching holes, or cutting.

Without getting into technicalities, a *topology* on a set is the kind of structure needed to make sense of limits, hence of continuity, for a continuous map is nothing but a limit-preserving map. Looking at things up to homeomorphisms means identifying a lot of seemingly different things. A well-known joke is that a topologist is someone who does not see the difference between a coffee cup and a donut. This is because the corresponding external surfaces are indeed topologically equivalent, that is, *homeomorphic*. Topology seeks to identify differences that are more profound than mere measurements, hence topology does not care to see the difference between a coffee cup and a donut. To distinguish two objects topologically, topologists use *invariants*, that is, properties that are left unchanged by homeomorphisms. Of course, if one object has this property and another does not, they cannot be homeomorphic. Hence, to a large extent, topology is the study of (topological) invariants.

Unlike objects studied in analysis such as spaces of functions, surfaces such as spheres, tori, and more generally manifolds, all have the same "local" topological structure: if you just look at a neighborhood of a given point, they all look like Euclidean space (topologically). Hence to distinguish two surfaces of the same dimension, differences in their global topological structures need to be identified. *Homotopy* is a (topological) tool to that end. It looks at the kind of closed curves that can be drawn on a surface and whether they can be deformed to a single point or not. In other words, it seeks to see if closed curves "catch" something or not. To illustrate this, consider Figure 1.

Think of the red curves as rubber bands placed around the surfaces. On the sphere, it is clear that you can remove the rubber band. However, the red curve on the torus, if it were a rubber band in that position, could not be removed without cutting either the rubber band or the torus. Homotopy formalizes these considerations and is thus instrumental in identifying, among other things, the number of "holes" in the surface, which turns out to be an invariant. Homotopy distinguishes for instance between the sphere

Math. Mag. **91** (2018) 16–19. doi:10.1080/0025570X.2017.1404797
MSC: Primary 57N05, Secondary 54B15, 54D05, 54D35
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

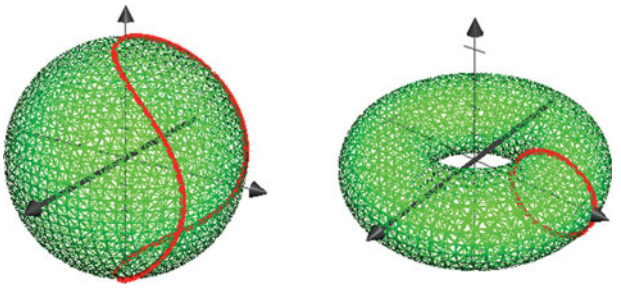


Figure 1 Jordan curves on the sphere and the torus.

and the torus. Homotopy theory however is more sophisticated than basic point-set topology, and what I want to discuss here is a work around to developing its machinery, in a simple case.

Namely, when introducing homotopy theory, it is often (e.g., [**1**], [**2**]) noted as motivation that it is difficult to distinguish topologically by elementary means the plane from the *punctured plane*, that is, the plane with one point removed. It is not hard to convince yourself that closed curves in the plane can be deformed to a single point, while curves on the punctured plane that "catch" the removed point cannot be. But this intuitive argument relies on the machinery of homotopy.

Yet, one may ask for an "elementary" argument to distinguish the plane from the punctured plane, and it seems hard to find alternatives in the literature. The idea presented here is a variant of a standard argument to distinguish the torus from the sphere—another standard motivation to introduce homotopy. Roughly speaking (in fact, too roughly to be really correct), an object is *connected* if it is "in one piece," and if it is not, each "piece" is a *connected component*. To work around homotopy, one may note that a *Jordan curve* (that is, an homeomorphic copy of the circle) separates the sphere into two connected components, while a Jordan curve on a torus may not separate it into two connected components. Such Jordan curves appear in Figure 1 where Jordan curves are drawn in red.

This type of argument does not work to distinguish the plane from the punctured plane. The purpose of this note is to observe that we can nevertheless adapt it to distinguish these two spaces, using *compactifications*—a term I will explain shortly. To be completely honest, I should point out that while the fact that a Jordan curve separates the plane or the sphere into two connected components is intuitively clear, proofs relying on elementary means are not that simple, even though not necessarily too hard, e.g., [**3**]. Therefore, one may argue whether my argument is truly more "elementary" than the classical one. Nevertheless, I think this is an interesting illustration of how compactifications can be used.

Compactness and compactifications

Compactness is a central concept of topology. In the case of objects living in Euclidean space, *compact* means bounded and *closed*, that is, it contains the limits of sequences on it. Hence, the sphere and torus that we considered above are compact subsets of three-dimensional space, while the plane is not bounded, hence not compact. The general definition is more technical, but basically ensures that many things converge, which turns out to be the key to many existence results, by relying on the existence of certain limit points by compactness.

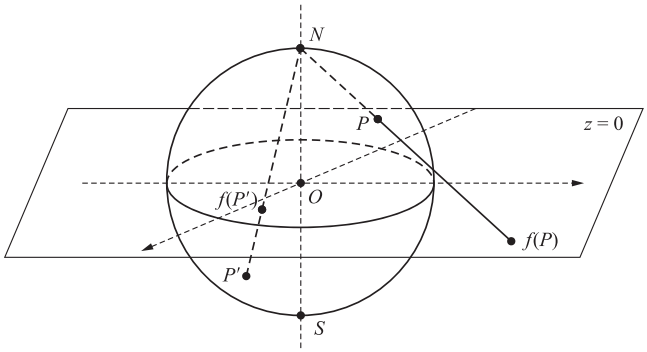


Figure 2 The stereographic projection.

A standard exercise when studying compactness it to show, for instance via stereographic projection, that the real line "plus one point" is homeomorphic to the circle, and that the plane "plus one point" is homeomorphic to the two-dimensional sphere. A homeomorphism between the sphere and the plane "plus one point" is pictured in Figure 2. Since the real line and the plane are not compact (they are not bounded) but the circle and sphere are, this exercise provides concrete geometric realizations of *one-point compactifications*. By a compactification Y of a topological space X, we mean a compact space that contains (a homeomorphic copy of) X as a dense subspace (that is, everything in Y is a limit of something in X). When $Y \setminus X$ is a singleton, we say that Y is a one-point compactification of X. The example of the sphere and the plane "plus a point" involving a stereographic projection is often presented before showing that every locally compact (that is, every point has a compact neighborhood, a property enjoyed in particular by the plane and the punctured plane) topological space admits a one-point compactification, which is unique up to homeomorphism.

Let us review this "exercise" more concretely:

The unit sphere $\mathbb{S}$ of $\mathbb{R}^3$ of equation $x^2 + y^2 + z^2 = 1$ without its "north pole" $N = (0, 0, 1)$ projects onto the "equatorial plane" $z = 0$ in the following way. To each point P of $\mathbb{S} \setminus \{N\}$, associate the point $f(P)$ of intersection of the plane $z = 0$ with the half-ray joining N to P, as described below and pictured in Figure 2.

The map $f : \mathbb{S} \setminus \{N\} \to \mathbb{R}^2$ is easily seen to be a homeomorphism. Indeed, if $P = (x, y, z)$ belongs to $\mathbb{S} \setminus \{N\}$, then $f(P)$ has coordinates $(\frac{x}{1-z}, \frac{y}{1-z})$ in $\mathbb{R}^2$ and if $M = (x, y)$ in $\mathbb{R}^2$, then $f^{-1}(M) = (\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{x^2+y^2-1}{1+x^2+y^2}) \in \mathbb{S}$. Thus, $\mathbb{R}^2$ can be identified with the subspace $\mathbb{S} \setminus \{N\}$ of the compact space $\mathbb{S}$ that is obtained from the plane by adding a single point corresponding to ∞: the north pole.

Let O denote the origin of the system of coordinates of $\mathbb{R}^3$, and of the equatorial plane $\mathbb{R}^2$ as well. Note that O is the image of the "south pole" $S = (0, 0, -1)$ under the stereographic projection f. Thus, if we consider the punctured plane $\mathbb{R}^2 \setminus \{O\}$ instead of the plane in the projection, the north pole corresponds to the compactifying point at ∞ while the south pole corresponds to the hole at O. In other words, $f : \mathbb{S} \setminus \{N, S\} \to \mathbb{R}^2 \setminus \{O\}$ is a homeomorphism from the sphere without its poles onto the punctured plane.

Thus, the one-point compactification of the punctured plane $\mathbb{R}^2 \setminus \{O\}$ is homeomorphic to the one-point compactification of $\mathbb{S} \setminus \{N, S\}$ which is easily seen to be the *pinched sphere* $\mathbb{S}^*$ obtained as the quotient space of $\mathbb{S}$ under the identification of the poles S and N, as pictured in Figure 3.

While it is intuitively clear that this realizes the desired quotient and thus the one point-compactification of $\mathbb{S} \setminus \{N, S\}$ (hence of the punctured plane), one may more

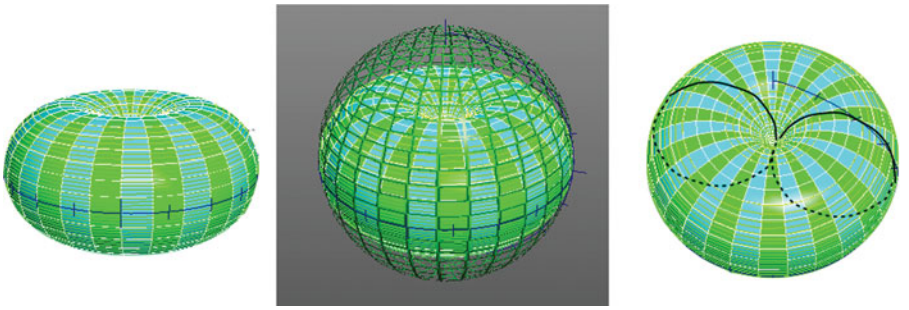


Figure 3 The pinched sphere $\mathbb{S}^*$.

concretely verify that $p : \mathbb{S} \to \mathbb{S}^*$ given in spherical coordinates by

$$p((1, \theta, \phi)) = (\sin\phi, \theta, \phi)$$

restricts to a homeomorphism from $\mathbb{S} \setminus \{N, S\}$ onto $\mathbb{S}^* \setminus \{O\}$ and that $p^{-1}(O) = \{S, N\}$.

As drawn on the rightmost picture of $\mathbb{S}^*$ in Figure 3, a figure eight with self-intersection at O can be drawn on $\mathbb{S}^*$. A single circle of this eight figure is a Jordan curve on $\mathbb{S}^*$ that does not disconnect $\mathbb{S}^*$. Thus, the plane and the punctured plane are not homeomorphic, because their one-point compactifications are not, for one is disconnected by Jordan curves while the other does not need to be.

REFERENCES

[1] Gamelin, T. W., Greene, R. E. (1983). *Introduction to Topology*. New York: Saunders.
[2] Lee, J. (2010). *Introduction to Topological Manifolds*. Vol. 940. New York: Springer Science & Business Media.
[3] Maehara, R. (1984). The Jordan curve theorem via the Brouwer fixed point theorem. *Am. Math. Monthly* 91(10):641–643.

Summary. After an informal short introduction to topology, I consider the problem of distinguishing the punctured plane from the plane topologically. I propose an alternative argument to the classical use of homotopy, relying instead on compactifications. Starting from the classical example of the stereographic projection realizing the sphere as the one-point compactification (a term that I explain) of the plane, I observe that the pinched sphere is the one-point compactification of the punctured plane. As the sphere is disconnected by a simple closed curve while the pinched sphere does not need to be, this provides an intuitive argument to distinguish topologically the plane from the punctured plane, without (explicitly) using homotopy.

FRÉDÉRIC MYNARD (MR Author ID: 662658, ORCID 0000-0002-1018-6748) received his Ph.D. in 1999 from the University of Burgundy (France) and is currently an associate professor of mathematics at New Jersey City University. His research focuses on general and categorical topology, and their applications to analysis. He is particularly interested in convergence spaces and their applications—the topic of a recent monograph of his, co-authored with Szymon Dolecki, which can be used as an alternative to an introductory topology textbook.

Of Puzzles and Partitions: Introducing Partiti

ANDRÉS EDUARDO CAICEDO
Mathematical Reviews
Ann Arbor, MI 48104
aec@ams.org

BRITTANY SHELTON
Albright College
Reading, PA 19612
bshelton@albright.edu

In each of the five issues for 2017, readers of this MAGAZINE found a PINEMI puzzle. Pinemi is the creation of Vietnamese puzzle enthusiast Thinh Van Duc Lai, who has also designed PARTITI, the puzzle that will run through this year's issues.

Partiti

Partiti is played on a 6×6 grid in which each cell contains a positive integer. To play, place one or more digits into each cell in such a way that the digits in a cell sum to the indicated positive integer and no digit appears more than once in a cell or between cells that are adjacent or share a corner.

The objective of the game can be described as finding unordered integer partitions of the given numbers consisting of distinct parts from $1, 2, \ldots, 9$ (subject to the additional restriction that the partitions for contiguous cells should use different parts). Such an integer partition of n consists of an increasing sequence of positive integers that sum to n. See Figure 1 for an example.

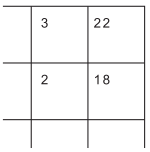


Figure 1 The top-right corner of this month's puzzle. We can solve some of it by noting that the only partition of 2 into distinct parts is 2 itself, and the only such partitions of 3 are $1 + 2$ and 3, but the former is excluded, since we cannot use 2 again. Though more information is needed to see what numbers go into the other two cells, the reader may want to note that $22 + 18 = 1 + 4 + 5 + 6 + 7 + 8 + 9$, the sum of the remaining digits, so all available numbers should be used between these two cells.

The puzzle for this month is at the end of this note. In what follows, we present some basic properties of integer partitions, take a very brief detour through partitions of infinite sets, and conclude with a few words about Partiti's creator Thinh Lai.

Integer partitions

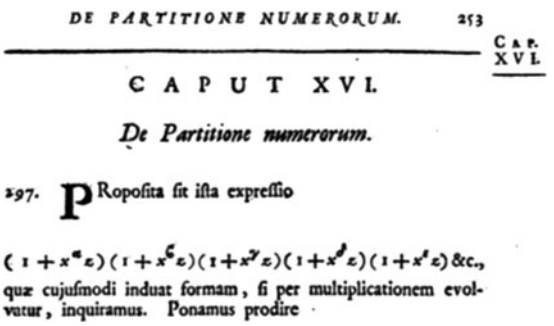
The study of partitions began with Euler. The number of integer partitions of n is often denoted $p(n)$. Hardy and Ramanujan worked out an analytic formula for $p(n)$; the

Math. Mag. **91** (2018) 20–23. doi:10.1080/0025570X.2018.1403233
MSC: Primary 00A08, Secondary 11P57

formula takes the form of an infinite series, and even just a few terms produce remarkably accurate approximations. As n increases, $p(n)$ grows faster than a polynomial, but slower than any exponential a^n, $a > 1$. More specifically, $p(n) \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{2n/3}}$, meaning that the quotient of these two expressions approaches one as n approaches infinity. However, puzzlers need not worry, since the possible positive integers in the cells of Partiti are quite modest, the largest potential entry in a cell being $9 + 8 + 7 + 6 + 5 + 4 = 39$, which could only occur in a corner surrounded by 1, 2, and 3.

More relevant than p in this context is the number of partitions of n into distinct parts, usually denoted $q(n)$. For example, $5 = 1 + 4 = 2 + 3$ are all the partitions of 5 into distinct parts, and therefore $q(5) = 3$. By convention, $q(0) = 1$. The function q has a somewhat more modest rate of growth than that of p, namely, $q(n) \sim \frac{1}{4\sqrt[4]{3n^3}} e^{\pi\sqrt{n/3}}$. The sequence $q(0), q(1), \ldots$ is sequence A000009 in the OEIS [3].

The first nontrivial result on this function q is Euler's theorem from 1748 [2] giving us that $q(n)$ coincides with the number of partitions of n into (not necessarily distinct) odd parts. For example, $5 = 1 + 1 + 3 = 1 + 1 + 1 + 1 + 1$ are all such partitions of 5 and, as predicted by the theorem, there are precisely 3 of them. See Figure 2 for the beginning of Euler's work on integer partitions.


DE PARTITIONE NUMERORUM. 253

CAP. XVI.

CAPUT XVI.

De Partitione numerorum.

297. Proposita sit ista expressio

$(1+x^{\alpha}z)(1+x^{\beta}z)(1+x^{\gamma}z)(1+x^{\delta}z)(1+x^{\varepsilon}z)$ &c.,

quae cujusmodi induat formam, si per multiplicationem evolvatur, inquiramus. Ponamus prodire

Figure 2 The opening of chapter 16 of Euler's *Introductio in Analysin Infinitorum* [2]. The book lays the foundations of mathematical analysis. It also introduces the theory of integer partitions, in this chapter.

There are several proofs of Euler's theorem. The one we proceed to sketch uses generating functions. It relies on observing that $\sum_{n=0}^{\infty} q(n)x^n$ can be represented as $\prod_{n=1}^{\infty}(1 + x^n)$: At least formally, this product can be expanded by picking from each factor $1 + x^n$ one of the two summands, with the understanding that in each product, all but finitely many times the summand 1 is the chosen one. Grouping together like terms, the coefficient of x^n in this expansion counts the number of ways the exponent n can be formed as a sum of distinct terms, which is precisely $q(n)$. For example, note that the only products that result in an x^5 term are $1 \cdot 1 \cdot 1 \cdot 1 \cdot x^5 = x \cdot 1 \cdot 1 \cdot x^4 = 1 \cdot x^2 \cdot x^3$, where in each product we have omitted the infinitely many remaining 1s.

Now, note that $\prod_{n=1}^{\infty}(1 + x^n) = \prod_{n=1}^{\infty} \frac{1-x^{2n}}{1-x^n} = \prod_{n=1}^{\infty} \frac{1}{1-x^{2n-1}}$, the latter equality holding because all numerators cancel out and the only denominators that survive are the ones with odd degree. Expanding this product reveals that it is the generating function for partitions into odd parts:

$$\begin{aligned}
\prod_{n=1}^{\infty} \frac{1}{1-x^{2n-1}} &= \prod_{n=1}^{\infty}(1 + x^{2n-1} + x^{2(2n-1)} + x^{3(2n-1)} + \cdots) \\
&= (1 + x + x^{2\cdot 1} + \cdots)(1 + x^3 + x^{2\cdot 3} + \cdots)(1 + x^5 + x^{2\cdot 5} + \cdots)\cdots \\
&= (1 + x + x^1x^1 + \cdots)(1 + x^3 + x^3x^3 + \cdots)(1 + x^5 + x^5x^5 + \cdots)\cdots \\
&= 1 + x + x^1x^1 + (x^1x^1x^1 + x^3) + (x^1x^1x^1x^1 + x^1x^3) + \cdots.
\end{aligned}$$

The argument above can be readily formalized either in terms of formal power series expansions or in terms of "genuine" power series (upon arguing that the series involved converge for $|x| < 1$).

Many other interesting results are known for q and other partition functions, see [1] for an introduction. These results are established by a wide variety of techniques, including combinatorial counting arguments, Ferrers diagrams, and others, and some are quite sophisticated, involving detailed analytical arguments, which entered the picture thanks to the joint work of Hardy and Ramanujan at the beginning of the twentieth century.

A small detour

The first-named author cannot help but mention that some of his own work involves the study of partitions, in this case partitions of infinite sets. This is part of the area of set theory called the partition calculus. As a simple example of the sort of problems one considers here, readers familiar with the distinction between countable and uncountable sets may enjoy verifying the following: Suppose the set $\mathbb{R}$ of reals is partitioned into countably many pieces, $\mathbb{R} = \bigcup_{n=1}^{\infty} A_n$. Then at least one of the sets A_n contains an infinite increasing sequence. Note the result fails for $\mathbb{Q}$ in place of $\mathbb{R}$ (we can split $\mathbb{Q}$ into countably many singletons). On the other hand, the result is not simply an artifact of $\mathbb{R}$ being uncountable (which ensures that one of the A_n is also uncountable), since not every uncountable ordered set contains an infinite increasing sequence.

About Partiti's creator

We hope readers enjoy Partiti and the many other puzzles we anticipate seeing from Thinh. They can learn more about Thinh himself in a recent piece on his work by Will Shortz that ran in The New York Times [4]. Thinh's puzzle Bar Code appeared for 14 weeks in the Sunday Magazine section of The Times.

Thinh has shown us a large variety of different puzzles of his own creation, all of a somewhat mathematical flavor. We asked him a few questions in preparation for this note. He shared with us that he solves all his puzzles manually, and uses the time it takes him to estimate their difficulty. He has surprised himself a few times with how hard some of his creations turned out to be. Although the name "Partiti" is probably self-explanatory, most of his puzzle names are inspired by Japanese puzzles, of which he confesses to be a big fan.

Thinh advises readers interested in creating their own mathematical puzzles that it is necessary to build a basic foundation and to read about numbers and logic puzzles. It took him five years to acquire this foundation himself. He indicates that some of the examples he has submitted to this MAGAZINE are harder than the ones he has had featured in The New York Times, and hopes to publish a book of his own puzzles.

Acknowledgment We thank Thinh Lai for his enthusiasm and help with the preparation of this note.

REFERENCES

[1] Andrews, G., Eriksson, K. *Integer Partitions*. Cambridge University Press, Cambridge, 2004.

[2] Euler, L. (1748). *Introductio in analysin infinitorum*. Apud Marcum-Michaelem Bousquet & socios, Lausanne. archive.org/details/bub_gb_jQ1bAAAAQAAJ.

[3] Sequence A000009. The on-Line encyclopedia of integer sequences. oeis.org/A000009.

[4] Shortz, W. Bar Code: a new infatuation poised for a puzzle craze. *The New York Times*, 2017, June 12. nyti.ms/2sfDb1f.

Summary. We introduce PARTITI, the puzzle that will run in this MAGAZINE this year, and use the opportunity to recall some basic properties of integer partitions.

ANDRÉS EDUARDO CAICEDO (MR Author ID: 684109) earned his B.Sc. in Mathematics from Universidad de los Andes, in Bogotá, Colombia, in 1996, and his Ph.D. from the University of California, Berkeley, in 2003. His research centers on set theory. From 2003–2005, he was a research assistant at the Kurt Gödel Research Center for Mathematical Logic of the University of Vienna. From 2005–2008 he was the Harry Bateman Research Instructor at the California Institute of Technology. He then joined the department of mathematics at Boise State University, where he remained until moving to Ann Arbor in 2015 as an associate editor at Mathematical Reviews. Andrés is married and has two kids. He enjoys comic books, Sudoku, and Candy Crush.

BRITTANY SHELTON (MR Author ID: 960329) earned her B.Sc. from Montclair State University, in 2007, and her Ph.D. from Lehigh University, in 2013. She is an assistant professor of mathematics at Albright College, Reading, PA. Her research interests include algebraic and enumerative combinatorics. She never gives up an opportunity to combine two of her passions: mathematics and puzzles.

PARTITI PUZZLE

17	3	10	19	3	22
8	17	9	4	2	18
9	2	16	1	12	13
17	10	4	8	18	2
15	3	5	2	5	12
23	4	16	18	20	8

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

The solution is on page 15.

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

How Franklin (May Have) Made His Squares

RONALD P. NORDGREN
Rice University
Houston, TX 77251
nordgren@rice.edu

Benjamin Franklin is well-known as a patriot, statesman, diplomat, writer, and scientist. In addition, he had a keen interest in numbers as described in great detail by Pasles [**13–15**]. In particular, Franklin was interested in the construction of magic squares and circles. He spent considerable effort on this endeavor as a young man (perhaps while clerking in the Pennsylvania legislature) and later in life. However, the motivation for his interest remains somewhat of a mystery. In a letter to an English mathematician he answered the question of the usefulness of his endeavors as follows [**15**, p. 125–126]:

> *Perhaps the considering and answering of such questions may not be altogether useless if it produces by practice an habitual readiness and exactness in mathematical disquisitions, which readiness may, on many occasions, be of real use. In the same way, may the making of these squares be of use.*

After stating that the construction of ordinary magic squares was easy, he added:

> *I had imposed on myself more difficult tasks, and succeeded in making other magic squares, with a variety of properties, and much more curious.*

Indeed, Franklin imposed complex requirements on his squares that differ from those of the usual magic squares. His special magic squares and their construction remain of interest to the present day. Unfortunately, Franklin did not disclose his methods of construction and various methods have been proposed. One method of constructing Franklin squares is analyzed here and other methods are discussed. We begin with definitions of various types of magic squares.

All rows and columns of a *semi-magic* square matrix must sum to an *index number* m. If its main diagonal and the cross diagonal also sum to m, then the square is *magic*. *Natural* $n \times n$ (*order* n) magic and semi-magic squares have elements $1, 2, \ldots, n^2$ for which

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

Magic squares originated in China about 2000 years ago according to Cammann [**4**]. They have received considerable attention in the mathematical literature over the years as detailed in books by Pickover [**16**] and Pasles [**15**, Chapter 2].

However, Franklin did not care for the two diagonal sum conditions on magic squares. Instead, he defined four families of *bent diagonals* which must sum to m. In the following families of order-6 squares, elements on the six bent diagonals (with wraparound) have the same symbol, see Figure 1. These four families are named for the direction of their "V" for future reference. Also, Franklin required that the elements on all left and right half-rows and all upper and lower half-columns of his squares must sum to $m/2$ which makes his squares semi-magic. This requirement forces natural Franklin

Math. Mag. **91** (2018) 24–32. doi:10.1080/0025570X.2017.1404789
MSC: Primary 05B15, Secondary 11C20, 00A08

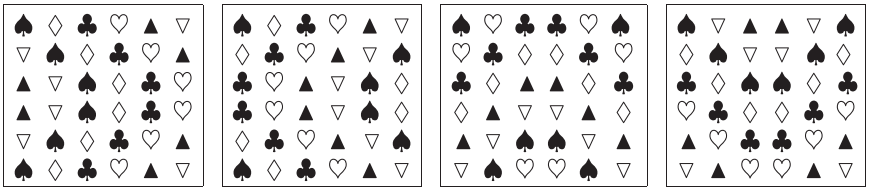


Figure 1 Families of bent diagonals.

squares to be of doubly even order ($n = 4k$). In addition, he required that elements in all 2×2 subsquares (including broken ones) of his order-n square sum to $4m/n$.

In 1769, Ben Franklin published an order-8 and an order-16 square that met his three sum conditions. Here is his order-8 square [15]:

$$F_8 = \begin{bmatrix} 52 & 61 & 4 & 13 & 20 & 29 & 36 & 45 \\ 14 & 3 & 62 & 51 & 46 & 35 & 30 & 19 \\ 53 & 60 & 5 & 12 & 21 & 28 & 37 & 44 \\ 11 & 6 & 59 & 54 & 43 & 38 & 27 & 22 \\ 55 & 58 & 7 & 10 & 23 & 26 & 39 & 42 \\ 9 & 8 & 57 & 56 & 41 & 40 & 25 & 24 \\ 50 & 63 & 2 & 15 & 18 & 31 & 34 & 47 \\ 16 & 1 & 64 & 49 & 48 & 33 & 32 & 17 \end{bmatrix}. \tag{2}$$

His published order-16 square is given by Pasles [13,15], Jacobs [7], Morris [8], and Nordgren [11]. Several other Franklin squares have been found among his papers as reported in [13–15].

Franklin's method of constructing his squares has been the subject of considerable speculation over the years since he did not indicate it. His two published squares can be constructed in various ways. In 1776, Euler formed a magic square by linear combination of two auxiliary squares. According to Pasles [15], this same composition method was used by Youle in 1813 to construct Franklin's two published squares and likely by Franklin himself. However, it is not clear how Franklin formed his two auxiliary squares, if indeed he used this method.

Another early effort, by Carus in [2], gives a direct construction of order-8 and order-16 Franklin squares that differ from those published by Franklin. A somewhat simpler direct construction by Jacobs [7] gives Franklin's two published squares, one of order 24, and supposedly those of order $8k$. Jacobs' direct method involves five prescribed steps for sequential placement of the integers in a natural Franklin square of order $8k$. His approach appears to arise from a generalization of the element pattern in Franklin's published order-8 and order-16 squares but a proof of its general applicability is lacking. Also, an order-32 Franklin square is presented by Behrforooz [3] without an indication of its method of construction, although it can be constructed by Jacobs' method.

Several direct methods of constructing Franklin squares are posted by Hurkens [6] who shows by exhaustive search that no natural order-12 Franklin squares exist. Also, Pasles [14] shows that there are no natural order-4 Franklin squares. Numerical generation of all natural order-8 Franklin squares is carried out by Schindel et al. [17]. A method of constructing nonnatural Franklin squares from Hilbert bases is given by Ahmed [1].

Here, we employ Euler's composition method to systematically construct Franklin squares of order $n = 8k$. We obtain two types of formulas for the elements of our two auxiliary squares that allow straightforward numerical formation of Franklin square

matrices of order $8k$. From these formulas, we verify that our auxiliary squares are suitable for construction of Franklin squares. Since our construction method leads to Franklin's published squares for $k = 1, 2$, it may be regarded as an extension of his methodology and may have been used by him.[1] This speculation is supported by an examination of the Eulerian composition of three order-8 Franklin squares by Pasles [15] and two order-16 Franklin squares by Morris [8] and Nordgren [11]. The two auxiliary squares for these squares follow regular patterns and it seems difficult to construct some of them by a direct method. The reader is invited to try!

Furthermore, our composite construction method produces the order-24 Franklin square of Jacobs [7] (shown in [11] with auxiliary squares) and the order-32 Franklin square of Behrforooz [3]. Although this correspondence lends support for Jacobs' direct method, an explicit connection between the two construction methods has not been found.

In a *pandiagonal square*, elements on all diagonals in both directions (including broken ones) sum to m. Some Franklin squares also are pandiagonal, including one-third of the order-8 natural ones [17]. In what follows, we show that a Franklin square can be transformed to a pandiagonal magic square in two ways but the converse is not true in general. Also, Nordgren [10] shows that order-$8k$ pandiagonal Franklin magic squares can be obtained from transformation of complete (or most-perfect) magic squares that are constructed and enumerated by Ollerenshaw and Brée [12].

Franklin square construction

The three sum conditions for a Franklin square are stated in the above introduction. In what follows, Franklin squares are treated as square matrices. For a Franklin matrix F_n of order-n, Euler's composition formula [5] can be written as

$$F_n = nQ_n + R_n + U_n, \tag{3}$$

where Q_n and R_n are order-n orthogonal matrices and U_n is the *unity matrix* with all elements 1. *Orthogonal matrices* are defined as having each ordered pair of elements in the same position in the two matrices occurring once and only once. A natural Franklin matrix F_n can be constructed by requiring that the *quotient matrix* Q_n and the *remainder matrix* R_n have elements $0, 1, \ldots, n-1$ repeated n times. If such Q_n and R_n are orthogonal and satisfy the three Franklin sum conditions with m replaced by $\hat{m} = n(n-1)/2$, then F_n forms a Franklin matrix. However, these conditions may not be necessary ones (see Nordgren [11] for counterexamples of a magic and a semi-magic square).

Order-8 square Franklin's order-8 matrix F_8 of (2) is obtained from (3) with

$$Q_8 = \begin{bmatrix} 6 & 7 & 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 7 & 6 & 5 & 4 & 3 & 2 \\ 6 & 7 & 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 7 & 6 & 5 & 4 & 3 & 2 \\ 6 & 7 & 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 7 & 6 & 5 & 4 & 3 & 2 \\ 6 & 7 & 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 7 & 6 & 5 & 4 & 3 & 2 \end{bmatrix}, \quad R_8 = \begin{bmatrix} 3 & 4 & 3 & 4 & 3 & 4 & 3 & 4 \\ 5 & 2 & 5 & 2 & 5 & 2 & 5 & 2 \\ 4 & 3 & 4 & 3 & 4 & 3 & 4 & 3 \\ 2 & 5 & 2 & 5 & 2 & 5 & 2 & 5 \\ 6 & 1 & 6 & 1 & 6 & 1 & 6 & 1 \\ 0 & 7 & 0 & 7 & 0 & 7 & 0 & 7 \\ 1 & 6 & 1 & 6 & 1 & 6 & 1 & 6 \\ 7 & 0 & 7 & 0 & 7 & 0 & 7 & 0 \end{bmatrix}. \tag{4}$$

[1] On the other hand, one referee states that this seems unlikely and another referee states that Jacobs' construction seems like something Franklin would construct.

The auxiliary matrices Q_8 and R_8 can be constructed in a systematic way that can be extended to higher-order matrices $n = 8k$. The quotient matrix Q_8 starts with 0 in row 1, column 3 and the entries continue sequentially with wraparound. Each element of the second row is the 7-complement of the element of the first row in the same column and the remaining rows are alternating copies of the first two rows. The remainder matrix R_8 can be constructed by first forming its main diagonal by sequentially positioning its elements starting with 0 in the lower right corner, 1 next upward, skipping 4 rows, then 2,3 upward, and finally 4, 5, 6, 7 downward. The rows are filled out with alternating 7-complements of its diagonal element and the diagonal element itself. The construction of the first row of Q_8 and the diagonal of R_8 are the crucial elements of our construction method. Next, we show that the construction of Q and R can be generalized to order $n = 8k$ and give a Franklin matrix from (3).

Higher-order squares The foregoing constructions for Q_8 and R_8 can be generalized to construct Franklin matrices of order $n = 8k$ by the following procedure:

1. Form the quotient matrix Q_n by entering 0 in column $n/4 + 1$ of row 1 and entering integers sequentially in the remaining columns with wraparound. Form the second row by taking the $n - 1$ complement of the entries in the same column of the first row. Form the remaining rows by alternating copies of the first two rows.
2. To form the remainder matrix R_n, first form its main diagonal by entering 0 in the lower-right corner and entering integers sequentially upward for $n/4$ elements. Skip to row $n/4$ and continue sequential entries upward to the upper-left corner. Continue sequential entries from row $n/4 + 1$ downward to fill out the diagonal. Take the $n - 1$ complement of the diagonal elements and alternate them with the diagonal elements to fill out the rows of R_n.
3. Combine Q_n, R_n, and U_n according to (3) to obtain a Franklin matrix F_n.

Formulas for the elements of Q_n and R_n can be expressed in two ways and used to verify that they are orthogonal and satisfy Franklin's three sum conditions.

Element formulas I In the first way, the elements of the quotient matrix Q_n can be written as

Q_n	**1**	**2**	...	$\hat{\mathbf{n}}-\mathbf{1}$	$\hat{\mathbf{n}}$	$\hat{\mathbf{n}}+\mathbf{1}$	$\hat{\mathbf{n}}+\mathbf{2}$	...	$\mathbf{n}-\mathbf{1}$	**n**
1	$3\hat{n}$	$3\hat{n}+1$	...	$n-2$	$n-1$	0	1	...	$3\hat{n}-2$	$3\hat{n}-1$
2	$\hat{n}-1$	$\hat{n}-2$	...	1	0	$n-1$	$n-2$	...	$\hat{n}+1$	$\hat{n}$
3	$3\hat{n}$	$3\hat{n}+1$	...	$n-2$	$n-1$	0	1	...	$3\hat{n}-2$	$3\hat{n}-1$
4	$\hat{n}-1$	$\hat{n}-2$	...	1	0	$n-1$	$n-2$	...	$\hat{n}+1$	$\hat{n}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\mathbf{n}-\mathbf{1}$	$3\hat{n}$	$3\hat{n}+1$	...	$n-2$	$n-1$	0	1	...	$3\hat{n}-2$	$3\hat{n}-1$
n	$\hat{n}-1$	$\hat{n}-2$	...	1	0	$n-1$	$n-2$	...	$\hat{n}+1$	$\hat{n}$

,

where $\hat{n} = n/4 = 2k$ and row/column numbers are in bold type. In view of the complementarity of odd and even row entries, it is easily seen that the half-columns sum to $n(n-1)/4 = \hat{m}/2$ as required. In the first row, the first element $3\hat{n}$ can be paired with the element in column $n/2$, namely, $\hat{n} - 1$ and the second element $3\hat{n} + 1$ with $\hat{n} - 2$, etc. Since there are $n/4$ such pairs that add to $n - 1$, the left half of the first row sums to $\hat{m}/2$ as required. Similarly for the right half of the first row. The proper half-row sums for the second row then follow from complementarity. The V-Right and V-Left

bent diagonals (Figure 1) sum to $n(n-1)/2 = \hat{m}$ due to complementarity. For $n = 8k$ (essential), the elements of the first V-Down bent diagonal can be summed pair-wise on the rows giving

$$\frac{n}{4}(3\hat{n} + 3\hat{n} - 1) + \frac{n}{4}(\hat{n} - 2 + \hat{n} + 1) = \frac{n}{2}(n-1) = \hat{m}. \tag{5}$$

The second V-Down bent diagonal sums to $\hat{m}$ in view of complementarity and the remaining V-Down bent diagonals are identical to the first two. The V-Up bent diagonals have the same elements as the V-Down bent diagonals and therefore they also sum to $\hat{m}$ for $n = 8k$. Furthermore, the sum of the elements in all 2 by 2 subsquares of Q_n is seen to be $2(n-1) = 4\hat{m}/n$ as required. Thus, the quotient matrix Q_n satisfies the Franklin sum conditions for $n = 8k$.

The remainder matrix can be written as

$$R_n = \begin{bmatrix} R_n^{(11)} & R_n^{(12)} \\ R_n^{(21)} & R_n^{(22)} \end{bmatrix}, \tag{6}$$

where $R_n^{(11)}$ and $R_n^{(22)}$ are $n/2$ by $n/2$ submatrices with elements given by

$R_n^{(11)}$	**1**	**2**	...	$\hat{\mathbf{n}}$	$\hat{\mathbf{n}}+\mathbf{1}$	$\hat{\mathbf{n}}+\mathbf{2}$	...	$\mathbf{2\hat{n}}$
1	$\mathbf{2\hat{n}-1}$	$2\hat{n}$	...	$2\hat{n}$	$2\hat{n}-1$	$2\hat{n}$	...	$2\hat{n}$
2	$2\hat{n}+1$	$\mathbf{2\hat{n}-2}$	...	$2\hat{n}-2$	$2\hat{n}+1$	$2\hat{n}-2$	...	$2\hat{n}-2$
⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋮
$\hat{\mathbf{n}}$	$3\hat{n}-1$	$\hat{n}$	...	$\hat{\mathbf{n}}$	$3\hat{n}-1$	$\hat{n}$	...	$\hat{n}$
$\hat{\mathbf{n}}+\mathbf{1}$	$2\hat{n}$	$2\hat{n}-1$	...	$2\hat{n}-1$	$\mathbf{2\hat{n}}$	$2\hat{n}-1$	...	$2\hat{n}-1$
$\hat{\mathbf{n}}+\mathbf{2}$	$2\hat{n}-2$	$2\hat{n}+1$	...	$2\hat{n}+1$	$2\hat{n}-2$	$\mathbf{2\hat{n}+1}$	...	$2\hat{n}+1$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
$\mathbf{2\hat{n}}$	$\hat{n}$	$3\hat{n}-1$	...	$3\hat{n}-1$	$\hat{n}$	$3\hat{n}-1$	...	$\mathbf{3\hat{n}-1}$

$R_n^{(22)}$	**1**	**2**	...	$\hat{\mathbf{n}}$	$\hat{\mathbf{n}}+\mathbf{1}$	$\hat{\mathbf{n}}+\mathbf{2}$	...	$\mathbf{2\hat{n}}$
1	$\mathbf{3\hat{n}}$	$\hat{n}-1$	...	$\hat{n}-1$	$3\hat{n}$	$\hat{n}-1$	...	$\hat{n}-1$
2	$\hat{n}-2$	$\mathbf{3\hat{n}+1}$	...	$3\hat{n}+1$	$\hat{n}-2$	$3\hat{n}+1$	...	$3\hat{n}+1$
⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋮
$\hat{\mathbf{n}}$	0	$n-1$	...	$\mathbf{n-1}$	0	$n-1$	...	$n-1$
$\hat{\mathbf{n}}+\mathbf{1}$	$\hat{n}-1$	$3\hat{n}$	...	$3\hat{n}$	$\hat{\mathbf{n}}-\mathbf{1}$	$3\hat{n}$	...	$3\hat{n}$
$\hat{\mathbf{n}}+\mathbf{2}$	$3\hat{n}+1$	$\hat{n}-2$	...	$\hat{n}-2$	$3\hat{n}+1$	$\hat{\mathbf{n}}-\mathbf{2}$	...	$\hat{n}-2$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
$\mathbf{2\hat{n}}$	$n-1$	0	...	0	$n-1$	0	...	**0**

with diagonal elements and row/column numbers in bold type. Also, $R_n^{(12)} = R_n^{(11)}$ and $R_n^{(21)} = R_n^{(22)}$ as is clear from their construction and is exhibited by R_8 of (4). Since the right and left half-rows of R_n consists of $n/4$ alternating complements, the two half rows sum to $n(n-1)/4 = \hat{m}/2$ as required. In view of its construction, the main diagonal entries consist of $n/2$ pairs of complements on its upper and lower half. Therefore, row entry complementarity forces the half columns also to consist of $n/2$ pairs of complements that sum to $n(n-1)/4 = \hat{m}/2$ for $n = 8k$ (essential). In view of the

diagonal construction and row-entry complementarity, the V-Right and V-Left bent diagonals contain the integers $0, 1, \ldots n-1$ that sum to $n(n-1)/2 = \hat{m}$ as required. The V-Down and V-Up bent diagonals consist of $n/2$ pairs of complements that sum to $\hat{m}$. Furthermore, the sum of the elements in all 2 by 2 subsquares of R_n is seen to be $2(n-1) = 4\hat{m}/n$ as required. Thus, R_n satisfies the Franklin sum conditions for $n = 8k$. Since Q_n and R_n satisfy the Franklin sum conditions, so does F_n from (3).

The orthogonality of the matrices Q_n and R_n can be established by considering pairs of their rows. For example, row 1 of R_n has elements $2\hat{n}-1, 2\hat{n}, 2\hat{n}-1, \ldots, 2\hat{n}$ and row $\hat{n}+1$ of R_n has elements $2\hat{n}, 2\hat{n}-1, 2\hat{n}, \ldots, 2\hat{n}-1$. Row 1 and row $\hat{n}+1$ of Q_n contain elements $0, 1, \ldots, n-1$ in the same order. Thus, the pairs from row 1 and row $\hat{n}+1$ of Q_n and R_n form all relevant pairs ending in $2\hat{n}$ and $2\hat{n}-1$. Continuation of this row pairing process leads to the conclusion that all relevant ordered pairs of Q_n and R_n occur once and only once, *i.e.*, Q_n and R_n are orthogonal as required. Thus, our composition construction method produces an order-$8k$ natural Franklin matrix.

Element formulas II The elements of the quotient and remainder matrices Q_n and R_n ($n = 8k$), constructed as indicated above, are given by

$$Q_n(i,j) = \frac{n-1}{2}[1+(-1)^i] - (-1)^i\left[j+\frac{3n}{4}-1\right] \bmod n,$$

$$R_n(i,j) = \begin{cases} \left(\frac{n}{2}-i\right)(-1)^{i+j} + \frac{n-1}{2}[1-(-1)^{i+j}], & 1 \le i \le \frac{n}{4}, \\ \left(i+\frac{n}{4}-1\right)(-1)^{i+j} + \frac{n-1}{2}[1-(-1)^{i+j}], & \frac{n}{4} < i \le \frac{3n}{4}, \\ (n-i)(-1)^{i+j} + \frac{n-1}{2}[1-(-1)^{i+j}], & \frac{3n}{4} < i \le n. \end{cases} \tag{7}$$

These formulas and (3) can be implemented in Maple®, MATLAB®, and Excel® to generate numerical Franklin squares of order $n = 8k$. The Heaviside step function can be used to enable the formula for R_n.

The elements Q_n and R_n given by (7) can be shown to satisfy the Franklin sum conditions on half-rows, half-columns, and the four families of bent diagonals. For example, the V-Right bent diagonals of Q_n have the sum

$$\sum_{i=1}^{n/2}\left[Q_n(i, i+\ell) + Q_n\left(i+\frac{n}{2}, \frac{n}{2}-i+1+\ell\right)\right] = \frac{n}{2}(n-1), \quad (\ell = 1, 2, \ldots, n).$$

Also, from (7) it can be shown that all 2 by 2 submatrices of Q_n and R_n sum to $2(n-1)$, *i.e.*,

$$Q_n(i,j) + Q_n(i,j+1) + Q_n(i+1,j) + Q_n(i+1,j+1) = 2(n-1),$$

and similarly for R_n. Thus, by (3), all 2 by 2 submatrices of a Franklin matrix F_n ($n = 8k$) sum to $2(n^2+1)$.

Transformation to pandigonal magic squares

We consider transformation of Franklin matrices to pandiagonal magic matrices (defined in the introduction). Following Nordgren [**9**], we define the order-n *shifter matrix* K_n as having elements $K_n(1,n) = 1$, $K_n(i, i-1) = 1$ ($i = 2, 3, \ldots n$), and all other elements 0. We define the order-n *reflection matrix* $\mathcal{R}_n$ as having elements 1 on the cross

diagonal and all other elements 0. For example,

$$K_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathcal{R}_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{8}$$

The matrix product K_nM_n shifts the elements of a matrix M_n down one row (bottom row to top) and M_nK_n shifts them one column left (first column to last). Matrix powers of K_n enable multiple row/column shifts. The matrix product $\mathcal{R}_nM_n$ reflects the elements of M_n about its horizontal axis, $M_n\mathcal{R}_n$ reflects the elements of M_n about its vertical axis, and $\mathcal{R}_nM_n\mathcal{R}_n$ rotates the elements of M 180°. Also, I_n is the order-n *identity matrix* and O_n is the order-n *zero matrix* with all elements 0.

As noted in [**9**], the pandiagonal sum property leads to

$$\sum_{i=1}^{n} [K_n]^i P_n [K_n]^i = mU_n \quad \text{and} \quad \sum_{i=1}^{n} [K_n]^{-i} P_n [K_n]^i = mU_n, \tag{9}$$

which provide a means of checking this property.

An order $n = 8k$ Franklin square F_n can be transformed to a pandiagonal magic square P'_n by permuting the columns in its lower half by means of the matrix formula

$$P'_n = \begin{bmatrix} I_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix} F_n + \begin{bmatrix} O_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & \mathcal{R}_{\tilde{n}} \end{bmatrix} F_n \begin{bmatrix} O_{\tilde{n}} & I_{\tilde{n}} \\ I_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix}, \tag{10}$$

where

$$P'_n = \begin{bmatrix} P'^{(11)}_{\tilde{n}} & P'^{(12)}_{\tilde{n}} \\ P'^{(21)}_{\tilde{n}} & P'^{(22)}_{\tilde{n}} \end{bmatrix}, \quad F_n = \begin{bmatrix} F^{(11)}_{\tilde{n}} & F^{(12)}_{\tilde{n}} \\ F^{(21)}_{\tilde{n}} & F^{(22)}_{\tilde{n}} \end{bmatrix},$$

$\tilde{n} = n/2 = 4k$, and $P'^{(ij)}_{\tilde{n}}$, $F^{(ij)}_{\tilde{n}}$, $I_{\tilde{n}}$, $O_{\tilde{n}}$, $\mathcal{R}_{\tilde{n}}$ are $\tilde{n} \times \tilde{n}$ submatrices, e.g. for $n = 4$

$$\begin{bmatrix} I_2 & O_2 \\ O_2 & O_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} O_2 & O_2 \\ O_2 & \mathcal{R}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{11}$$

Note that normal matrix multiplication applies in (10). The V-Right and V-Left bent diagonals of F_n in (Figure 1) become the Down-Right and Down-Left broken diagonals of P'_n, respectively. Similarly, F_n can be transformed to a pandiagonal matrix P''_n by permuting the rows in its right half by means of the matrix formula

$$P''_n = F_n \begin{bmatrix} I_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix} + \begin{bmatrix} O_{\tilde{n}} & I_{\tilde{n}} \\ I_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix} F_n \begin{bmatrix} O_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & \mathcal{R}_{\tilde{n}} \end{bmatrix}. \tag{12}$$

The V-Down and V-Up bent diagonals of F_n in Figure 1 become the Down-Right and Down-Left broken diagonals of P''_n, respectively. The magic row/column sum conditions are satisfied by P'_n and P''_n since F_n satisfies half-column and half-row sum conditions that are unchanged by the transformations. Thus, P'_n and P''_n are pandiagonal magic squares. The transformations (10) and (12) provide a means of verifying the four Franklin bent-diagonal sum conditions on F_n by using (9) to verify that P'_n and P''_n are pandiagonal. Nordgren [**10**] gives simpler matrix formulae for the three Franklin sum conditions.

For example, application of (10) and (12) to the Franklin square (2) gives the pandiagonal magic squares

$$P_8' = \begin{bmatrix} 52 & 61 & 4 & 13 & 20 & 29 & 36 & 45 \\ 14 & 3 & 62 & 51 & 46 & 35 & 30 & 19 \\ 53 & 60 & 5 & 12 & 21 & 28 & 37 & 44 \\ 11 & 6 & 59 & 54 & 43 & 38 & 27 & 22 \\ 48 & 33 & 32 & 17 & 16 & 1 & 64 & 49 \\ 18 & 31 & 34 & 47 & 50 & 63 & 2 & 15 \\ 41 & 40 & 25 & 24 & 9 & 8 & 57 & 56 \\ 23 & 26 & 39 & 42 & 55 & 58 & 7 & 10 \end{bmatrix},$$

$$P_8'' = \begin{bmatrix} 52 & 61 & 4 & 13 & 42 & 39 & 26 & 23 \\ 14 & 3 & 62 & 51 & 24 & 25 & 40 & 41 \\ 53 & 60 & 5 & 12 & 47 & 34 & 31 & 18 \\ 11 & 6 & 59 & 54 & 17 & 32 & 33 & 48 \\ 55 & 58 & 7 & 10 & 45 & 36 & 29 & 20 \\ 9 & 8 & 57 & 56 & 19 & 30 & 35 & 46 \\ 50 & 63 & 2 & 15 & 44 & 37 & 28 & 21 \\ 16 & 1 & 64 & 49 & 22 & 27 & 38 & 43 \end{bmatrix},$$

which can be checked for pandiagonality using (9).

Given a pandiagonal magic square P_n, the inverse transformations from (10) and (12) give

$$\begin{aligned} M_n' &= \begin{bmatrix} I_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix} P_n + \begin{bmatrix} O_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & \mathcal{R}_{\tilde{n}} \end{bmatrix} P_n \begin{bmatrix} O_{\tilde{n}} & I_{\tilde{n}} \\ I_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix}, \\ M_n'' &= P_n \begin{bmatrix} I_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix} + \begin{bmatrix} O_{\tilde{n}} & I_{\tilde{n}} \\ I_{\tilde{n}} & O_{\tilde{n}} \end{bmatrix} P_n \begin{bmatrix} O_{\tilde{n}} & O_{\tilde{n}} \\ O_{\tilde{n}} & \mathcal{R}_{\tilde{n}} \end{bmatrix}, \end{aligned} \tag{13}$$

where M_n' and M_n'' are not Franklin matrices in general since M_n' satisfies the V-Right and V-Left bent-diagonal sum conditions whereas M_n'' satisfies the V-Down and V-Up bent-diagonal sum conditions. Furthermore, M_n' and M_n'' are not likely to satisfy the other two Franklin sum conditions. Therefore, a pandiagonal magic matrix cannot be transformed to a Franklin matrix in general. However, as noted earlier, Schindel et al. [**17**] found that one-third of the order-8 natural Franklin squares are pandiagonal.

REFERENCES

[1] Ahmed, M. (2004). How many squares are there, Mr Franklin? *Am. Math. Monthly* 111:394–410.

[2] Andrews, W. (1908). Magic Squares and Cubes. Chicago: The Open Court Publishing Company (reprinted by Nabu Public Domain Reprints). archive.org/details/MagicSquaresAndCubes_754.

[3] Behrforooz, H. (2005). Behrforooz-Franklin 32 by 32 magic square. *J. Recreational Math* 33:2004–2005.

[4] Cammann, S. (1960). The evolution of magic squares in China. *J. Am. Oriental Soc.* 80:116–124.

[5] Euler, L. (1849). De quadratis magicis. *Commentationes Arithmeticae* 2:593–602 (Delivered to the St. Petersburg Academy October 17, 1776), (Translation: arxiv.org – arXiv:math/0408230v6[math.CO]).

[6] Hurkens, C. (2007). Plenty of Franklin magic squares, but none of order 12. win.tue.nl/bs/spor/2007-06.pdf.

[7] Jacobs, C. (1971). A reexamination of the Franklin square. *Math. Teacher* 64:55–62.

[8] Morris, D. (2009). Best Franklin squares. bestfranklinsquares.com.

[9] Nordgren, R. (2012). On properties of special magic squares. *Linear Algebra Appl.* 437:2009–2025.

[10] ——— (2017). On Franklin and complete magic square matrices. *Fibonacci Quart.* 54:304–318.

[11] ——— (2017). Eulerian composition of certain Franklin squares, arXiv:1703.06488.

[12] Ollerenshaw, K., Brée, D. (1998). Most-Perfect Pandiagonal Magic Squares: Their Construction and Enumeration. Southend-on-Sea, UK: The Institute of Mathematics and its Applications.

[13] Pasles, P. (2001). The lost squares of Dr. Franklin: Ben Franklin's missing squares and the secret of the magic circle. *Am. Math. Monthly* 108:489–511.

[14] ——— (2006). A bent for magic. *Math. Mag.* 79:3–13.
[15] ——— (2008). Benjamin Franklin's Numbers. Princeton, NJ: Princeton University Press.
[16] Pickover, C. (2002). The Zen of Magic Squares, Circles, and Stars. Princeton, NJ: Princeton University Press (also a Kindle e-book).
[17] Schindel, D., Rempel, M., Loly, P. (2006). Enumerating the bent diagonal squares of Dr Benjamin Franklin FRS. *Proc. Roy. Soc. A* 462:2271–2279.

Summary. Franklin squares of order 8*k* are constructed by Euler's composite method with specified forms for the two orthogonal auxiliary squares. Two types of formulas are given for elements of the auxiliary squares that are shown to be orthogonal and to satisfy Franklin's three sum conditions. Squares of order 8 and 16 agree with Franklin's published squares and those of order 24 and 32 agree with squares previously constructed by a direct method. It is shown that Franklin squares can be transformed to pandiagonal magic squares in two ways but the converse is not true in general.

RONALD P. NORDGREN (MR Author ID: 600461) is a retired professor of civil and mechanical engineering from Rice University. He holds degrees from the University of Michigan and the University of California, Berkeley with specialty in the mechanics of solids and applied mathematics. He spent 27 years in exploration and production R & D with Shell Development Company in Houston, Texas where he was elected to the National Academy of Engineering. He became interested in magic squares after viewing a 6 by 6 one in China (photo in [**9**]) and he has written several papers on the subject. Ron and his wife reside near Boulder, Colorado where they enjoy hiking, snowshoeing, and classical music.

The Metric Metric on S_4

BRET J. BENESH
College of Saint Benedict and Saint John's University
Saint Joseph, MN 56374
bbenesh@csbsju.edu

My favorite band is currently Metric. My kids' favorite band is currently Metric. Our favorite video is "Gimme Sympathy" [1], in which the band members switch instruments several times. Below, we translate their movement among the different instruments into the language of permutations. We can then use these permutations to define a distance function, or metric, on the set of all permutations of four objects, known as the symmetric group of degree 4 (denoted S_4). To do this, we employ the *word metric*, which is an idea from the field of geometric group theory (see [3], for example). In the end, we define the deliciously named *Metric metric on* S_4.

The video

The video for "Gimme Sympathy" features the members of Metric switching instruments in a curious manner. The camera might show the guitarist, then the drummer, and then return to the guitarist—only to find that a different person is now playing guitar. We will use the movement of the band members to obtain a set of permutations.

Impressively, this video was shot in one take, so there are no editing tricks. The video can be seen at the following two links, the first of which is the official music video with the second being a behind-the-scences video that shows how they were able to shoot the video without editing.

1. http://www.youtube.com/watch?v=LqldwoDXHKg
2. http://www.youtube.com/watch?v=jHt5caARmh0

The video starts with the lead singer and keyboardist, Emily Haines, alone and preparing to play some music with her bandmates. She soon moves to the rehearsal space, and we quickly see the band's usual formation: Haines on lead vocals, James Shaw on guitar, Joshua Winstead on bass, and Joules Scott-Key on drums. Soon after, the camera focuses on the guitarist Shaw. The camera then focuses on the drum kit, although we are surprised to see that Haines is now playing drums—several of the band members traded instruments while the camera was fixed on Shaw.

This movement can be described in the language of permutations. We will identify 1 with the location of the microphone, 2 with the guitar, 3 with the bass, and 4 with the drum kit. We can now define a permutation. We know that Shaw was a fixed point, as he was playing his guitar the entire time, so 2 maps to 2. Similarly, the vocalist (Haines) moves to the drums, so 1 maps to 4. However, we do not know what the bassist (Winstead) or drummer (Scott-Key) did. It could be that the vocalist and drummer simply swapped positions, which would yield the permutation (1, 4). But it could also be that the vocalist went to the drums, the drummer went to the bass, and the bassist went to the microphone; in this case, the permutation would be (1, 4, 3).

Math. Mag. **91** (2018) 33–36. doi:10.1080/0025570X.2018.1405697
MSC: Primary 20B30

Fortunately, the "making of" video clears up this ambiguity, showing that Haines moves to the drums and Scott-Key moves to the bass. Therefore, the first permutation is $(1, 4, 3)$ since the vocalist moved to drums, the drummer moved to bass, the bassist moved to vocals, and the guitarist remained at the guitar.

TABLE 1: The permutations in the first and fourth steps require viewing [2] to eliminate ambiguity.

Position	Vocals	Guitar	Bass	Drums	Permutation
0	Haines	Shaw	Winstead	Scott-Key	—
1	Winstead	Shaw	Scott-Key	Haines	$(1, 4, 3)$
2	Shaw	Winstead	Scott-Key	Haines	$(1, 2)$
3	Shaw	Haines	Scott-Key	Winstead	$(2, 4)$
4	Winstead	Haines	Scott-Key	Shaw	$(1, 4)$
5	Winstead	Scott-Key	Haines	Shaw	$(2, 3)$
6	Scott-Key	Shaw	Haines	Winstead	$(1, 4, 2)$
7	Haines	Shaw	Winstead	Scott-Key	$(1, 4, 3)$

We repeat this process for every rearrangement in the video and summarize the results in Table 1. Thus, we now have a set of permutations

$$M_0 = \{(1, 4, 3), (1, 2), (2, 4), (1, 4), (2, 3), (1, 4, 2)\}$$

that we will use later in this paper.

Word metrics on groups

The obvious thing to do when a band named Metric hands you a set of permutations is to try to make a metric from it. Recall that a metric on a set S is a function d that takes pairs of elements of S to nonnegative real numbers, and such a function d must fulfill the following for all $x, y, z \in S$:

1. $d(x, y) \geq 0$
2. $d(x, y) = 0$ if and only if $x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, y) + d(y, z) \geq d(x, z)$.

There are several possible metrics to use, but a common metric for groups is the *word metric*, which we describe. Let G be a group with a subset W such that W is closed under inverses and the smallest subgroup containing every element of W is G (i.e., W generates G, and we call W a *generating set that is closed under inverses*). We now define a metric from W by defining two functions. First, we define a norm on G: for all $g \in G$, denote by $|g|$ the minimal number of elements of W that are needed to multiply together to yield g (we define the product of 0 elements of W to be the identity e of G, so $|e| = 0$). This is known as the *word norm* of g with respect to W.

We now use this word norm to define the following: let $d : G \times G \to \mathbb{N}$ be defined by $d(x, y) = |xy^{-1}|$ for all $x, y \in G$. (Note the parallel to the distance function on $\mathbb{R}$, where the distance between $x, y \in \mathbb{R}$ is $|x - y|$; here, we simply use an element's inverse to mimic subtraction.) We claim that this function d defines a metric on G. Let x, y, and z be elements of G. Then

1. $d(x, y) \geq 0$: We have $d(x, y) = |xy^{-1}|$, which is just the word norm of some element of G and therefore nonnegative.
2. $d(x, y) = 0$ if and only if $x = y$: If $d(x, y) = 0$, then $0 = d(x, y) = |xy^{-1}|$; the only element with word norm 0 is e, so we must have $e = xy^{-1}$, or $x = y$. If $x = y$, then $d(x, y) = d(x, x) = |xx^{-1}| = |e| = 0$.
3. $d(x, y) = d(y, x)$: This is the reason why we require W to be closed under inverses. First, note that if $|x| = n$, then $|x^{-1}| = n$. To see this, suppose that x can be written as $w_1 w_2 \dots w_n$ for some elements $w_i \in W$; then $x^{-1} = w_n^{-1} \dots w_2^{-1} w_1^{-1}$. Since we ensured that W is closed under inverses, we know that $w_i^{-1} \in W$. Therefore, $|x| \geq |x^{-1}|$. By a symmetric argument, $|x^{-1}| \geq |x|$, so we conclude they are equal. This gives us
$$d(x, y) = |xy^{-1}| = |(xy^{-1})^{-1}| = |yx^{-1}| = d(y, x).$$
4. $d(x, y) + d(y, z) \geq d(x, z)$: Let $n = |xy^{-1}|$ and $m = |yz^{-1}|$. Then there exist $w_i \in W$ and $u_j \in W$ such that $xy^{-1} = w_1 w_2 \dots w_n$ and $yz^{-1} = u_1 u_2 \dots u_m$. Then
$$xz^{-1} = x(y^{-1}y)z^{-1} = (xy^{-1})(yz^{-1}) = (w_1 w_2 \dots w_n)(u_1 u_2 \dots u_m),$$
so $d(x, z) \leq n + m = d(x, y) + d(y, z)$.

Therefore, the word metric is indeed a metric. We now create a word metric from M_0.

The Metric metric on S_4

We now use M_0 to make a word metric. We first note that the 3-cycles in M_0 do not have an inverse in M_0, which creates a problem. For example, note that $(1, 2), (1, 4), (1, 4, 2) \in M_0$ but $(1, 2, 4) \notin M_0$, so (if multiplication is done left-to-right)
$$d((1, 2, 4), e) = |(1, 2, 4)e^{-1}| = |(1, 2, 4)| = |(1, 2)(1, 4)| = 2$$
while
$$d(e, (1, 2, 4)) = |e(1, 2, 4)^{-1}| = |(1, 4, 2)| = 1,$$
so $d((1, 2, 4), e) \neq d(e, (1, 2, 4))$, and M_0 yields distances that are not necessarily symmetric.

Our solution is to extend M_0 to a set M that includes all inverses of elements of M_0; this will ensure that our word metric is symmetric. We simply add the inverses of the two 3-cycles:
$$M := \{(1, 4, 3), (1, 2), (2, 4), (1, 4), (2, 3), (1, 4, 2), (1, 3, 4), (1, 2, 4)\}.$$
The only other requirement for M to be the foundation of a word metric is that M must generate S_4; we can see this in Table 2, since every $x \in S_4$ can be written as a product of elements in M.

Therefore, we can define a function $d : S_4 \times S_4 \to \mathbb{N}$ from the norm described in Table 2; the function d is a metric by the previous section. Therefore, we have defined the Metric metric on S_4, and we can say the distance between, say, $(2, 4)$ and $(1, 3, 2)$ is
$$d((2, 4), (1, 3, 2)) = |(2, 4)(1, 3, 2)^{-1}| = |(2, 4)(1, 2, 3)| = |(1, 2, 4, 3)| = 2.$$

TABLE 2: Multiplication is done left-to-right, so (1, 2)(1, 3) is (1, 2, 3), not (1, 3, 2).

Element	Product in M	Norm	Element	Product in M	Norm
e	empty product	0	(1,2,4)	(1,2,4)	1
(1,2)	(1,2)	1	(1,4,2)	(1,4,2)	1
(1,3)	(1,4)(1,4,3)	2	(1,3,4)	(1,3,4)	1
(1,4)	(1,4)	1	(1,4,3)	(1,4,3)	1
(2,3)	(2,3)	1	(2,3,4)	(2,3)(2,4)	2
(2,4)	(2,4)	1	(2,4,3)	(2,4)(2,3)	2
(3,4)	(1,4,3)(1,4)	2	(1,2,3,4)	(1,2)(1,3,4)	2
(1,2)(3,4)	(1,4,3)(1,4,2)	2	(1,2,4,3)	(1,2)(1,4,3)	2
(1,3)(2,4)	(1,4,2)(1,4,3)	2	(1,3,2,4)	(2,3)(1,3,4)	2
(1,4)(2,3)	(1,4)(2,3)	2	(1,3,4,2)	(2,4)(1,3,4)	2
(1,2,3)	(2,3)(1,2)	2	(1,4,2,3)	(2,3)(1,4,2)	2
(1,3,2)	(1,2)(2,3)	2	(1,4,3,2)	(2,3)(1,4,3)	2

Conclusion

We were able to use the permutations from the "Gimme Sympathy" video to create the Metric metric on S_4. This is not much more than an amusing result, but the use of word metrics on groups is an important idea with uses beyond providing mathematicians with an excuse to watch music videos on YouTube. Indeed, word metrics are a common tool in the field of geometric group theory, which is a field that uses geometric methods to better understand groups.

REFERENCES

[1] "Gimme Sympathy" [Official Music Video], YouTube.com. metricmusic. youtube.com/watch?v=LqldwoDXHKg.

[2] "Gimme Sympathy" [Behind The Scenes], YouTube.com. metricmusic. youtube.com/watch?v=jHt5caARmh0.

[3] Löh, C. (2011). Geometric group theory, an introduction. mathematik.uni-regensburg.de/loeh/teaching/ggt_ws1011/lecture_notes.pdf.

Summary. We define a metric on the symmetric group on four elements based on the music video for "Gimme Sympathy" by the band Metric. In the video, the four band members rearrange themselves according to a set of permutations. We extend this set to include inverses and use this to define a metric on the group. This gives us the aptly named *Metric metric on* S_4.

BRET J. BENESH (MR Author ID: 797761) is an Associate Professor of Mathematics at the College of Saint Benedict and Saint John's University in Central Minnesota. His research interests are finite group theory and combinatorial game theory, as well as learning how to teach better. He lives with his wife and two children, only one of whom is named after a type of mathematical function.

A Functional Equation View of an Addition Rule

MIHÁLY BESSENYEI
GRÉTA SZABÓ
University of Debrecen
Debrecen, Egyetem tér 1, 4032 Hungary
besse@science.unideb.hu
szabogreta55@gmail.com

We came across this exercise [**16**, Problem 923] in a problem-solving seminar, where one of the topics was functional equations: *Determine all functions* $f : \mathbb{R} \to \mathbb{R}$ *satisfying*

$$f(x+y) = \frac{f(x)+f(y)}{1+f(x)f(y)}. \tag{1}$$

In spite of our efforts, we were not able to understand either the hints or the solution. The hints state that the only possible solutions are constant functions. The critical part of the reasoning reads as follows: *If a solution* f *satisfies* $f(0) = 0$*, then the substitution* $x = -y$ *shows that* f *is identically zero.*

However, this substitution merely shows that f must be an odd function. Therefore, the official solution cannot be complete. Unfortunately, neither is it correct, since the following calculations demonstrate that $f(x) = \tanh(x)$ satisfies the same addition rule (and clearly $f(0) = 0$):

$$\tanh(x+y) = \frac{\sinh(x+y)}{\cosh(x+y)} = \frac{\sinh(x)\cosh(y)+\cosh(x)\sinh(y)}{\cosh(x)\cosh(y)+\sinh(x)\sinh(y)} = \frac{\tanh(x)+\tanh(y)}{1+\tanh(x)\tanh(y)}.$$

Hence we wondered: *How can the solutions of equation* (1) *be described completely?* First, we determine differentiable solutions to equation (1), and then we examine the necessity of differentiability. A careful analysis of the calculation suggests two ways to drop the extra assumption of differentiability. Proceeding the first way, we express the general solutions in terms of exponential functions. The second way links the general solutions to additive functions.

Prelude: preliminary experiments

Throughout this note, we assume $f(x)f(y) \neq -1$ for all $x, y \in \mathbb{R}$, so equation (1) makes sense. Note that substituting $x = y = 0$ in equation (1) leads to the algebraic equation $t^3 = t$ for $t = f(0)$ and hence $f(0) \in \{-1, 0, 1\}$. Assume now that $f(x_0) = 1$ for some $x_0 \in \mathbb{R}$. Then expanding $f(x + x_0)$ with the addition formula shows that $f(x) = 1$ for all $x \in \mathbb{R}$. Similar reasoning implies that $f(x) = -1$ for all $x \in \mathbb{R}$, provided there exists some $x_0 \in \mathbb{R}$ with $f(x_0) = -1$. These cases are trivial. Therefore, in the rest of the note, only the remaining case of $f(0) = 0$ is investigated. The previous remarks ensure that *the range of* f *does not include* $\{-1, 1\}$*, when* $f(0) = 0$.

Math. Mag. **91** (2018) 37–41. doi:10.1080/0025570X.2018.1403230
MSC: Primary 39B22, Secondary 00A05, 26A06

To get the complete solution, assume that the function we are seeking is as "beautiful" as we want it to be, by letting it be differentiable everywhere. Of course, this is not required, since equation (1) may be true without this assumption. But we hope, on one hand, that we can apply some standard tools of analysis, and on the other hand, that later we will be lucky enough to get rid of the assumption. We consider the rearranged form

$$f(x+y)(1+f(x)f(y)) = f(x)+f(y),$$

and differentiate it with respect to the variable x. Then substituting $y = -x$, one gets the separable differential equation:

$$f'(x) = f'(0)(1-f^2(x)). \tag{2}$$

In order to solve the differential equation, information about the range of f is needed. The observations at the beginning of this section imply that there is no x with $|f(x)| > 1$. If there were one, the intermediate value theorem would imply that either $f(x_0) = 1$ or $f(x_0) = -1$ for some x_0. Then f would be identically equal either to 1 or to -1, contradicting $f(0) = 0$. Knowing this, equation (2) can be separated and solved (where we let $c = f'(0)$):

$$\begin{aligned} cx &= \int_0^x \frac{f'(t)}{1-f^2(t)}dt = \frac{1}{2}\int_0^x \frac{f'(t)}{1+f(t)}dt + \frac{1}{2}\int_0^x \frac{f'(t)}{1-f(t)}dt \\ &= \frac{1}{2}[\log(1+f(t))]_{t=0}^x - \frac{1}{2}[\log(1-f(t))]_{t=0}^x \\ &= \frac{1}{2}\log(1+f(x)) - \frac{1}{2}\log(1-f(x)) = \log\sqrt{\frac{1+f(x)}{1-f(x)}}. \end{aligned}$$

Solving for f in the above equation yields

$$f(x) = \frac{\exp(2cx)-1}{\exp(2cx)+1} = \frac{\exp(cx)-\exp(-cx)}{\exp(cx)+\exp(-cx)} = \frac{\sinh(cx)}{\cosh(cx)} = \tanh(cx). \tag{3}$$

We had already shown that $f(x) = \tanh(x)$ is a solution to equation (1). From equation (3), differentiable solutions to equation (1) are of the form $y = \tanh(cx)$. Can we relax the assumption of differentiability? Let us try to calculate the difference quotient at x, where $x \in \mathbb{R}$ is fixed arbitrarily. Applying equation (1), we arrive at

$$\frac{f(x+h)-f(x)}{h} = \frac{1}{h}\cdot\left[\frac{f(x)+f(h)}{1+f(x)f(h)} - f(x)\right] = \frac{f(h)}{h}\cdot\frac{1-f^2(x)}{1+f(x)f(h)}.$$

If f is differentiable at zero, then $f(h)/h$ has a finite limit at zero, which has already been denoted by c. Moreover, f has to be continuous at zero, hence $f(h) \to f(0) = 0$ as $h \to 0$. Therefore, differentiability at zero implies *differentiability everywhere*. Furthermore, as a side effect, equation (2) follows as $h \to 0$. We summarize the above thoughts in the following theorem.

Theorem 1. *Consider those functions that are differentiable at zero and vanish at zero. Such a function $f : \mathbb{R} \to \mathbb{R}$ is a solution to equation (1) if and only if $f(x) = \tanh(cx)$ for some $c \in \mathbb{R}$.*

However, one question still remains: Is it possible to get rid of the differentiability condition completely?

Intermezzo: additive functions

Cauchy's functional equation is related to additive group homomorphisms of the real numbers and satisfies

$$a(x+y) = a(x) + a(y). \tag{4}$$

Restricting this equation to the rationals, any solution turns out to satisfy $a(x) = cx$, where $c \in \mathbb{Q}$. Is the situation similar for the reals? For a long time, this question challenged mathematicians. For interesting historical details, see the Aczél's book [**2**]. After several efforts, the general solution was obtained by Hamel [**12**]. He proved that there exist irregular solutions of the Cauchy equation, that is, additive functions on $\mathbb{R}$ differing from $a(x) = cx$. Irregular solutions have extremely strange behavior: their graphs are dense on the plane, and they are not measurable. They are neither differentiable nor continuous. Additive functions $a(x) = cx$ are commonly referred to as regular additive functions.

The Cauchy functional equation has many important and surprising applications including to Hilbert's third problem [**10**, **13**] and to Buffon's needle problem [**4**]. Both of these problems can be found among the topics of *Proofs from the Book* [**3**].

Of course, the Cauchy equation plays a crucial role in the field of functional equations, too. For example, consider a related equation, the exponential Cauchy equation:

$$g(x+y) = g(x)g(y). \tag{5}$$

It can be proved that either g is identically zero or $g = \exp \circ a$, where a is a solution to the Cauchy equation. In other words, the general solution can be expressed via additive functions. Tracing back the solution of a particular equation to the solution of a distinguished one is a typical method of this field. In ref. [**9**], Cauchy considered two further related equations of the form $f(xy) = f(x) + f(y)$ and $f(xy) = f(x)f(y)$, which he reduced to equations (4) and (5), respectively. Additive and exponential functions are the protagonists of the rest of this note.

Finale: answer completed

Returning to the original problem, the first part of equation (3) suggests expressing f in terms of an exponential function. This observation leads to the first characterization result.

Theorem 2. *A function $f : \mathbb{R} \to \mathbb{R}$ is a solution to equation (1) satisfying $f(0) = 0$ if and only if there exists an exponential function $g : \mathbb{R} \to \mathbb{R}$ such that*

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

Proof. As it had been observed earlier, equation (1) and the initial condition imply that $f(x) \neq 1$, so that

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

Then using equation (1),

$$g(x+y) = \frac{1 + f(x+y)}{1 - f(x+y)} = \frac{1 + f(x)f(y) + f(x) + f(y)}{1 + f(x)f(y) - f(x) - f(y)}$$
$$= \frac{1 + f(x)}{1 - f(x)} \cdot \frac{1 + f(y)}{1 - f(y)} = g(x)g(y).$$

In other words, g is an exponential function. For the converse statement, note that $g(x) + 1 > 0$ since g is an exponential function. This fact ensures that f is well defined. Moreover,

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

Thus, f fulfills equation (1). ■

Now let us turn our attention to the last part of equation (3). The $\tanh(cx)$ represents the solution as the composition of a bijective solution and an additive function. The second characterization result gives an alternative approach via this observation.

To formulate this result, the concept of groupoids is needed. A pair $(G, *)$ is called a *groupoid*, if G is nonempty and $* : G \times G \to G$ is an operation. A mapping $a : G \to G$ is said to be a *homomorphism* on the groupoid $(G, *)$ if $a(x * y) = a(x) * a(y)$ whenever $x, y \in G$. In other words, a satisfies the Cauchy functional equation on G.

Theorem 3. *Let $(G, *)$ be a groupoid, $X \neq \emptyset$, and $F : X^2 \to X$ a given function. If the equation $f(x * y) = F(f(x), f(y))$ has a bijective solution $\varphi : G \to X$, then all solutions have the form $\varphi \circ a$, where $a : G \to G$ is a homomorphism.*

Proof. Direct calculation shows that $f = \varphi \circ a$ is a solution. For the converse statement, consider the mapping $a = \varphi^{-1} \circ f$ where f is a solution and φ is a bijective solution of the equation in the theorem. Then, for all $x, y \in G$,

$$\begin{aligned}a(x * y) &= \varphi^{-1}(f(x * y)) = \varphi^{-1}(F(f(x), f(y))) \\ &= \varphi^{-1}(F(\varphi(a(x)), \varphi(a(y)))) = \varphi^{-1}(\varphi(a(x) * a(y))) = a(x) * a(y).\end{aligned}$$

This proves that a is a homomorphism, as required. ■

Applying Theorem 3 in the particular settings $G = \mathbb{R}$ and $X = (-1, 1)$, and using the fact that $\varphi = \tanh$ is a bijective solution with $f(0) = 0$, the desired connection between equation (1) and additive functions can be obtained easily: *Any solution $f : \mathbb{R} \to \mathbb{R}$ of equation (1) satisfying $f(0) = 0$ can be written in the form $f = \tanh \circ a$, where $a : \mathbb{R} \to \mathbb{R}$ is an additive function.*

At first glance, the formulae in Theorem 2 and Theorem 3 seem to be quite alien to one another. But, in fact, they are the same. We invite the reader to show their equivalence. What have we gained with this small adventure? Besides an elementary approach, now we have the complete set of solutions to equation (1). Theorem 2 and Theorem 3 apply when differentiability at zero fails, showing that this extra regularity is a serious restriction in Theorem 1.

Postlude: some comments

Functional equations are a mainstay in competitions. Besides the motivating exercise, the books of Brodskii–Slipenko [**7**], Lajkó [**15**], and Small [**17**] demonstrate this fact convincingly. Elementary problems can also have effect in recent researches (see [**5**, **6**, **11**]). For further details, see Kuczma's green book [**14**], which gives a deep overview of this field.

The paper of Caccioppoli [8] and also a part of Aczél's monograph [1] deal with equations of the form $f(x+y) = F(f(x), f(y))$. Among the applications, equation (1) is also considered. However, their approach is completely different from ours, and their methods require assumptions like monotonicity, continuity, and differentiability.

Surprisingly, Theorem 2 and Theorem 3 present the complete solution via simple calculations, while Theorem 1 gives only a partial answer using advanced tools. (Though in the case of Theorem 3, some algebraic background is definitely needed.) This demonstrates that an adequate approach can be both effective and simple simultaneously. Still, the role of Theorem 1 is not negligible: without it, any adequate approach would not have been found.

Acknowledgments The authors wish to express their gratitude to professor Gyula Maksa for the valuable comments on this note, and also to professor Zsolt Páles for suggesting the approach via Theorem 3. This research has been supported by the Hungarian Scientific Research Fund (OTKA) Grants K–111651.

REFERENCES

[1] Aczél, J. (1966). *Lectures on Functional Equations and Their Applications*, Mathematics in Science and Engineering, Vol. 19, New York, NY: Academic Press, pp. 5–12.
[2] Aczél, J. (Ed.) (1984). In: *Functional Equations: History, Applications and Theory*. Mathematics and Its Applications, Dordrecht: Reidel.
[3] Aigner, M., Ziegler, G. (2004). *Proofs from the Book*, 3rd ed., Berlin: Springer-Verlag.
[4] Barbier, J. E. (1860). Note sur le probléme de l'aiguille et le jeu du joint couvert. *J. Math. Pures Appl.* 5(2):273–286.
[5] Bessenyei, M. (2010). Functional equations and finite groups of substitutions. *Amer. Math. Monthly.* 117(10):921–927.
[6] Bessenyei, M., Horváth, G., Kézi, C. G. (2012). Functional equations and group substitutions. *Expo. Math.* 30(3):283–294.
[7] Brodskii, V. S., Slipenko, A. K. (1986). *Functional Equations* (in Russian). Kiev, USSR: Visa Skola.
[8] Caccioppoli, R. (1928). L'équazione funzionale $f(x+y) = F(f(x), f(y))$. *Giorn. Mat. Battaglini, III. Ser.* 66:69–74.
[9] Cauchy, A. L. (1821). *Cours d'Analyse de l'École Royale Polytechnique*. Partie I, Chapitre V, Cambridge University Press, 2009.
[10] Dehn, M. (1902). Üeber den Rauminhalt. *Math. Ann.* 55:465–478.
[11] Gong, X.-B., Shi, Y.-G. (2014). Linear functional equations involving Babbage's equation. *Elem. Math.* 69(4):195–204.
[12] Hamel, G. (1905). Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung: $f(x+y) = f(x) + f(y)$. *Math. Ann.* 60:459–462.
[13] Jessen, B., Karpf, J., Thorup, A. (1968). Some functional equations in groups and rings. *Math. Scand.* 22:257–265.
[14] Kuczma, M. (2009). *An Introduction to the Theory of Functional Equations and Inequalities. Cauchy's Equation and Jensen's Inequality*, 2nd ed. , Basel, Switzerland: Birkhäuser Verlag.
[15] Lajkó, K. (2005). *Functional Equations in Competition Problems* (in Hungarian). Debrecen, Hungary: University Press of Debrecen.
[16] Róka, S. (2006). *2000 Problems in Elementary Mathematics* (in Hungarian). Budapest, Hungary: TypoTex.
[17] Small, C. G. (2007). *Functional Equations and How to Solve Them*, New York, NY: Springer Science+Business Media, LLC.

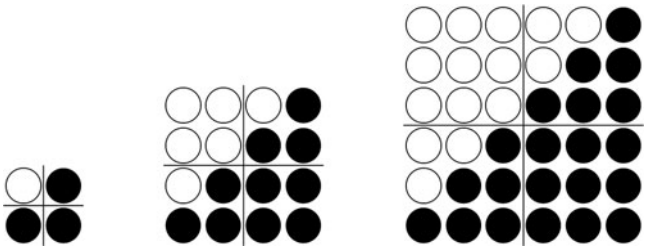
Summary. Motivated by a problem posed in a competition textbook, we give the general solution of a functional equation related to the addition theorem of the hyperbolic tangent function.

MIHÁLY BESSENYEI (MR Author ID: 705563) received his Ph.D. in mathematics in 2005, from the University of Debrecen, Hungary. He currently does his teaching and research activity there as an Associate Professor. Besides mathematics, he loves music, nature, and running. Although he considers these loves unrequited, he tries to do his best to meet them as frequently as possible.

GRÉTA SZABÓ (MR Author ID: 1199226) received her B.Sc. in mathematics in 2015, and earned her M.Sc. in applied mathematics in 2017 both from University of Debrecen, Hungary. Besides functional equations, she is interested in discrete and computational geometry. In her free time, she likes to reconnect with nature by hiking and birdwatching.

Proof Without Words: On Sums of Squares and Triangles

ANDRZEJ PIOTROWSKI
University of Alaska Southeast
Juneau, AK 99801
apiotrowski@alaska.edu



$$T_k = \sum_{j=1}^{k} j \quad \Longrightarrow \quad \sum_{k=1}^{2n} T_k = 4\sum_{k=1}^{n} k^2.$$

Remarks. If the diagram above is extended to contain the odd squares too, then it becomes clear that $\sum_{k=1}^{n} k^2 = \sum_{k=1}^{n} T_k + \sum_{k=1}^{n-1} T_k$, which can also be derived from the results given in either of the proofs without words [4] or [5]. Thus,

$$\sum_{k=1}^{2n} T_k = 4\left(\sum_{k=1}^{n} T_k + \sum_{k=1}^{n-1} T_k\right) = 4T_n + 8\sum_{k=1}^{n-1} T_k,$$

which can also be derived by summing the squares of the even integers using the result given in the proof without words [2] and the definition of T_n. Finally, the reader is invited to compare this to other proofs without words involving sums of squares and triangular numbers, especially [1] and [3].

REFERENCES

[1] Guy, R. K. (1993). Identities for triangular numbers, In: Nelsen, R. B. ed. *Proofs Without Words: Exercises in Visual Thinking*, Vol. 1, Washington, DC: Mathematical Association of America, 104.
[2] Landauer, E. G. (1985). Proof without words: Square of an even positive integer. *Math. Mag.* 58(4):236.
[3] Logothetti, D. (1987). Proof without words: Alternating sums of squares. *Math. Mag.* 60(5):291.
[4] Nelsen, R. (1995). Proof without words: Alternating sums of triangular numbers. *Math. Mag.* 68(4):284.
[5] Page, W. (1982). Proof without words: Count the dots. *Math. Mag.* 55(2):97.

Summary. We visually display a relationship between sums of squares and the sum of an even number of triangular numbers. Connections to some proofs without words appearing in the literature are briefly discussed.

ANDRZEJ PIOTROWSKI (MR Author ID: 1036554) received his B.S. and M.S. degrees from the University of New Hampshire and his Ph.D. degree from the University of Hawai'i at Mānoa. He is currently an Associate Professor of Mathematics at the University of Alaska Southeast in Juneau, AK.

Math. Mag. **91** (2018) 42. doi:10.1080/0025570X.2018.1404885
MSC: Primary 05A19

On Volumes of Hyper-Ellipsoids

SHAHNAWAZ AHMED
Queen Mary University of London
London, UK
shahnawaz.ahmed@qmul.ac.uk

ELIAS G. SALEEBY
The American University of Iraq Sulaimani
Sulaimania, Iraq
esaleeby@yahoo.com

Given a geometric object in Euclidian space, a natural question that arises is how to evaluate its volume and surface area. As we know from calculus, such geometric measures are often expressed in terms of multiple integrals. With a few exceptions, such integrals are often hard to resolve analytically especially in higher dimensions. To obtain estimates of the values of these integrals, one has to resort to numerical methods, and often, in higher dimensions, the Monte Carlo method is the only method that is practical. Closed-form expressions for geometric measures are of value in developing and testing new numerical methods. Such measures are available for spheres and cubes in $\mathbf{R}^n$, however, these objects may not serve the purpose fully. So the natural thing to do is to examine closely related convex objects with fewer symmetries.

The first family of objects that comes to mind, that could serve our objectives, are n-dimensional ellipsoids. It turns out that generalizations of n-dimensional ellipsoids are also well-known, and they are interesting objects to look at as well—as some of them are not necessarily convex. The purpose of this note is to develop a rather simple method of deriving the volumes of a large collection of generalized ellipsoids. In view of the well-known fact that the volume of the unit ball shrinks with increasing dimension, we consider a sub-collection of n -dimensional l_{2m} unit balls which fit more tightly in the cube than the unit ball, and discuss how their volumes decrease with increasing dimension. For estimating the volumes numerically, we employ an elementary Monte Carlo (MC) method that has a wider application. Finally, we take a brief look at volumes of revolution in higher dimensions.

Volumes of generalized super-ellipsoids

In this section, we derive a formula for the volumes of generalized super-ellipsoid balls defined by the equation $\sum_{i=1}^{n} |\frac{x_i}{c_i}|^{p_i} \leq 1$, $p_i > 0$. The generalized unit balls in $\mathbf{R}^n$ are defined as the sets

$$\{(y_1, \ldots, y_n) : |y_1|^{p_1} + |y_2|^{p_2} + \cdots + |y_n|^{p_n} \leq 1,\ p_i > 0\}.$$

For $p_i = p \geq 1$, p a positive integer, and $i = 1, 2, \ldots, n$, we have the n-dimensional l_p balls. Generalized super-ellipsoid balls can be easily transformed into generalized unit balls. Formulas for volumes of generalized super-ellipsoid balls were obtained by Lejeune Dirichlet [4] and were rediscovered recently by Wang [10]. In this section, we point out another derivation method initiated by Hein [6]. Due to the elementary nature of the problem of evaluating the volume integrals, one would expect that the methods utilized to be somewhat related. The common thread is due to the appearance

Math. Mag. **91** (2018) 43–50. doi:10.1080/0025570X.2018.1404834
MSC: Primary 51M25, Secondary 65C05, 97G30

of integrals related to the gamma and beta functions and the more flexible Gauss' hypergeometric function. In the study of calculus and differential equations, these functions are not identified as elementary functions and belong to the class of "special functions." Often special functions have integral or power series representations and satisfy functional or differential equations. For the convenience of the reader, we recall the definitions of these functions.

Definition 1 (Euler gamma function). The Euler gamma function is a higher transcendental function defined by the integral

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \ x > 0.$$

It is easy to see that $\Gamma(1) = 1$. Integration by parts yields the important formula $\Gamma(x+1) = x\Gamma(x)$, which allows us to extend the definition of the factorial to other real numbers, and for a positive integer n, we see that $\Gamma(n+1) = n!$.

A closely related function also introduced by Euler is the beta function.

Definition 2 (Beta function). The beta function is defined by the integral

$$B(x, y) = \int_0^1 u^{x-1}(1-u)^{y-1}\, du; \quad x, y > 0.$$

The formula $B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ relates the beta function to the gamma function. For further properties of the gamma and beta functions, see [2].

Another special function which plays a central role in the theory of special functions and in the theory of ordinary differential equations is the Gaussian hypergeometric function. This function is a generalization of the geometric series $\Sigma_{n=0}^{\infty} x^n$.

Definition 3 (Gaussian hypergeometric function). Gauss' hypergeometric function is defined by the power series

$${}_2F_1(a, b; c; z) = \sum_{k\geq 0} \frac{(a)_k\,(b)_k}{(c)_k} \frac{z^k}{k!}, \ |z| < 1,$$

where a, b, c are complex numbers, $c \neq 0, -1, -2, \ldots$, and $(\alpha)_k$ is the Pochhammer symbol defined as $(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$, $n > 0$, $(\alpha)_0 = 1$ for $\alpha \neq 1$.

For convenience, denote ${}_2F_1$ by F. The following integral

$$\int (a - x^m)^{\frac{d}{n}}\, dx = a^{\frac{d}{n}} x F\left(\frac{1}{m}, -\frac{d}{n}; 1 + \frac{1}{m}; \frac{x^m}{a}\right), \tag{1}$$

for $|x^m| < a$, will be useful to us below. Hein employed equation (1) to compute the volumes of three-dimensional super-ellipsoids. A large number of functions are special cases of the Gaussian hypergeometric function. The limiting behavior of $F(a, b; c; z)$ is discussed in [8] (chapter 4).

We denote the n-dimensional sphere centered at the origin and of radius r by $S_n(r)$. The volume of $S_n(r)$ is given by $v_n(r) = \frac{\pi^{\frac{n}{2}} r^n}{\Gamma(1+\frac{n}{2})}$. S_n are special cases of the n-dimensional ellipsoids E_n given by $\sum_{i=1}^n \frac{x_i^2}{c_i^2} = 1$. The volume of E_n centered at the origin is given by $V_{E_n} = v_n(1) \prod_{i=1}^n c_i$, where the c_i's are the lengths of the semi-axes. Whenever $n > 3$, S_n and E_n are also sometimes referred to in the literature as hyper-spheres and hyper-ellipsoids, respectively. Herein, we refer to them as the n-sphere and the n-ellipsoid, respectively. We reserve the prefix "hyper" to denote hyper-ellipsoids

$E_n(m)$ defined by the equation $\sum_{i=1}^{n}\left(\frac{x_i}{c_i}\right)^{2m}$ $m = 2, 3, \ldots$. For convenience, and to differentiate from super-ellipsoids and hyper-ellipsoids, we refer to geometric objects associated with the Fermat varieties described by $\sum_{i=1}^{n} x_i^{2m} = 1$, $m > 1$, as Fermatoids. Fermatoids are equiaxial hyper-ellipsoids and are n-dimensional l_{2m} unit spheres. The symmetry about the diagonals possessed by Fermatoids make them closer relatives and generalizations of the n-sphere, and thus play a similar role. For instance, upon using a linear transformation of the form $y_i = \frac{x_i}{c_i}$, the equation of a hyper-ellipsoid can be converted into that of a Fermatoid. Now we give some notation that we employ in the proof of Theorem 1.

Definition 4 (n-dimensional generalized super-ellipsoid ball). The n-dimensional generalized super-ellipsoid ball is defined as

$$E(n, p) = \left\{(x_1, \ldots, x_n) : \sum_{i=1}^{n}\left|\frac{x_i}{c_i}\right|^{p_i} \leq 1,\, p_i > 0\right\}.$$

The positive orthant of $E(n, p)$ is defined as

$$E'(n, p) = \left\{(x_1, \ldots, x_n) : \sum_{i=1}^{n}\left(\frac{x_i}{c_i}\right)^{p_i} \leq 1,\ x_i \geq 0,\ p_i > 0\right\}.$$

The positive orthant of the generalized unit ball is defined as

$$E''(n, p) = \left\{(y_1, \ldots, y_n) : \sum_{i=1}^{n} y_i^{p_i} \leq 1,\ y_i \geq 0,\ p_i > 0\right\}.$$

For $p_i = 2m$, m a positive integer greater than 1, $i = 1, 2, \ldots, n$, we have the Fermatoid ball.

The following theorem gives a formula for the volumes of generalized super-ellipsoids.

Theorem 1. *The volumes of the generalized super-ellipsoids $E(n, p)$ are given by*

$$V_E(n, p) = 2^n \frac{\prod_{i=1}^{n} c_i \Gamma\left(1 + \frac{1}{p_i}\right)}{\Gamma\left(1 + \sum_{i=1}^{n} \frac{1}{p_i}\right)};\ p_i > 0, n = 1, 2, 3, \ldots. \tag{2}$$

Proof. The volume of $E(n, p)$ is given by the integral

$$2^n \int_{E'(n,p)} dx_1 \cdots dx_n = 2^n \left(\prod_{i=1}^{n} c_i\right) \int_{E''(n,p)} dy_1 \cdots dy_n,$$

where equality comes from the affine transformation

$$(y_1, \ldots, y_n) = \left(\frac{x_1}{c_1}, \ldots, \frac{x_n}{c_n}\right).$$

Let P_{i-1} be denoted by

$$F\left(\frac{1}{p_{i-1}}, -\sum_{k=i}^{n} \frac{1}{p_k}; 1 + \frac{1}{p_{i-1}}; 1\right).$$

The method we utilize below is similar to the one usually used to evaluate the iterated integral for the volume of $S_3(1)$,

$$V = 8\int_0^1 dy_1 \int_0^{(1-y_1^2)^{\frac{1}{2}}} \left(1 - y_1^2 - y_2^2\right)^{\frac{1}{2}} dy_2.$$

Employing equation (1) and a result due to Gauss (Theorem 18 in [8], p. 47)

$$F(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}.$$

For Re $(c - a - b) > 0$, and c is neither zero nor a negative integer, we obtain

$$V = \frac{16}{3}F\left(\frac{1}{2}, -\frac{1}{2}, \frac{3}{2}, 1\right) = \frac{16}{3}\left(\frac{\sqrt{\pi}}{2}\right)^2 = \frac{4}{3}\pi.$$

Therefore, we start first by integrating over y_n, and employ equation (1) to obtain a reduction in the dimension to get

$$\begin{aligned}
\int_{E''(n,p)} dy_1 \cdots dy_n &= \int_{E''(n-1,p)} dy_1 \cdots dy_{n-1} \int_0^{\left(1-\sum_{i=1}^{n-1} y_i^{p_i}\right)^{\frac{1}{p_n}}} dy_n \\
&= \int_{E''(n-1,p)} \left(1 - \sum_{i=1}^{n-1} y_i^{p_i}\right)^{\frac{1}{p_n}} dy_1 \cdots dy_{n-1} \\
&= \int_{E''(n-2,p)} dy_1 \cdots dy_{n-2} \int_0^{\left(1-\sum_{i=1}^{n-2} y_i^{p_i}\right)^{\frac{1}{p_{n-1}}}} \left(1 - \sum_{i=1}^{n-1} y_i^{p_i}\right)^{\frac{1}{p_n}} dy_{n-1} \\
&= P_{n-1} \int_{E''(n-2,p)} \left(1 - \sum_{i=1}^{n-2} y_i^{p_i}\right)^{\frac{1}{p_n}+\frac{1}{p_{n-1}}} dy_1 \cdots dy_{n-2}.
\end{aligned}$$

Next, repeating the process and using equation (1) again, we get

$$\begin{aligned}
&P_{n-1} \int_{E''(n-2,p)} \left(1 - \sum_{i=1}^{n-2} y_i^{p_i}\right)^{\frac{1}{p_n}+\frac{1}{p_{n-1}}} dy_1 \cdots dy_{n-2} \\
&= P_{n-1} \cdot P_{n-2} \int_{E''(n-3,p)} \left(1 - \sum_{i=1}^{n-3} y_i^{2m}\right)^{\frac{1}{p_n}+\frac{1}{p_{n-1}}+\frac{1}{p_{n-2}}} dy_1 \cdots dy_{n-3}.
\end{aligned}$$

Continuing the reduction process, we arrive at

$$\prod_{i=3}^{n} P_{i-1} \int_0^1 \left(1 - y_1^{p_1}\right)^{\sum_{j=2}^{n} \frac{1}{p_j}} dy_1 = \prod_{i=2}^{n} P_{i-1};$$

which by Gauss' theorem yields the volume in equation (2). ∎

For $c_i = 1$, $i = 1, 2, \ldots, n$ and $p_i = p$, where p is an integer, we obtain the volumes of the n-dimensional l_p balls (or Fermatoid balls for $p > 2$, even). Both the method of Wang and the method of Dirichlet employed the transformations $y_i = (\frac{x_i}{c_i})^{p_i}$, $i = 1, \ldots, n$ and made use of the integral $\int |x_1|^{i_1} \cdots |x_n|^{i_n} dx$ over $E'(n, p)$, expressed in terms of gamma functions. Wang employed a recurrence method in his derivation, while Dirichlet made clever use of the beta function integral to reduce the problem to

integrating over an n -simplex [4,10]. The method we employ is a straightforward extension of Hein's method to higher dimensions and for all $p_i > 0$, and it is somewhat simpler as it does not involve the use of the Jacobian of the above transformation.

A well-known observation [9] is that the volume of the unit sphere ($r = 1$) is maximal at $n = 5$, and the volume of the unit sphere approaches 0 as n goes to infinity. In contrast, the n-dimensional unit cube $[\frac{-1}{2}, \frac{1}{2}]^n$ has volume one. Note also that the unit ball is contained in the cube $[-1, 1]^n$, which has vertices at a distance equal to $\sqrt{n}$ from the center. A rough explanation for the decay of the volume of the sphere is that in higher dimensions, more measure of the cube is being shaved-off to obtain the unit sphere. This explanation is based on the fact that all points inside the n-sphere are within a unit distance from its center, whereas for the n -dimensional unit cube we have points that lie at a distance $\sqrt{\frac{n}{2}}$ away from the center, and so most of the volume of the unit cube is outside the unit sphere. Moreover, a statistical argument ([3], p. 184) shows that the expected length of an n-dimensional vector picked at random from the unit sphere equals $\frac{n}{n+1}$. This means that most of the volume of a high-dimensional sphere is contained near the surface. One can ask if this is also true for generalized unit balls, and how this can be characterized.

Toward this end, the integral in equation (1) should be of help in evaluating the volume content of the end caps of Fermatoids with respect to a chosen coordinate plane. Furthermore, note that the behavior of the volume of the n-ellipsoid $E_n = v_n(1) \prod_{i=1}^n c_i$, does not behave like the volumes of $S_n(1)$ in general, as it depends on the limit of the product $\prod_{i=1}^n c_i$. We can ask if the volumes of Fermatoids behave similarly to the volumes of $S_n(1)$. We know from equation (2) that the volumes of Fermatoids are given by $V_F(n, m) = 2^n \frac{(\Gamma(1+\frac{1}{2m}))^n}{\Gamma(1+\frac{n}{2m})}$. First, it is clear that for a fixed n, if we let $m \to \infty$, we end up with 2^n, which is the volume of the n-dimensional cube with sides equal 2. On the other hand, fix m and let n go to infinity. By Stirling's formula, we have

$$\Gamma(1 + n\alpha) \sim \sqrt{2\pi n\alpha}\left(\frac{n\alpha}{e}\right)^{n\alpha}.$$

For $\alpha = \frac{1}{2m}$, $m \geq 1$, it is not difficult to show that

$$\lim_{n\to\infty} V_{F(n,m)} = \lim_{n\to\infty} \frac{2^n(\Gamma(1+\alpha))^n}{\sqrt{2\pi n\alpha}\left(\frac{n\alpha}{e}\right)^{n\alpha}} = 0.$$

In other words, once m is fixed, the volume of the Fermatoid behaves much like $S_n(1)$, and hence, it attains a maximal volume for a certain dimension $n_{\max}$. Just as in the case of $S_n(1)$, this dimension can be determined numerically.

A Monte Carlo method

Before arriving at equation (2), we initially computed the volumes using a geometric Monte Carlo method that we now describe. First, generate a uniformly distributed set of points on the surface of the unit sphere. It is well known that by generating u_i, $1 \leq i \leq n$, from the standard normal distribution, and after normalizing by the square root of the sums of squares, $||u||$, we obtain a uniformly distributed set of points (with respect to Lebesgue measure) on the unit sphere. This method is due to Muller [7]. Generate a sample of N points on the sphere and employ these points to generate and send random rays out from the origin to intercept the surface of the ellipsoid. Start at

$V = 0$. If the length of the ray to the intersection point is R, we compute

$$V = V + v_n(R) = V + \frac{\pi^{\frac{n}{2}} R^n}{\Gamma\left(1 + \frac{n}{2}\right)},$$

and so the final hyper-volume is approximately V/N. In other words, in this Monte Carlo method, we estimate the volume of interest by the average of the volumes of spheres in a spherical covering. Carrying out some numerical experiments on some convex-shaped objects with known volumes shows that this method gives remarkably good estimates for the volumes (see Tables 1 and 2).

TABLE 1: Volumes of Fermatoids of degree 4 obtained from equation (2) and estimates computed from the MC method using 100,000 points, for different dimensions n.

n	MC	Equation (2)
2	3.7085	3.7091
3	6.4768	6.4820
4	10.8050	10.7995
5	17.2798	17.2792
6	26.6993	26.6975

TABLE 2: Volumes of hyper-ellipsoids computed as in Table 1, of degree 4, for various dimensions and semi-axes c_i (taken in arithmetic progression).

n	c_i	MC	Equation (2)
2	1, 2	7.4081	7.4163
3	1, 2, 3	38.9441	38.8919
4	1, 2, 3, 4	259.399	259.188
5	1, 2, 3, 4, 5	2083.2	2073.5
6	1, 2, 3, 4, 5, 6	19,287	19,222

Volumes of revolution

In this last section, we briefly look at computing certain hyper-volumes of revolution. Herein, we are particularly interested in volumes of revolution generated by revolving the graph of the nonnegative function

$$y = f(x) = \left(1 - x^{2m}\right)^{\frac{1}{2m}}, -1 \le x \le 1$$

about an axis. The graph of this function can be viewed as a curve in the xy-plane embedded in $\mathbf{R}^n$. Revolving this graph about an axis spawns a new axis orthogonal to the xy-plane [5]. For the detailed derivation of the formulas for the volumes and hypersurface areas of revolution about either the x- or the y -axis in dimensions greater than 3, see [1,5]. The volumes of revolution about the x-axis and the y-axis are given by

$$V_x(n, m) = v_{n-1}(1) \int_{-1}^{1} f^{n-1}(x)dx \text{ and } V_y(n, m) = a_{n-1}(1) \int_{-1}^{1} x^{n-2} f(x)dx,$$

respectively, where $v_n(1)$ is the volume of the n-dimensional unit ball, and $a_n(1)$ is the surface area of the n-dimensional unit sphere. The surface area $a_n(r)$ of $S_n(r)$ is given by $\frac{dv_n(r)}{dr} = \frac{n}{r}v_n(r)$. Thus, we have

$$V_x(n,m) = \frac{2\pi^{\frac{n-1}{2}}}{\Gamma\left(1+\frac{n-1}{2}\right)}\int_0^1 (1-x^{2m})^{\frac{n-1}{2m}}dx = \frac{2\pi^{\frac{n-1}{2}}\Gamma\left(1+\frac{n-1}{2m}\right)\Gamma\left(\frac{1}{2m}\right)}{m(n-1)\Gamma\left(1+\frac{n}{2m}\right)\Gamma\left(\frac{n-1}{2}\right)}.$$

Indeed, in this special case, it is easy to show that $V_x(n,m)$ is identical to $V_y(n,m)$—this also can be deduced from the symmetry of the equation. In Table 3, we present some values obtained for V_x, and for comparison purposes, we give the corresponding values of $V_E(n,m)$ obtained from equation (2). Note that, for $n = 2$, the volumes and volumes of revolution about the x-axis and y-axis are identical for a given m. For $m = 1$, we obtain the volume of the n-sphere.

TABLE 3: Volumes of revolution for the function $y = (1-x^{2m})^{\frac{1}{2m}}$, $-1 \le x \le 1$ for different values of m and n.

n	m	V_E	$V_x = V_y$
2	1	3.1416	3.1416
2	2	3.7081	3.7081
2	3	3.8552	3.8552
2	4	3.9138	3.9138
4	2	10.7995	6.9789
4	3	13.1287	7.6298
4	4	14.2005	7.9134
6	2	26.6975	8.1329
6	3	40.8018	9.1870
6	4	48.5768	9.6717

It is sometimes of interest to examine the ratios of volumes. From our volume expressions, we see that this would lead to examining the ratios of gamma functions. We caution the reader that this has to be carried out carefully by choosing the appropriate approximation for the gamma function in order to avoid contradictory results. For instance, instead of using Stirling's formula we utilized above, an extended version of Stirling's approximation, for example, $\Gamma(1+x) \sim \sqrt{2\pi}x^{x+\frac{1}{2}}e^{-x+\frac{1}{12x}}$, may prove to be more appropriate.

Acknowledgment The authors would like to thank the referees and the editor Michael Jones for helpful suggestions to improve the presentation of this article.

REFERENCES

[1] Aberra, D., Agrawal, K. (2007). Surface of revolution in n dimensions. *Int. J. Math. Educ. Sci. Technol.* 38:843–852.
[2] Artin, E. (1964). *The Gamma Function*. New York, USA: Holt, Reinhardt and Winston.
[3] DasGupta, A. (2011). *Probability for Statistics and Machine Learning*. New York: Springer.
[4] Edwards, J. (1922). *A Treatise on the Integral Calculus*, Vol II. NY: Chelsea Pub. Co.
[5] Eisenberg, B. (2004). Surfaces of revolution in four dimensions. *Math. Magazine* 77:379–386.
[6] Hein, P. (2001). Superellipse. `matematiksider.dk/piethein.html`.
[7] Muller, M. E. (1959). A note on a method for generating points uniformly on n-dimensional spheres. *Commun. ACM* 2:19–20.
[8] Rainville, E. D. (1960). *Special Functions*. NY: The MacMillan Co.

[9] Smith, D. J., Vamanamurthy, M. K. (1989). How small is a unit ball? *Math. Magazine* 62:101–107.
[10] Wang, X. (2005). Volumes of generalized unit balls. *Math. Magazine* 78:390–395.

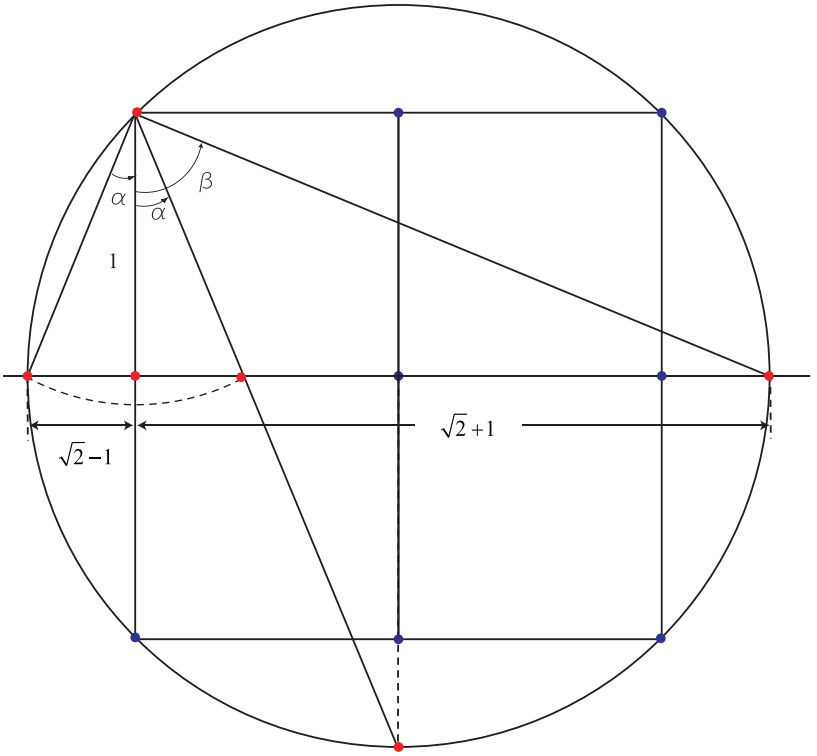
Summary. Volumes of objects in higher dimensions are of interest in higher-dimensional geometry. In this note, we give a sampling of some of the topics that are considered in this field. We first provide an alternative proof for a formula for the volumes of generalized super-ellipsoids that was obtained by Dirichlet in the 19th century. Then we employ a geometric Monte Carlo method that allows us to estimate hyper-volumes numerically. We then end the presentation with a brief discussion on the volumes of revolution in higher dimensions.

SHAHNAWAZ AHMED (MR Author ID: 952783) holds a M.Sc. in mathematics from the Indian Institute of Technology Madras (IITM). He is currently a graduate student in computer vision at Queen Mary University of London. He has research interests in modeling 3D data obtained from consumer depth-cameras, as well as multiple-view reconstruction from images. He is also interested in inverse problems and numerical linear algebra.

ELIAS G. SALEEBY (MR Author ID: 624911) holds a Ph.D. in mathematics and a Ph.D. in chemical engineering, both from the University of Arkansas. He is currently an assistant professor of mathematics at the American University of Iraq. He has research interests in function theory and nonlinear PDE of complex variables. He also likes to consider some problems in statistics and economics that require mathematical analysis.

Proof Without Words: Three Arctangent Identities

ÁNGEL PLAZA
Universidad de Las Palmas de Gran Canaria
Las Palmas, Spain
angel.plaza@ulpgc.es



$$\alpha = \arctan(\sqrt{2}-1) = \frac{\pi}{8}, \quad \beta = \arctan(\sqrt{2}+1) = \frac{3\pi}{8},$$
$$\beta - \alpha = \arctan(\sqrt{2}+1) - \arctan(\sqrt{2}-1) = \frac{\pi}{4}.$$

Summary. Visual proof of three arctangent identities involving $\arctan(\sqrt{2}-1)$ and $\arctan(\sqrt{2}+1)$.

ÁNGEL PLAZA (MR Author ID: 350023, ORCID 0000-0002-5077-6531) received his masters degree from Universidad Complutense de Madrid in 1984 and his Ph.D. from Universidad de Las Palmas de Gran Canaria in 1993, where he is a Full Professor in Applied Mathematics.

Math. Mag. **91** (2018) 51. doi:10.1080/0025570X.2018.1404888
MSC: Primary 51M04, Secondary 97G60
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

Designing for Minimum Elongation

NIELS CHR. OVERGAARD
Centre for Mathematical Sciences
Lund University, Sweden
nco@maths.lth.se

In a short paper [10], Verma and Keller suggested and found the solution to the following optimal design problem: Consider a heavy rope hanging vertically from a fixed support (the ceiling) and stretched due to its own weight and that of an additional load attached at the rope's lower end. Assume in addition that the volume and length of the unstretched rope are known. How should one taper the rope so that its elongation becomes smallest possible? The paper, which appeared in SIAM Review some 30 years ago under the headline "Classroom Notes in Applied Mathematics," had the pedagogical aim of showing how to solve a real world problem (in elasticity) using methods from calculus of variations. In the present paper, we prove the optimality of the solution found in [10], something which is missing in the original paper. This is achieved by reformulating the problem as a variational problem in standard form, unconstrained and with fixed end-points. The paper uses ideas from calculus of variations, but reading the optimality proof actually does not require any deep prior knowledge of the subject. For the particularly interested readers, the ideas used in the optimality proof will be examined further and put into their right context afterward.

The rest of the paper is organized as follows: In the next three sections we first give the problem formulation, recall some facts from the calculus of variations, and then reproduce the original solution from [10], using essentially the same notation as in that paper, but a different coordinate system (for convenience alone). The subsequent two sections cover the reformulation of the problem and its solution, including the proof of optimality. The principles underlying the optimality proof is then discussed in a separate section. The penultimate section briefly mentions two other solutions to the problem. Finally, we make some concluding remarks concerning the didactic merits of the problem and its solution.

Problem formulation

The unstretched rope is assumed to have length L, a fixed total volume V, and a uniform density ρ. The weight of the attached load is denoted W. The problem is essentially one-dimensional, and it turns out to be convenient to choose a coordinate system with the origin at the lower end of the rope and whose axis points vertically up toward the ceiling, see Figure 1. Thus, $x = 0$ is (always) where the additional load is attached to the rope and the ceiling is located at $x = L$ when the rope is unstretched. The unstretched rope (gravitational constant $g = 0$) is taken as reference state. Denote by $A(x)$ the cross-sectional area at x of the unstretched rope. Each positive function $A : [0, L] \to \mathbb{R}_+$ corresponds to a possible design of the rope. (The shape of each cross-section is not considered relevant for the model.) The volume of the unstretched rope is a functional $\mathcal{V}[A]$ of the design A, and those designs which are admissible for the

Math. Mag. **91** (2018) 52–61. doi:10.1080/0025570X.2017.1404790
MSC: Primary 49-01, Secondary 74K05, 74P

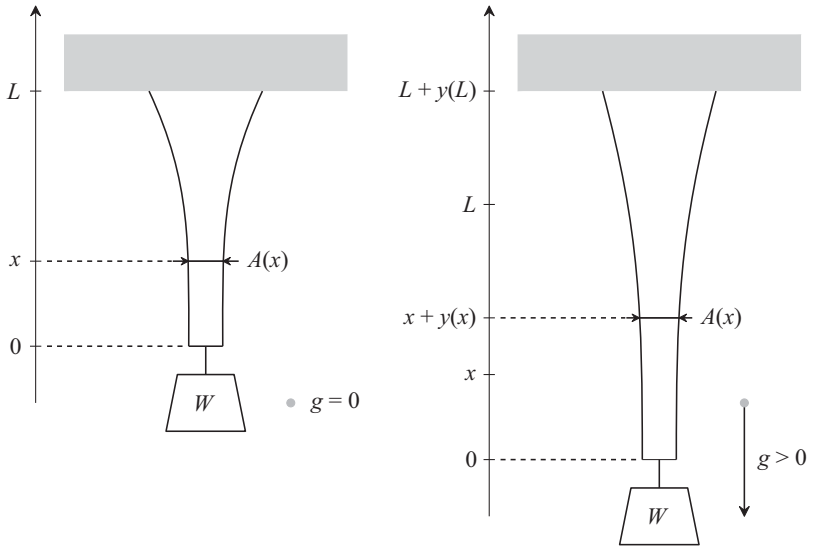


Figure 1 The rope in its unstretched (reference-) state, left, and in the stretched equilibrium state, right.

minimum elongation problem satisfy

$$\mathcal{V}[A] := \int_0^L A(x)\,dx = V. \tag{1}$$

When the load is applied (when $g > 0$), the rope is stretched, and the point originally at x will be displaced by an amount $y(x)$ relative to its initial position, and its coordinate at equilibrium is consequently $x + y(x)$. The displacement $y(x)$ is a monotone increasing function of x and $y(L)$ corresponds to the total elongation of the rope.

For moderate displacements, Hooke's law of elasticity applies. It states that the stress at x, $\sigma(x)$, is proportional to the strain $y'(x)$. The stress at x equals the combined weight of the applied load and the mass of the part of the rope lying below x, divided by the cross-section area at x. Hooke's law therefore implies that

$$Ey'(x) = \sigma(x) = \frac{1}{A(x)}\left[W + \rho g \int_0^x A(x')\,dx'\right],$$

where E is Young's modulus of the material and g is the gravitational constant. Since $y(0) = 0$, we find

$$y(x) = \int_0^x \frac{1}{EA(x')}\left[W + \rho g \int_0^{x'} A(x'')\,dx''\right] dx'.$$

The total elongation of the rope is

$$\mathcal{E}[A] := y(L) = \int_0^L \frac{1}{EA(x')}\left[W + \rho g \int_0^{x'} A(x'')\,dx''\right] dx', \tag{2}$$

and our task is to find a piecewise continuous function $A : [0, L] \to \mathbb{R}_+$ which minimizes the expression (2) and simultaneously satisfies the integral constraint (1). In the calculus of variations such a problem is known as an *isoperimetrical problem*. The above functional is, however, not written in the standard form. Before we do that, let us recall some basic terminology and results from the calculus of variations.

Euler's equation and extremals

The standard problem[1] of the calculus of variations is to find a function y_0 which minimizes an integral (or functional) of the form

$$\mathcal{J}[y] = \int_a^b F(x, y(x), y'(x))\, dx,$$

where the function $F = F(x, y, z)$, the Lagrangian, is a sufficiently smooth function defined for points (x, y, z) in an open domain in $\mathbb{R}^3$. The minimum of the integral is taken over all continuously differentiable functions y, defined on the interval $[a, b]$, such that each y satisfies the two conditions, (i) the end point conditions $y(a) = \alpha$ and $y(b) = \beta$, where α and β are given real numbers, and (ii) the point $(x, y(x), y'(x))$ belongs to the domain of F for all $x \in [a, b]$. Such functions are said to be admissible for the problem and the set of admissible functions is denoted $\mathcal{A}$.

Suppose that y_0 is a minimizer of the integral, *i.e.*, $y_0 \in \mathcal{A}$ and $\mathcal{J}[y_0] \le \mathcal{J}[y]$ for all $y \in \mathcal{A}$. Then y_0 satisfies the differential equation:

$$\frac{\partial F}{\partial y}(x, y_0(x), y_0'(x)) - \frac{d}{dx}\frac{\partial F}{\partial y'}(x, y_0(x), y_0'(x)) = 0, \quad a < x < b, \tag{3}$$

where $\partial F/\partial y'$ denotes the partial derivative $\partial F/\partial z$. This is *Euler's equation* (or Euler-Lagrange's equation), and its proof can be found in any introduction to the calculus of variations. However, the result is so central to the subject that we prefer to recall the proof here: We first make a "variation" of the minimizer y_0 by defining, for any test function $h \in C_0^1 := \{\eta \in C^1([a, b]) : \eta(a) = \eta(b) = 0\}$, a new function $y_\epsilon(x) = y_0(x) + \epsilon h(x)$. This function meets the required end point-conditions (i) for all real ϵ, because $h(a) = h(b) = 0$, and so belongs to $\mathcal{A}$ when ϵ is sufficiently close to zero, so that condition (ii) is also satisfied. Now, since y_0 is a minimizer over $\mathcal{A}$ it is certainly a minimizer over the subset of $\mathcal{A}$ consisting of the variations y_ϵ. Therefore, the function $\epsilon \mapsto \int_a^b F(x, y_0 + \epsilon h, y_0' + \epsilon h')\, dx$ attains its minimum value at the interior point $\epsilon = 0$, hence its derivative must vanish there. If we differentiate under the integral sign and set $\epsilon = 0$, the well-known necessary condition for optimality is obtained:

$$\int_a^b F_{y_0}(x)h(x) + F_{y_0'}(x)h'(x)\, dx = 0, \quad \text{for all } h \in C_0^1. \tag{4}$$

Here $F_{y_0}(x) = F_y(x, y_0(x), y_0'(x))$ and $F_{y_0'}(x) = F_{y'}(x, y_0(x), y_0'(x))$ are the partial derivatives $\partial F/\partial y$ and $\partial F/\partial y'$, respectively, evaluated along the graph $y = y_0(x)$ of the minimizer. Next, let us define $\theta(x) = \int_a^x F_{y_0}(t)\, dt$. Since F_{y_0} is continuous (being a composition of the continuous partial derivative F_y with y_0 and y_0', which are both continuous), this anti-derivative is a continuously differentiable function. Using $\theta'(x) = F_{y_0}(x)$, we may integrate the first term in (4) by parts to rewrite the necessary condition as

$$\int_a^b \{-\theta(x) + F_{y_0'}(x)\}h'(x)\, dx = 0, \quad \text{for all } h \in C_0^1,$$

where the contribution from the boundary terms vanish because $h(a) = h(b) = 0$. We now appeal to the following famous result by Paul du Bois-Reymond, whose proof may be found in, e.g. Pars' book [**8**, section 2.1, Lemma 2].

[1] In the older literature, standard problems are often referred to as "problems of the simplest kind in the calculus of variations."

Lemma. *If $g : [a, b] \to \mathbb{R}$ is continuous and*

$$\int_a^b g(x)h'(x)\,dx = 0,$$

for all $h \in C^1([a, b])$ with $h(a) = h(b) = 0$, then $g(x) = C$ for all $x \in [a, b]$ for some constant C.

The lemma implies that

$$-\theta(x) + F_{y_0'}(x) = C$$

for some constant C. Since θ is differentiable and the right-hand side is constant, it follows that the function $F_{y_0'}$ is likewise differentiable. Differentiation on both sides of the equality gives Euler's equation (3), and the proof is complete.

This was the "sophisticated" proof of Euler's equation. The "standard proof" is slightly simpler, but comes at the price of an extra assumption: the minimizer y_0 needs to be twice continuously differentiable, and not just continuously differentiable as all the other members of $\mathcal{A}$ (and how do we know this?). It uses integration by parts in the second term of the necessary condition (4) and appeals to the so-called "fundamental lemma of the calculus of variations," see, for instance, [**6**, Lecture 2].

Finally, we remark that any solution of Euler's equation (3) is called an *extremal* of the functional $\mathcal{J}$. This term is unfortunate, of course, since a solution of Euler's equation is not necessarily a minimizer, nor a maximizer, of $\mathcal{J}$. An extremal need not even belong to $\mathcal{A}$. So when we say something like "y is an extremal of the functional $\mathcal{J}$" we simply mean that y solves the corresponding Euler equation.

The original solution

The minimizer of $\mathcal{E}$ in (2) subject to the constraint (1) was found in [**10**] by appealing to Euler's rule, which is the equivalent in the calculus of variations to the Lagrange multiplier rule of ordinary calculus. Euler's rule is a necessary condition for optimality in isoperimetrical problems, see, e.g. Pars [**8**, section 6.5] or Gelfand and Fomin [**3**, section 12.1]. For the present problem it states that if the design A_0 minimizes the elongation $\mathcal{E}[A]$ among all designs A satisfying the volume constraint (1), and if A_0 is not an extremal of $\mathcal{V}$, then A_0 is an extremal of the augmented functional,

$$\mathcal{E}[A] - \lambda\mathcal{V}[A],$$

where λ is the (Lagrange-)multiplier. Notice that $\mathcal{V}$ does not have any extremals.

Before attacking this problem, it was suggested in [**10**] to rewrite the above functional in terms of a new dependent variable B defined by

$$B(x) = \frac{W}{\rho g} + \int_0^x A(x')\,dx'. \tag{5}$$

Thus, $B(x)$ may be interpreted as $1/\rho g$ times the weight of the attached load together with the weight of the part of the rope which lies below the level x in the reference state. Using that $B' = A$ gives $\mathcal{E}[A] = (\rho g/E)\int_0^L B(x)/B'(x)\,dx$. Therefore, if we define

$$\hat{\mathcal{E}}[B] := \int_0^L F(B, B')\,dx \quad \text{where} \quad F(B, B') = B/B', \tag{6}$$

the problem becomes that of determining the extremals of the functional $(\rho g/E)\hat{\mathcal{E}}[B] - \lambda\mathcal{V}[B']$, that is, finding the solutions of Euler's equation:

$$\left(\frac{\partial}{\partial B} - \frac{d}{dx}\frac{\partial}{\partial B'}\right)\left(\frac{\rho g}{E}\frac{B}{B'} - \lambda B'\right) = 0.$$

Differentiation and simplification yields the following second-order differential equation,

$$BB'' = (B')^2.$$

After one integration this becomes $B' = KB$, for some real constant K, and after another integration, $B(x) = B(0)e^{Kx}$. The definition (5) implies $B(0) = W/\rho g$ and differentiation, $A = B'$, gives $A(x) = (KW/\rho g)e^{Kx}$. The constant K is determined using (1). The result is the admissible extremal:

$$A_0(x) = K_0\frac{W}{\rho g}e^{K_0 x}, \quad 0 \le x \le L, \tag{7}$$

where $K_0 = L^{-1}\ln(1 + V\rho g/W)$.

This is the only extremal that satisfies the volume constraint, and is therefore tacitly assumed in [10] to be the minimizer of the elongation. This may seem plausible, from a practical point of view. However, from the mathematical point of view the situation is unsatisfactory until optimality has been proved (or disproved). Our main aim is to provide the missing proof of optimality (in fact we propose three). However, textbooks on calculus of variations usually do not state nor prove any sufficient conditions for (local or global) optimality of extremals of isoperimetrical problems, so there is no off-the-shelf method to help us here. Instead we achieved our goal by observing that the problem can be reformulated as a standard problem in the calculus of variations. We thereby avoid the isoperimetrical constraint and the use of Euler's rule all together, and are able to give a simple proof of the optimality of A_0.

Before proceeding let us mention that in [10] it was observed that the stress corresponding to $A_0(x)$ is constant, $\sigma(x) = \rho g/K_0$ for $0 \le x \le L$. This observation was stated in the very last couple of lines and was not used. In structural optimization, a uniform stress distribution is usually taken as a sign that an admissible design is optimal.[2] That such a condition is sufficient for optimality may very well be true. It is, however, far from obvious.

The reformulation

Our new approach to the minimum elongation problem is again based on the introduction of the new dependent variable B in (5) and minimization of the elongation functional $(\rho g/E)\hat{\mathcal{E}}[B]$. But instead of using the condition (1) as a constraint, we observe that it implies the end point conditions $B(0) = W/\rho g$ and $B(L) = V + W/\rho g$ on the admissible functions B. This was, curiously enough, overlooked in [10].

The problem is now a standard problem in the calculus of variations, whose precise formulation we summarize here: Let $\mathcal{D}^1$ denote the space of piecewise continuously differentiable functions defined on the interval $[0, L]$. We say that a function f is piecewise differentiable on the interval $[a, b]$ if f is continuous on the interval and if there exists a positive integer N and numbers $a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b$, all dependent on f, such that f is continuously differentiable on each of the subintervals

[2] So the present author was told by a colleague at the Faculty of Engineering.

$[x_{i-1}, x_i]$, $i = 1, \ldots, N$, with finite one-sided derivatives at the end points. Set

$$\mathcal{A} = \{B \in \mathcal{D}^1 \mid B(x) > 0, B'(x) > 0 \text{ for } x \in [0, L] \text{ and } B(0) = \alpha, B(L) = \beta\},$$

where $\alpha = W/\rho g$ and $\beta = V + W/\rho g$. Notice that our choice of $\mathcal{A}$ permits designs $A = B'$ which have jump discontinuities. We want to find $B_0 \in \mathcal{A}$ which solves

$$\min_{B \in \mathcal{A}} \hat{\mathcal{E}}[B], \tag{8}$$

where $\hat{\mathcal{E}}$ is the functional defined in (6). We have dropped the inessential factor $\rho g/E$.

The optimality proof

Here comes the very short but complete solution to the reformulated problem (8) which implies that A_0 in (7) is the correct solution of the minimum elongation problem.

First, for any number $K > 0$ (whose exact value K_0 will be chosen at the end of the proof), observe that the integral

$$\mathcal{K}[B] := K^{-2} \int_0^L 2K - B'/B \, dx = K^{-2}[\, 2Kx - \ln(B(x))\,]_0^L$$

depends only on the value of B at the interval end points, and therefore has the same value for all $B \in \mathcal{A}$. It follows that the functional

$$\mathcal{I}[B] := \hat{\mathcal{E}}[B] - \mathcal{K}[B] = \frac{1}{K} \int_0^L \frac{KB}{B'} + \frac{B'}{KB} - 2 \, dx \tag{9}$$

has the same minimizer in $\mathcal{A}$ as $\hat{\mathcal{E}}$. A special case of the arithmetic-geometric inequality states that $a + a^{-1} \geq 2$, $a > 0$, where equality holds if and only if $a = 1$. This implies that the integrand of $\mathcal{I}$ is nonnegative, and therefore $\mathcal{I}[B] \geq 0$ for all $B \in \mathcal{A}$. If we can find $B_0 \in \mathcal{A}$ such that $\mathcal{I}[B_0] = 0$, then B_0 is a minimizer of $\mathcal{I}$, and therefore a solution of (8). But $\mathcal{I}[B_0] = 0$ is possible if and only if the integrand of (9) is identically zero on the interval $[0, L]$. By the equality case in the arithmetic-geometric inequality, $B_0' = KB_0$ throughout the interval, and therefore $B_0(x) = B_0(0)e^{Kx}$. The end point conditions $B(0) = \alpha$ and $B(L) = \beta$ are satisfied if we take K to be $K_0 = L^{-1} \ln(\beta/\alpha)$. It follows that

$$B_0(x) = \alpha e^{K_0 x} \in \mathcal{A} \tag{10}$$

is the unique solution of (8) and differentiation, $A_0 = B_0'$, leads to the optimal design solution (7).

How to find the optimality proof

As shown in the previous section, a logically complete solution of (8) may be given without first appealing to Euler's equation. Such a solution depends, however, upon guessing the form of the integral $\mathcal{K}$ correctly, which may not always be easy. It is therefore interesting to see how an algorithm, due to Hilbert,[3] allows us to find $\mathcal{K}$ if we know sufficiently many of the extremals of the problem, and thus prove optimality.

[3] Hilbert's method appeared in his famous lecture delivered at the International Congress of Mathematicians in Paris, 1900. A written version of the lecture was published as two papers (in German) in two different German journals. An English translation [4], where precise references to the two German originals may also be found, was published by the American Mathematical Society in 1902.

We first determine all the extremals of (8). Our derivation, which is inspired by similar ones found in Bliss [1], has the advantage that we immediately get that minimizers are smooth, and not merely piecewise smooth, functions. Assume B_0 is a solution to the minimization problem. The necessary condition for optimality in (4), which can be extended to piecewise differentiable functions, states that $\int_0^L F_{B_0}h + F_{B_0'}h'\,dx = 0$ for all $h \in \mathcal{D}_0^1 := \{\eta \in \mathcal{D}^1 \mid \eta(0) = \eta(L) = 0\}$. Here $\mathcal{D}_0^1$ has replaced C_0^1 as our set of test functions. Since $F = B/B'$, the necessary condition becomes

$$0 = \int_0^L \frac{h}{B_0'} - \frac{B_0 h'}{{B_0'}^2}\,dx = \int_0^L \frac{B_0^2}{{B_0'}^2}\Big(-\frac{h}{B_0}\Big)'\,dx$$

for all $h \in \mathcal{D}_0^1$. Since $B_0 > 0$ on $[0, L]$ we see that every member $\eta \in \mathcal{D}_0^1$ can be written in the form $\eta = -h/B_0$ for some $h \in \mathcal{D}_0^1$. It follows that $\int_0^L (B_0/B_0')^2\eta'\,dx = 0$ for all such η. The lemma of du Bois-Reymond, stated earlier, implies that $(B_0/B_0')^2$ is a constant throughout the interval, that is, $B_0' = KB_0$ on $[0, L]$ for some number $K > 0$. Since B_0 is continuous, so is B_0', hence B_0 is continuously differentiable and it immediately follows that the minimizer of the functional has to be found among functions of the form

$$B_0(x) = Ce^{Kx}, \tag{11}$$

where $C > 0$ and $K > 0$. When the endpoint conditions are taken into account, we find that the admissible extremal (*i.e.*, the one in $\mathcal{A}$) is precisely the function (10).

The functional $\mathcal{K}[B]$, used in the optimality proof, is known as Hilbert's invariant integral, since its value is the same for all $B \in \mathcal{A}$. We are now going to indicate, briefly, how the extremals of the elongation functional are used in the construction of this integral. The method is general and may be applied to any Lagrangian F.

First, we fix the value of K to be $K_0 = L^{-1}\ln(\beta/\alpha)$, see (10), and consider the subset of the extremals (11) given by

$$B(x; C) = Ce^{K_0 x}, \quad C > 0. \tag{12}$$

This subset is called a field of extremals because it has the following properties: It is a one-parameter family of functions. Each member of the family is an extremal of the functional, and through each point (x, B) of the domain $0 \le x \le L$ and $B > 0$ there passes exactly one such extremal. It is easy to verify that each member of the field satisfies a first-order differential equation of the form

$$y' = p(x, y(x)).$$

In fact, each member of the field (12) satisfies the linear first-order, constant coefficient differential equation $B' = K_0 B$, that is, $p(x, B) = K_0 B$.

Hilbert's invariant integral $\mathcal{K}$ is defined by

$$\mathcal{K}[B] = \int_0^L M(x, B(x)) + B'(x)N(x, B(x))\,dx, \tag{13}$$

where the functions M, N are defined in terms of the Lagrangian $F(B, B')$ as

$$M(x, B) = F(B, p(x, B)) - p(x, B)F_{B'}(B, p(x, B))$$

and

$$N(x, B) = F_{B'}(B, p(x, B)).$$

It can be shown that for this choice of M and N, the integrand of $\mathcal{K}$ becomes an exact differential, see Mesterton-Gibbons [6, Lecture 12] or any of the other textbooks in

the references. This means that there exists a function $U = U(x, B)$ such that for any function $y = y(x)$,

$$\frac{d}{dx}U(x, y(x)) = M(x, y(x)) + y'(x)N(x, y(x)).$$

As a consequence $\mathcal{K}[B] = U(L, B(L)) - U(0, B(0))$, which depends only on the end point values of B. This is the desired invariance.

For the Lagrangian $F = B/B'$, we have $F_{B'} = -B/B'^2$, and, since $p(x, B) = K_0 B$, the above definitions give

$$M = \frac{B}{p(x, B)} - p(x, B)\frac{-B}{p(x, B)^2} = \frac{2B}{p(x, B)} = \frac{2}{K_0},$$

and

$$N = \frac{-B}{p(x, B)^2} = \frac{-1}{K_0 B^2}.$$

If these expressions are substituted into (13), we recover the invariant integral of the optimality proof. Solving the system $\partial U/\partial x = M$, $\partial U/\partial B = N$ leads to $U(x, B) = K_0^{-2}(2K_0 x - \ln(B))$ (plus a constant, chosen here to be zero), which we also recognize from the optimality proof. Notice that such an explicit expression for U is not really needed once the invariance has been established by other means. All that is needed to complete the optimality proof are M and N.

Finally, notice that the admissible extremal B_0 found in (10) is one of the extremals in the field (12) and that $\mathcal{K}[B_0] = \hat{\mathcal{E}}[B_0]$. The latter fact is easily checked by substituting the relation $B_0' = p(x, B_0)$ into the definition of $\mathcal{K}$. Now, if the integrand of the functional $\mathcal{I} = \hat{\mathcal{E}} - \mathcal{K}$ is nonnegative,[4] as is the case when $F = B/B'$, then $\mathcal{I}[B] \geq 0$ for all admissible B. The functional $\mathcal{I}$ is therefore clearly minimized by B_0. The invariance of $\mathcal{K}$ now ensures the success of our method because it implies that a minimizer of $\mathcal{I}$ is also a minimizer of $\hat{\mathcal{E}}$.

Alternative solutions

The functional $\int_0^L F(B, B')\,dx$ is written in standard from, but the actual Lagrangian $F(B, B') = B/B'$ is far from common. It does not occur in any of the examples or problems in such classical texts on the calculus of variations as Elsgolc [2], Pars [8], Gelfand and Fomin [3], and Sagan [9] nor in the modern introductions to the subject Kot [5] and Mesterton-Gibbons [6]. Placing two conditions, $B > 0$ and $B' > 0$, on the admissible functions in $B \in \mathcal{A}$ is not common either. However, the special form of F allows us to come up with two alternative (and opportunistic) solutions of the minimum elongation problem.

The first alternative solution uses that members of $\mathcal{A}$ are positive and strictly increasing. The only differential calculus rule which involves a quotient with a derivative in the denominator is the one for differentiation of the inverse of a function. This suggests that we apply the substitution $x = u(t) := B^{-1}(t)$ in the integral which defines $\hat{\mathcal{E}}$. It then turns out that the problem (8) is equivalent to the minimization of the integral

$$\int_\alpha^\beta t u'(t)^2\,dt$$

[4] This integrand is known as Weierstrass' excess function.

over the set of piecewise differentiable functions u with $u > 0$ and $u' > 0$ on the interval $[\alpha, \beta]$ which satisfy the end point-conditions $u(\alpha) = 0$ and $u(\beta) = L$. The admissible extremal of this problem is

$$u_0(t) = L\frac{\ln(t/\alpha)}{\ln(\beta/\alpha)},$$

whose inverse coincides with B_0 in (10). The optimality of u_0 is easily established by direct verification, that is, by showing the identity $\int_\alpha^\beta tu'^2\,dt - \int_\alpha^\beta tu_0'^2\,dt = \int_\alpha^\beta t(u' - u_0')^2\,dt$, for all admissible u, and use that the right-hand side is nonnegative. We leave the details, including those in the optimality proof, to the student as a means for further study. One only needs Euler's equation to do this.

The second alternative solution of (8) is based on the ansatz $B(x) = \alpha\exp(u(x))$ which transforms the problem into minimization of the integral $\tilde{\mathcal{E}}(u) = \int_0^L 1/u'(x)\,dx$ over piecewise differentiable functions satisfying $u' > 0$, $u(0) = 0$, and $u(L) = \ln(\beta/\alpha) =: \gamma$. If we use that $\gamma = \int_0^L u'(x)\,dx$, it follows from Schwarz' inequality that $\gamma\tilde{\mathcal{E}}(u) \geq L^2$. Equality occurs if and only if u' and $1/u'$ are linearly dependent, which implies that u' is constant. This allows us to determine the optimal u and leads to the solution (10). (Observe that the stress in the rope is $\sigma(x) = \rho g/u'(x)$, so in this formulation of the problem, constancy of stress turns out to imply the optimality of an admissible design. Compare this to the remark made at the end of our presentation of the original solution.) Once more we leave the details to the reader as an instructive exercise.

We close this section with the remark that the minimum elongation problem can be generalized to ropes with variable density and nonlinear stress-strain relations, see [7]. For nonconstant density two distinct elongation problems emerges; one with a constant mass-constraint and another with a constant volume-constraint. The above ansatz can be used to simplify the analysis of the first of these two problems considerably (compared to [7].) The second problem seems to be much harder.

Concluding remarks

The purpose of the present paper is pedagogical and it has been written with the beginning student of the calculus of variations in mind. For this reason we did not hesitate to give the original derivation as well as three new complete solutions of the minimum elongation problem. The minimum elongation problem has some nice features: It involves, as already mentioned, a variational problem in standard form with an uncommon Lagrangian function. Moreover, the (first) optimality proof, which is basically an application of Weierstrass' sufficient condition for a strong relative minimum, is simple but not trivial, as a direct verification of the optimality of B_0 does not seem possible, or at least not easy to find. Finally, the methods of the third solution seem capable of generalization to more complicated minimum elongation problems.

I have been teaching a course on calculus of variations at my home university for a couple of years now. In the latest realization of this course I decided to hand out a set of journal articles for the students to read and present orally. Most of these papers consider relatively simple real-world applications formulated as variational problems and the solutions are based on methods taught in the course. The inspiring paper by Verma and Keller was one of these articles. It is my hope that the present paper may serve future students and teachers of the calculus of variations as [10] has served me and my pupils.

Acknowledgments The author wishes to thank his colleagues Tomas Persson, Mikael Persson Sundqvist, and Anders Holst, the former two for reading and offering their comments on the first draft of this paper and the latter for suggesting the elegant second alternative solution.

REFERENCES

[1] Bliss, G. A. (1925). *Calculus of Variations*. The Carus Mathematical Monographs. Vol. 1. LaSalle, Illinois: Mathematical Association of America.
[2] Elsgolc, L. E. (1961). *Calculus of Variations*. Oxford: Pergamon Press.
[3] Gelfand, I. M., Fomin, S. V. (1963). *Calculus of Variations*. Mineola, New York: Prentice-Hall.
[4] Hilbert, D. (1902). Lecture delivered before the international congress of mathematicians at Paris in 1900. *Bull. Am. Math. Soc.* VIII(2):473–479.
[5] Kot, M. (2014). *A First Course in the Calculus of Variations*, Student Mathematical Library. Vol. 72. Providence: American Mathematical Society.
[6] Mesterton-Gibbons, M. (2009). *A Primer on the Calculus of Variations and Optimal Control Theory*, Student Mathematical Library. Vol. 50. Providence: American Mathematical Society.
[7] Negron-Marrero, P. V. (2003). The hanging rope of minimum elongation for a nonlinear stress-strain relation. *J. Elasticity* 71:133–155.
[8] Pars, L. A. (1962). *An Introduction to the Calculus of Variations*. London: Heinemann Educational Books Ltd.
[9] Sagan, H. (1969). *Introduction to the Calculus of Variations*. New York: McGraw-Hill.
[10] Verma, G. R., Keller, J. B. (1984). Hanging rope of minimum elongation. *SIAM Rev.* 26(4):569–571.

Summary. We reconsider the variational problem of finding the shape of a vertically hanging rope such that its elongation, due to the rope's own weight and that of a load attached at its lower end, is minimum. The known solution is recalled and the missing proof of optimality is supplied.

NIELS CHRISTIAN OVERGAARD (MR Author ID: 804032) Grew up in Denmark where he earned his M.Sc. in mathematics and physics from Aalborg University. He moved to Sweden where he received a Ph.D. in mathematics from Lund University in 2002. His mathematical interests include application of variational calculus to image analysis problems and PDE models for bacterial colony growth. In his spare time he enjoys cycling and has taken part in several cyclosportives such as the local 300 km *Vätternrundan* or the 430 km *Jothunheimen Runt* in Norway.

Tiling One-Deficient Rectangular Solids with Trominoes in Three and Higher Dimensions

ARTHUR BEFUMO
Yale University
New Haven, CT 06520
arthur.befumo@yale.edu

JONATHAN LENCHNER
IBM Research Africa
Catholic University of East Africa
Nairobi 00100 Kenya
jonathan.lenchner@ke.ibm.com

The polyomino, a term coined by Solomon Golomb in 1953–1954 [**6**] and later popularized by Martin Gardner in one of his *Scientific American* columns of 1960 [**5**], has long been a popular topic in recreational mathematics, and has even made its way into pop culture through games like Tetris and Blokus. A polyomino is a finite collection of edge-connected, equal-sized squares in the plane. Problems often involve tiling the entire plane, or regions of it, using certain sets of polyominoes [**6–8**]. Perhaps the most famous theorem involving polyominoes is Solomon Golomb's tromino theorem, which also happens to be an incredibly elegant example of inductive reasoning. A tromino is a polyomino consisting of just three squares. Golomb's theorem, first proved in [5], states that any "chess board" of size $2^N \times 2^N$ with a single square removed can be completely tiled by trominoes of the type shown in Figure 1,which we will call

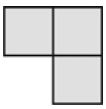


Figure 1 The L-tromino.

an L-tromino. While this fact may not be entirely obvious at first glance, the proof is so simple and beautiful it bears repeating. The case $N = 0$ is trivial, since after removing the one square there is nothing to cover. Now consider a $2^N \times 2^N$ board for $N \geq 1$, and divide the board into a 2×2 array of $2^{N-1} \times 2^{N-1}$ boards as shown in Figure 2(a). Any square removed from the board must fall in one of the four smaller $2^{N-1} \times 2^{N-1}$ boards. Without loss of generality, we suppose it is the board on the bottom left. Now place an L-tromino as shown in Figure 2(b) and complete the proof by tiling each of the $2^{N-1} \times 2^{N-1}$ boards, now with one tile missing from each, by induction.

In 1985–1986, Chu and Johnsonbaugh [**3**,**4**] generalized this result to show that as long as 3 does not divide evenly into K and $K \neq 5$, then if you remove a square from a $K \times K$ board, the resulting board can always be tiled by L-trominoes. They also characterized which rectangular boards can be so tiled. They called boards with tiles removed **deficient**, and specifically, boards with one tile removed **1-deficient**. Additionally, we shall say that if a board (deficient or otherwise) is tilable by L-trominoes then it is **L-tilable**.

Math. Mag. **91** (2018) 62–69. doi:10.1080/0025570X.2017.1404796
MSC: Primary 52C22, Secondary 05B45, 05B50

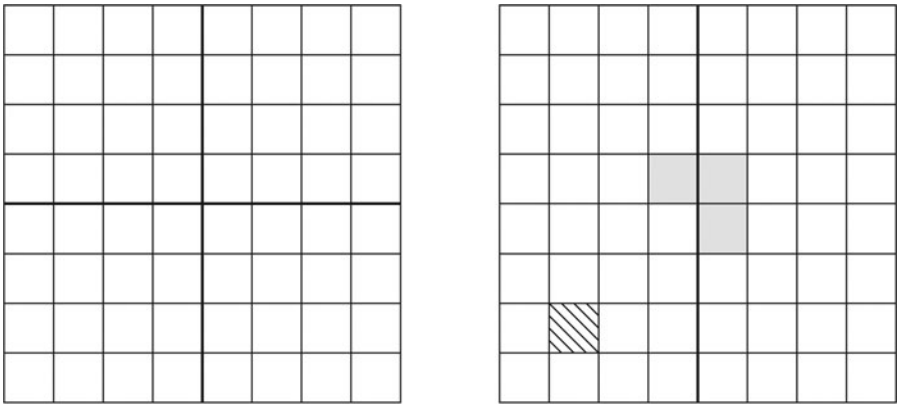


Figure 2 (a) A $2^N \times 2^N$ board divided into 2×2 array of $2^{N-1} \times 2^{N-1}$ boards. (b) After a square has been removed from the bottom-left $2^{N-1} \times 2^{N-1}$ board, a tromino is placed such that a square is taken from each of the other three $2^{N-1} \times 2^{N-1}$ boards. Each of the smaller boards may then be covered by basic trominoes by induction.

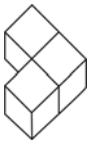


Figure 3 The solid L-tromino.

In 2008, Starr went on to study the problem in three dimensions, first showing that 1-deficient cubical boards of edge length 2^N for any N are tilable by solid L-trominoes (Figure 3) [**10**], and then generalizing the result to show that 1-deficient cubical boards of arbitrary edge length K with $K \equiv 1 (\text{mod } 3)$ are L-tilable [**9**].

In this paper, we show that any 1-deficient rectangular solid of dimensions $K \times L \times M$, where $KLM \equiv 1 (\text{mod } 3)$ and $K, L, M > 1$, is L-tilable. The proof is very simple and as a corollary gives a simplified proof of Starr's result for 1-deficient cubical boards. In addition, we extend the result to all higher dimensions showing that one may always tile a 1-deficient $K_1 \times \cdots \times K_N$ board with $K_1 \cdots K_N \equiv 1 (\text{mod } 3)$, $N \geq 3$, where at least three of the $K_i > 1$.

The only other tromino (solid or otherwise), in addition to the L-tromino, is the straight tromino. To complete our study, we show that, in contrast to the case of the L-tromino, it is *never* possible to cover a 1-deficient $K_1 \times \cdots \times K_N$ board, where $K_1 \cdots K_N \equiv 1 (\text{mod } 3)$, and at least *one* of the $K_i > 1$, with straight trominoes if we are free to choose the square/cube/hypercube to be removed.

L-trominoes in 3D

In what follows we shall refer to a generic 1-deficient two-dimensional board of size $K \times L$ as a $(K \times L) - 1$ board and a 1-deficient three-dimensional board of size $K \times L \times M$ as a $(K \times L \times M) - 1$ board. Moreover, a $K \times L$ board has K rows and L columns, while a $K \times L \times M$ board has K levels, L rows, and M columns. We say that a board of size $(K \times L) - 1$ or $(K \times L \times M) - 1$ is **generically L-tilable** if it is L-tilable regardless of the square or cube removed.

In the arguments that follow, note that by taking slices in each of the different axis-aligned directions, we may freely regard a $K \times L \times M$ board as the union of K boards of size $L \times M$, L boards of size $K \times M$, or M boards of size $K \times L$.

Lemma 1. *If $K, L, M > 1$ then if a two-dimensional board of size $(KL \times M) - 1$ is generically L-tilable then so is a three-dimensional $(K \times L \times M) - 1$ board.*

Proof. For illustration purposes consider a $2 \times 2 \times 4$ board. We picture this board as two 2×4 boards, one on "top" and one on the "bottom," as in Figure 4. If we slide the

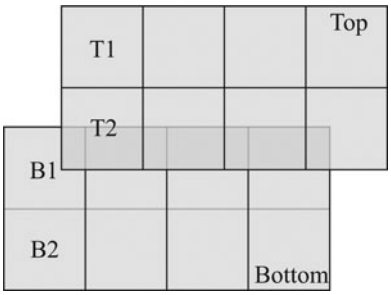


Figure 4 A $2 \times 2 \times 4$ board pictured as two 2×4 boards, one on "top" and one on the "bottom."

board on the bottom to the left, flip it horizontally, and glue the edges together we get a 2×8 board as pictured in Figure 5. In the 2×8 board note that an L-tromino

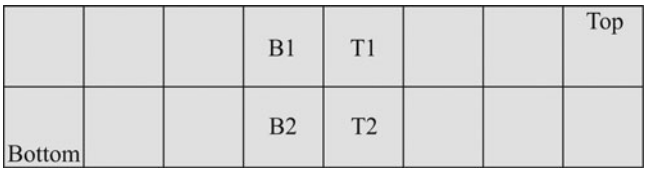


Figure 5 After sliding the "bottom" board in Figure 4 to the left, flipping horizontally, and gluing the adjacent edges one is left with the above 2×8 board.

that crosses the middle vertical line corresponds to an L-tromino that extends from the top to the bottom in the left-most column in the original 3D board. It easily follows that any tiling by L-trominoes of a $(2 \times 8) - 1$ board corresponds to an analogous tiling of the associated $(2 \times 2 \times 4) - 1$ board. Hence, if a $(2 \times 8) - 1$ board is generically tilable then so is a $(2 \times 2 \times 4) - 1$ board. If we instead had a $5 \times 2 \times 4$ board, say, then we would think of this board as 5 boards of size 2×4 and analogously lay them out right to left, starting with the top one on the right, and would then flip every other one horizontally, beginning with the second from the top. In the resultant 2×20 board, if we number the vertical lines in left to right order with the numbers 1 through 21, then any L-tromino that crosses vertical line $1 + 4j$ for $1 \leq j < 5$ corresponds to an L-tromino that would have spanned vertically adjacent boards in the left-most column. Laying out the K boards of size $L \times M$ in the statement of the lemma in analogous fashion, we see that any L-tiling of an $(L \times KM) - 1$ board corresponds to L-tiling of a $(K \times L \times M) - 1$ board and the lemma is proved. ■

We will also need the following result that readily follows from [1].

Theorem 1 (Ash and Golomb [1]). *Any $(K \times L) - 1$ board with $KL \equiv 1 (\mathrm{mod}\ 3)$ and $K, L > 1$ is L-tilable as long as $K, L \notin \{2, 5\}$.*

Theorem 2. *A $(K \times L \times M)$-1 board, where $KLM \equiv 1 (\mathrm{mod}\ 3)$ and $K, L, M > 1$ is always L-tilable.*

Proof. Suppose we have a $(K \times L \times M) - 1$ board as in the statement of the theorem. Note that we cannot have all of $K, L, M \in \{2, 5\}$ since then $KLM \equiv 2 (\mathrm{mod}\ 3)$. Without loss of generality we may assume that $M \notin \{2, 5\}$. In addition, it must be that $KL \notin \{2, 5\}$. It follows by Theorem 1 that any board of size $(KL \times M) - 1$ is L-tilable and by Lemma 1 that our $(K \times L \times M)$-1 board is L-tilable. The theorem follows. ■

L-trominoes in higher dimensions

Let us first consider the 4D case, and for illustrative purposes, consider a $2 \times 2 \times 2 \times 5$ board. First, we consider a $2 \times 2 \times 5$ board, which we think of as two parallel 2×5 boards, one on "top" and one on "bottom." See Figure 6. To get the $2 \times 2 \times 2 \times 5$

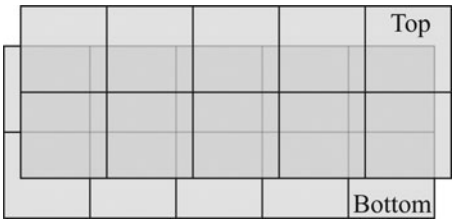


Figure 6 A $2 \times 2 \times 5$ board, thought of as two parallel 2×5 boards, one on "top" and one on "bottom."

board, take two identical copies of the two parallel 2×5 boards and consider each cube (which we are picturing as squares) on one of the $2 \times 2 \times 5$ boards to be connected to the identically positioned cube in the other $2 \times 2 \times 5$ board. In other words, each square in the top 2×5 board is thought of as being connected to each square on the associated top 2×5 board, and analogously for the bottom squares. If we reorder the component 2×5 boards of the top $2 \times 2 \times 5$ board (so that the former "top" 2×5 board is now on the bottom) and slide this reordered $2 \times 2 \times 5$ board on top of the other one, we get a $4 \times 2 \times 5$ board, as in Figure 7.

This board is different from the $2 \times 2 \times 2 \times 5$ 4D board in some ways, and in some ways similar. For example, the corresponding squares on the associated "top" layers are connected, but those on the associated bottom layers are not. However, this embedding of the 4D $2 \times 2 \times 2 \times 5$ board onto the 3D $4 \times 2 \times 5$ board gives rise to a bijective mapping ϕ of cells where

$$\phi(i, j, k, l) = (i, j, k + (l - 1)(5 - 2k)), \tag{1}$$

or in a form that is easier to generalize to boards with other dimensions:

$$\phi(i, j, k, l) = \begin{cases} (i, j, k), & \text{if } l = 1, \\ (i, j, 5 - k), & \text{if } l = 2. \end{cases} \tag{2}$$

The cell c with coordinates (i, j, k, l) is in the ith row, jth column, kth level, and lth hyper-layer. The map ϕ has the key properties:

(i) if cells c, c' are adjacent in the $4 \times 2 \times 5$ board, then $\phi^{-1}(c), \phi^{-1}(c')$ are adjacent in the $2 \times 2 \times 2 \times 5$ board, and

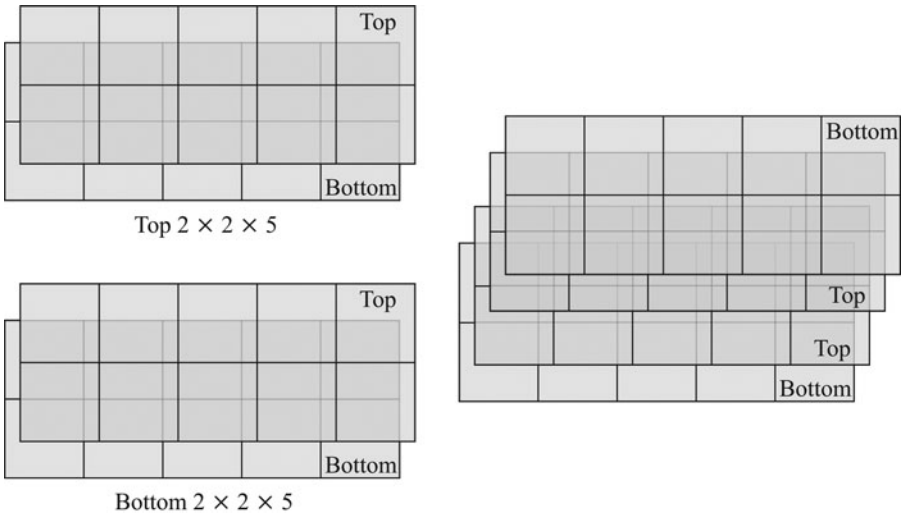


Figure 7 On the left a $2 \times 2 \times 2 \times 5$ board thought of as two parallel $2 \times 2 \times 5$ boards with squares in corresponding locations thought of as being connected to one another. If we reorder the boards of the "top" $2 \times 2 \times 5$ board and slide them over the "bottom" $2 \times 2 \times 5$ board, we get an embedding ϕ of the $2 \times 2 \times 2 \times 5$ board into a $4 \times 2 \times 5$ board. If L denotes three cells that comprise an L-tromino in the $4 \times 2 \times 5$ board then $\phi^{-1}(L)$ is also an L-tromino in the $2 \times 2 \times 2 \times 5$ board.

(ii) if cells c, c', c'' are not collinear in the $4 \times 2 \times 5$ board then $\phi^{-1}(c)$, $\phi^{-1}(c')$, $\phi^{-1}(c'')$ are not collinear in the $2 \times 2 \times 2 \times 5$ board.

Hence, the inverse image of cells forming an L-tromino on the $4 \times 2 \times 5$ board form an L-tromino on the $2 \times 2 \times 2 \times 5$ board. As a result, tiling of the $4 \times 2 \times 5$ board by L-trominoes gives rise, under the bijection ϕ^{-1}, to a tiling of the $2 \times 2 \times 2 \times 5$ board. There was nothing special about our choice of a $2 \times 2 \times 2 \times 5$ board. For a generic $L \times K \times J \times I$ board, in other words where $i \in \{1, \ldots, I\}$, $j \in \{1 \ldots, J\}$, $k \in \{1, \ldots, K\}$, and $l \in \{1, \ldots, L\}$, the analog of equation (2) is

$$\phi(i, j, k, l) = \begin{cases} (i, j, k + (l-1)K), & \text{if } l \text{ is odd,} \\ (i, j, lK - k + 1), & \text{if } l \text{ is even.} \end{cases} \tag{3}$$

Now ϕ is a bijection between the cells of an $L \times K \times J \times I$ board and an $LK \times J \times I$ board. For odd values of the hyper-parameter l, the boards are layered from bottom to top, while for even values of l the boards are layered from top to bottom. As before, the inverse image of cells forming an L-tromino on the $LK \times J \times I$ board form an L-tromino on the $L \times K \times J \times I$ board. Hence, Theorem 2 immediately gives way to the analogous proof for 4D boards.

More generally, for a board of dimensions $K_N \times \cdots \times K_1$, $N \geq 4$, and cell with generic coordinates $c = (i_1, \ldots, i_N)$, the embedding ϕ will be a composition of embeddings $\phi_4 \circ \cdots \circ \phi_N$ where

$$\phi_N(i_1, \ldots, i_N) = \begin{cases} (i_1, \ldots, i_{N-2}, i_{N-1} + (i_N - 1)K_{N-1}), & \text{if } i_N \text{ is odd,} \\ (i_1, \ldots, i_{N-2}, i_N K_{N-1} - i_{N-1} + 1), & \text{if } i_N \text{ is even.} \end{cases} \tag{4}$$

ϕ_N then embeds the $K_N \times \cdots \times K_1$ board bijectively onto the associated $K_N K_{N-1} \times K_{N-2} \times \cdots \times K_1$ board, and preserves L-trominoes under ϕ_N^{-1}.

The composition $\phi = \phi_4 \circ \cdots \circ \phi_N$ thus is a bijective embedding of the $K_N \times \cdots \times K_1$ board onto a $(K_N \cdots K_3) \times K_2 \times K_1$ board that preserves L-trominoes under ϕ^{-1}. Hence, Theorem 2 yields the following result.

Theorem 3. *A $(K_1 \times \cdots \times K_N)$-1 board, where $K_1 \cdots K_N \equiv 1 (\mathrm{mod}\ 3)$, for $N \geq 3$ and some three of the $K_i > 1$, is always L-tilable.*

Tiling with straight trominoes

It has often been noted, for example, in [7,8], that the standard 8×8 chessboard cannot be tiled by straight trominoes if an arbitrary square is removed. The argument follows by 3-coloring the chessboard as in Figure 8. For clarity of exposition, we have

0	1	2	0	1	2	0	1
2	0	1	2	0	1	2	0
1	2	0	1	2	0	1	2
0	1	2	0	1	2	0	1
2	0	1	2	0	1	2	0
1	2	0	1	2	0	1	2
0	1	2	0	1	2	0	1
2	0	1	2	0	1	2	0

Figure 8 3-coloring of a standard 8×8 chessboard. There is one more square labeled "0" than "1" or "2". A straight tromino necessarily covers squares with each of the three numbers. Considering symmetries we then easily see that removing any square other than the ones highlighted results in a board that cannot be covered by straight trominoes.

chosen to use the numbers 0, 1, and 2 to designate the three colors. The 3-coloring is performed by proceeding in diagonal bands starting at the bottom left. The reason we begin with a 2 will become apparent somewhat later; for the time being, take this to be an arbitrary choice. Note that any straight tromino necessarily covers a tile of each of the three colors. Since the total number of 0's is one more than the total number of 1's or 2's, if we remove a square numbered 1 or 2, we obviously cannot tile the remainder of the board with straight trominoes. Symmetry arguments allow us to rule out tiling with straight trominoes except in the case of the highlighted squares. Though in what follows we shall not care about such exceptional cases that may be tilable, it is worth noting that removal of such tiles does indeed enable a tiling, as shown in Figure 9.

An analogous impossibility of tiling argument shows that any rectangular board of dimension $K \times L$ with $KL \equiv 1 (\mathrm{mod}\ 3)$, and one of $K, L > 1$ cannot be tiled with straight trominoes, regardless of square removed. We again 3-color the board in diagonal bands, note that there necessarily are more of one color than another (since 3 does not divide evenly into KL), and make sure to remove a square of the less colored variety.

Interestingly, such a diagonal coloring is also possible and leads to the same conclusion in three and all higher dimensions. Figure 10 demonstrates such a coloring in 3D, in the case of a $4 \times 4 \times 4$ board. We now reveal the method to our coloring. In any dimension, let the cell at location $(i_1, \ldots, i_N)$ be colored with color $i_1 + \cdots + i_N (\mathrm{mod}\ 3)$.

It is immediate that any straight tromino covers a cell of each color, and moreover that removing a cell of a least used color necessarily results in an uncoverable board.

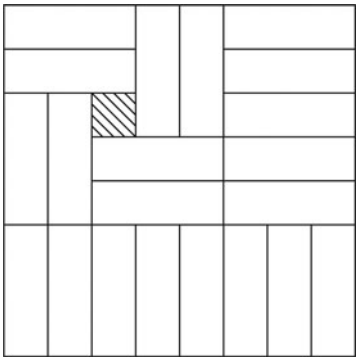


Figure 9 Tiling the standard 8×8 chessboard with trominoes after one of the highlighted squares from Figure 8 has been removed.

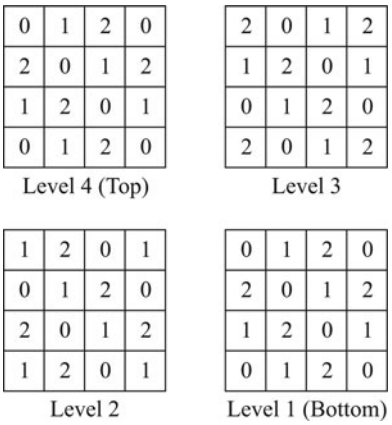


Figure 10 Three coloring of a generic 3D board as exemplified on a $4 \times 4 \times 4$ board.

Hence, we can state the following theorem.

Theorem 4. *A* $(K_1 \times \cdots \times K_N) - 1$ *board, where* $K_1 \cdots K_N \equiv 1 (\mathrm{mod}\ 3)$*, and at least* one *of the* $K_i > 1$*, is never generically tilable by straight trominoes (in other words, tilable regardless of the cell removed).*

In fact, using a K-coloring rather than a 3-coloring, *i.e.,* coloring cell $(i_1, \ldots, i_N)$ with color $i_1 + \cdots + i_N (\mathrm{mod}\ K)$, we see that the conclusion of Theorem 4 remains true for more than just trominoes.

Theorem 5. *A* $(K_1 \times \cdots \times K_N) - 1$ *board, where* $K_1 \cdots K_N \equiv 1 (\mathrm{mod}\ K)$*, and at least* one *of the* $K_i > 1$*, is never generically tilable by straight polyominoes of length* $K \geq 2$ *(in other words, tilable regardless of the cell removed).*

Concluding remarks

In this paper, we have presented an exhaustive study of covering 1-deficient boards with L-shaped and straight trominoes. It is natural to next study tiling 2-deficient boards

with L-trominoes. Ash and Golomb [1] initiated such a study for rectangles in the plane, while Starr [9] studied these problems for the case of 3D cubical boards. We conjecture that for any K, it is always possible to L-tile a K-deficient $K_1 \times \cdots \times K_N$ board, with all $K_i > 1$ and $K_1 \cdots K_N \equiv 1 (\mathrm{mod}\ 3)$, as long as N is sufficiently large.

Theorem 5 can clearly be extended to show that tiling K-deficient boards is not, in general, possible for straight trominoes. However, if we remove a cell of the necessary color so that an equal number of cells of each color remain regardless of what symmetry operation is performed on the board, do we always get a board that is tilable by straight trominoes?

Finally, in [2] it is shown that one can tile $(3^N \times 3^N) - 1$ boards with three specially chosen tetrominoes and $(4^N \times 4^N) - 1$ boards with three specially chosen pentominoes. Can these same families of tetrominoes and pentominoes tile more general 1-deficient boards in two, three, or possibly higher dimensions?

Acknowledgments The authors gratefully acknowledge the anonymous referees, one for suggesting a substantial simplification of our main argument, and the other for helping them improve the presentation.

REFERENCES

[1] Ash, J. M., Golomb, S. W. (2003). Tiling deficient rectangles with trominoes. *Math. Magazine* 77(1):46–55.
[2] Befumo, A., Lenchner, J. (2014). Extensions of Golomb's tromino theorem. Presented at the Fall Workshop in Computational Geometry, University of Connecticut, Storrs. fwcg14.cse.uconn.edu/program/wp-content/uploads/sites/863/2014/10/fwcg2014_submission_10.pdf.
[3] Chu, I.-P., Johnsonbaugh, R. (1985-86). Tiling boards with trominoes. *J. Recreat. Math.* 18:188–193.
[4] Chu, I.-P., Johnsonbaugh, R. (1986). Tiling deficient boards with trominoes. *Math. Magazine* 59:34–40.
[5] Gardner, M. (1960). More about the shapes that can be made with complex dominoes. *Sci. Am.* 203(5):186–194.
[6] Golomb, S. W. (1954). Checker boards and polyominoes. *Am. Math. Monthly* 61:675–682.
[7] Golomb, S. W. (1996). *Polyominoes: Puzzles, Patterns, Problems , and Packings*. 2nd ed. Princeton, NJ: Princeton University Press.
[8] Martin, G. E. (1991). *Polyominoes. A Guide to Puzzles and Problems in Tiling*. Washington, D.C.: Mathematical Association of America.
[9] Starr, N. (2008). Tromino tiling deficient cubes of any side length. http://arxiv.org/abs/0806.0524.
[10] Starr, N. (2008). Tromino tiling deficient cubes of side length 2^n. *Geombinatorics* XVIII(2):72–87.

Summary. A classic theorem of Solomon Golomb's states that if you remove a square from a chess board of size $2^N \times 2^N$ then the resulting board can always be tiled by L-shaped trominoes (polyominoes of three squares). We show that if you remove a cube (hyper-cube) from a board of size $K_1 \times \cdots \times K_N$, where $K_1 \cdots K_N \equiv 1 (\mathrm{mod}\ 3)$, for $N \geq 3$, and at least three of the $K_i > 1$, then the remaining board can always be tiled by solid L-shaped trominoes. This extends 2D results of Chu and Johnsonbaugh from the 80s and results of Starr's on 3D cubical boards from 2008. We also study the analogous problem for straight trominoes, showing that the same types of boards are never generically tilable (*i.e.*, tilable regardless of square/cube/hypercube removed) using straight trominoes.

ARTHUR BEFUMO (MR Author ID: 1183745) at age 14 simultaneously enrolled in the University of Montana and Hellgate High School, opting to take all of his math classes at the University. During his junior year Arthur was first introduced to Golomb's Tromino Theorem. He is currently pursuing a B.S. in Computer Science and Mathematics at Yale University, Class of 2019. In his free time Arthur fences, composes electronic music, and avoids complicated integration like the plague.

JONATHAN LENCHNER (MR Author ID: 782105, ORCID 0000-0002-9427-8470) earned a Ph.D. in mathematics late in life from Polytechnic University (now the NYU Tanden School of Engineering) in 2008. He has spent almost 20 years at IBM and became Chief Scientist of IBM's African research labs, one in Nairobi, the other in Johannesburg, in May of 2016. He learned about Golomb's Theorem from Arthur on a family visit, while Arthur was still in high school.

Proof Without Words: Products of Odd Squares and Triangular Numbers

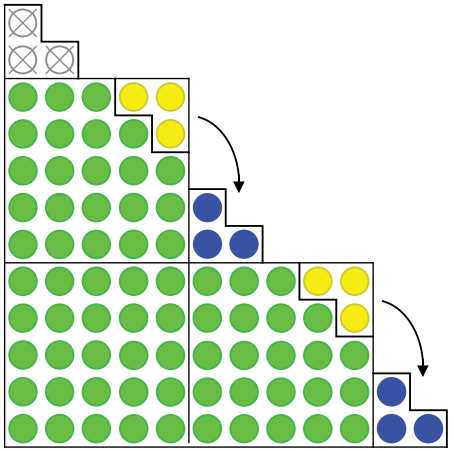
BRIAN HOPKINS
Saint Peter's University
Jersey City, NJ 07306
bhopkins@saintpeters.edu

Write $T_n = 1 + \cdots + n$ for the nth triangular number.

Theorem. $(2k+1)^2 \cdot T_n = T_{(2k+1)n+k} - T_k$ for $n, k \in \mathbb{N}$.

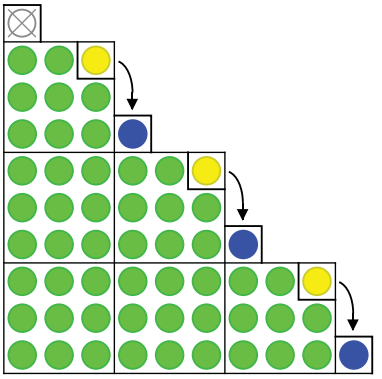
Proof.

E.g., for $k = 2, n = 2$:



$5^2 \cdot T_2 = T_{12} - T_2$.

E.g., for $k = 1, n = 3$:



$3^2 \cdot T_3 = T_{10} - T_1$. ■

See [1] for another visual approach to this result.

Summary. We visually demonstrate an identity equating the product of an odd number squared and a triangular number to a difference of triangular numbers.

REFERENCE

[1] Nelsen, R. B. (1994). Proofs without words: A triangular identity. *Math. Mag.* 67:293. doi.org/10.2307/2690851. Also in Nelsen, R. B. (1993). *Proofs Without Words*, Vol. 1. Washington, DC: Mathematical Association of America, p. 105.

BRIAN HOPKINS (MR Author ID: 734157) is a professor of mathematics at Saint Peter's University who is finishing up a term as editor of *The College Mathematics Journal*. Among his research interests are Bulgarian solitaire and the sand pile model, examples of dynamics on integer partitions, which motivated this proof.

Math. Mag. **91** (2018) 70. doi:10.1080/0025570X.2017.1409534
MSC: Primary 05A19
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

PROBLEMS

EDUARDO DUEÑEZ, *Editor*
Spelman College

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

JOSÉ *A. GÓMEZ*, Facultad de Ciencias, UNAM, Mexico; CODY PATTERSON, University of Texas at San Antonio; MARÍA LUISA PÉREZ-SEGUÍ, Universidad Michoacana SNH, Mexico; RICARDO A. SÁENZ, Universidad de Colima, Mexico; ROGELIO VALDEZ, Centro de Investigación en Ciencias, UAEM, Mexico; *Assistant Editors*

Proposals

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

2036. *Proposed by Dan Stefan Marinescu, Hunedoara City and Leonard Giugiuc, Drobeta Turnu-Severin, Romania.*

Let a and b be real numbers with $a < b$. Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(tx + (1-t)y) \leq \max\{f(x), f(y)\}$ for all $x, y \in [a, b]$ and $t \in [0, 1]$. Prove that if $f(a) = 0$ and $\int_a^b f(x)\,dx = 0$ then $\int_a^b f(x)g(x)\,dx \geq 0$ for all increasing functions $g : [a, b] \to \mathbb{R}$.

2037. *Proposed by Ioana Mihăilă, Cal Poly Pomona, Pomona, CA.*

A point D lies on the hypothenuse $\overline{BC}$ of a right triangle $\triangle ABC$ so that $AB = BD$. Let P be the point on the circumcircle of $\triangle ADC$ such that $\angle APB$ is a right angle, and let L be the midpoint of $\overline{AD}$. Show that $\overline{PC}$ is perpendicular to $\overline{PL}$.

2038. *Proposed by Eugène Delacroix, Lycée Thérèse d'Avila, France and Su Pernu Mero, Valenciana GTO, Mexico.*

Given any real-valued random variable X, let $A_{11}, A_{12}, A_{21}, A_{22}$ be independent random variables that have the same distribution as X, and let

$$\widetilde{X} = \min_i \max_j A_{ij} = \min\{\max\{A_{11}, A_{12}\}, \max\{A_{21}, A_{22}\}\}.$$

(Although $\widetilde{X}$ does not directly depend on X but rather on the variables A_{ij}, its probability distribution is uniquely determined by that of X.) Define a sequence $\{X_0, X_1, \ldots, X_n, \ldots\}$ recursively by $X_0 = X$ and $X_{n+1} = \widetilde{X}_n$. Prove that, as $n \to \infty$,

Math. Mag. **91** (2018) 71–77. doi:10.1080/0025570X.2018.1411650

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.

X_n tends in distribution to a discrete random variable Z taking at most two values. Characterize the distribution of Z in terms of the distribution of X.

2039. *Proposed by Baris Burcin DEMIR, Ali Naili Erdem Anatolian High School, Ankara, Turkey.*

Given a triangle $\triangle ABC$, let $\mathcal{M}$ be the locus of all midpoints P of segments $\overline{DE}$ that divide $\triangle ABC$ into equivalent (i.e., equal-area) parts, where both D and E lie on some side (or possibly a vertex) of $\triangle ABC$. Compute the ratio of the area of the region enclosed by $\mathcal{M}$ to the area of $\triangle ABC$.

2040. *Proposed by George Stoica, Saint John, New Brunswick, Canada.*

Fix a positive integer n. Let A be an $n \times n$ complex matrix such that $A^n = 0$. For any complex number $z \neq 0$ and positive integer m, prove that there exists a matrix B such that $B^m = 0$ and $A + z^m I = (B + zI)^m$, where I denotes the $n \times n$ identity matrix.

Quickies

1077. *Proposed by Julien Sorel, Piatra Neamt, PNI, Romania.*

Find all integers k such that $0 \leq k < 100$ and the binomial coefficient $\binom{99}{k}$ is not divisible by 3.

1078. *Proposed by Greg Oman, University of Colorado, Colorado Springs, CO.*

Let Ω be an uncountable well-ordered set such that, for all $\alpha \in \Omega$, the set

$$\Omega_{<\alpha} := \{\beta \in \Omega : \beta < \alpha\}$$

of predecessors of α is countable. (Ω is order-isomorphic to the first uncountable ordinal ω_1.) We will call any collection $(x_\alpha : \alpha \in \Omega)$ of real numbers indexed by Ω an Ω-*sequence*. We say that an Ω-sequence (x_α) *converges to the real number* r if given a positive number ϵ there exists $\alpha \in \Omega$ such that $|x_\beta - r| \leq \epsilon$ for all $\beta \in \Omega$ such that $\beta \geq \alpha$, and that it *is eventually equal to the constant* r if the preceding property holds for $\epsilon = 0$. Is there a convergent Ω-sequence that is not eventually constant?

Solutions

Sums of positive and negative numbers in an open set — February 2017

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

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

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

The proof will only assume that S contains at least one negative number as well as some open neigborhood of a positive number. Let T be the set of all finite sums of elements of S. By hypothesis there are positive real numbers x, y, δ such that $-y \in S$ and S includes the open interval $I = (x - \delta, x + \delta)$. Since $\mathbb{Q}^+$ is dense in $\mathbb{R}^+$, there are positive integers m, n such that $|x/y - n/m| < \delta/y$, hence $|mx - ny| < m\delta$. It follows

that the open interval $I_{m,n} := (mx - ny - m\delta, mx - ny + m\delta)$ contains 0. It is clear that $I_{m,n} = mI - ny$ is precisely the set of sums of m elements taken from I and n times the element $-y$, and moreover T includes all such sums. Given any $z \in \mathbb{R}$, there is some positive integer k such that $z/k \in I_{m,n}$; hence, $z = k(z/k) = z/k + z/k + \cdots + z/k \in T$, since T is obviously closed under addition.

Also solved by Ulrich Abel (Germany), Elton Bojaxhiu (Albania) & Enkel Hysnelaj (Australia), Paul Budney, John Christopher, Timothey V. Craine, Charles Degenkolb, Robert L. Doucette, Gregory Dresden, Joseph DiMuro, Dmitry Fleischman, Isaac Garfinkle (student) Rafe Jones, Abhay Goel, Russell A. Gordon, Tom Jager, Kelly Jahns, Reiner Martin (Germany), Missouri State University Problem Solving Group, Northwestern University Math Problem Solving Group, Eugene A. Herman, Edward Schmeichel, Skidmore College Problem Group, Philip Straffin, John Tolle, Edward T. White, Stuart V. Witt, and the proposer.

A family of integrals with value $\pi/8$ — February 2017

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

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

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

Solution by Ulrich Abel, Technische Hochschule Mittelhessen, Friedberg (Germany). Under the stated hypotheses on f, we show that

$$I(a) := \int_{\sqrt{2}-1}^{\sqrt{2}+1} \frac{dx}{(1+x^2)(1+a^{f(x)})} = \frac{\pi}{8}.$$

Using $(\sqrt{2}-1)(\sqrt{2}+1) = 1$, the change of variable $x = 1/t$ shows that $I(a) = I(a^{-1})$. From the identity

$$\frac{1}{1+a^u} + \frac{1}{1+a^{-u}} = 1 \qquad \text{for all } u \in \mathbb{R},$$

we obtain

$$I(a) = \frac{I(a) + I(a^{-1})}{2} = \frac{1}{2}\int_{\sqrt{2}-1}^{\sqrt{2}+1} \frac{dx}{1+x^2} = \frac{\arctan(\sqrt{2}+1) - \arctan(\sqrt{2}-1)}{2} = \frac{\pi}{8},$$

where the last equality follows from $\arctan 1 = \pi/4$ and the identity

$$\arctan x - \arctan y = \arctan\left(\frac{x-y}{1+xy}\right) \qquad \text{for } xy > -1.$$

Also solved by Michel Bataille (France), Gerald E. Bilodeau, Brian Bradie, Bruce S. Burdick, Prithwijit De (India), Robert L. Doucette, Dmitry G. Fleischman, Isaac Garfinkle, Michael Goldenberg & Mark Kaplan, Russell A. Gordon, Raymond N. Greenwell, GWstat Problem Solving Group, Lixing Han, Eugene A. Herman, Tom Jager, Isaac E. Leonard (Canada), Weiping Li, James Magliano, Soumitra Mandal (Chandar Nagore, India), Soumitra Mandal (Kolkata, India), Rituraj Nandan, Northwestern University Math Problem Solving Group, Moubinool Omarjee (France), Paolo Perfetti (Italy), Ángel Plaza (Spain) Ravi Prakash (India), Shafiqur Rahman (Bangladesh), Don Redmond, Michael Reid, Edward Schmeichel, Seán M. Stewart (Australia), John Tolle, Robert W. Vallin, Michael Vowe (Switzerland), Edward T. White, John Zacharias, and the proposer. There was one incomplete or incorrect solution.

Counting lattice points in a tetrahedron **February 2017**

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

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

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

Solution by Marty Getz, University of Alaska Fairbanks, and Dixon Jones, Fairbanks, AK.

The cube $\mathcal{C} = [0, n] \times [0, n] \times [0, n]$ contains $(n + 1)^3$ lattice points. Let $O'(n, n, n)$, $A'(n, 0, 0)$, $B'(0, n, 0)$, $C'(0, 0, n)$ be the vertices of $\mathcal{C}$ diagonally opposite to O, A, B, C, respectively. The complement of $\mathcal{T}$ in $\mathcal{C}$ is the union of the four congruent solid trirectangular tetrahedra $O'ABC$, $OA'BC$, $OAB'C$, $OABC'$ each with one face (ABC, resp. OBC, resp. OAC, resp. OAB) removed. Clearly, each of these four tetrahedra contains $n(n + 1)(n + 2)/6$ lattice points (the nth tetrahedral number). It follows that $\mathcal{T}$ contains

$$N = (n + 1)^3 - 4\left(\frac{n(n + 1)(n + 2)}{6}\right) = \frac{1}{3}(n + 1)(n^2 + 2n + 3)$$

lattice points.

Also solved by Pedro Acosta, Michel Bataille, Elton Bojaxhiu (Albania) & Enkel Hysnelaj (Australia), Robin Chapman (UK), John Christopher, Con Amore Problem Group (Denmark), Timothy Crane, Robert L. Doucette, Habib Y. Far, GW University Math Problems Group, GWstat Problem Solving Group, Eugene Herman, Lucyna Kabza, Hidefumi Katsuura, Alejandro Mahillo (Spain), Peter McPolin (UK), Rituraj Nandan, Northwestern University Math Problem Solving Group, Zachery Peterson, Rob Pratt, Edward Schmeichel, Skidmore College Problem Group, Philip Straffin, and the proposer.

Anti-automorphisms of $\mathbb{Z}_n$ **February 2017**

2014. *Proposed by Gaitanas Konstantinos, Greece.*

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

Solution by Isaac Garfinkel (student) and Rafe Jones, Carleton College, Northfield, MN.

An anti-automorphism of $\mathbb{Z}_n$ exists precisely for n odd. Clearly, f is an anti-morphism precisely when $g(x) := f(x) - x$ is an injective function from $\mathbb{Z}_n$ to itself. If n is odd, it is clear that $f(x) = -x$ is an anti-automorphism of $\mathbb{Z}_n$ since $g(x) = f(x) - x = -2x$ is an injective function on $\mathbb{Z}_n$ (because 2 is invertible in $\mathbb{Z}_n$ when n is odd) and f is obviously bijective. Conversely, consider any anti-automorphism f on $\mathbb{Z}_n$ for some $n \geq 2$. Let $g(x) := f(x) - x$ as above, so g is injective. Since f is a bijection from $\mathbb{Z}_n$ to itself, it is a permutation of $\mathbb{Z}_n$. In any k-cycle $(x_1 x_2 \ldots x_k)$ of the permutation f (where $x_2 = f(x_1)$, ..., $x_k = f(x_{k-1})$, $x_1 = f(x_k)$), we have

$$\begin{aligned} &g(x_1) + g(x_2) + \ldots + g(x_{k-1}) + g(x_k) \\ &\quad = [f(x_1) - x_1] + [f(x_2) - x_2] + \ldots + [f(x_{k-1}) - x_{k-1}] + [f(x_k) - x_k] \\ &\quad = (x_2 - x_1) + (x_3 - x_2) + \ldots + (x_k - x_{k-1}) + (x_1 - x_k) = 0. \end{aligned}$$

(Note that this reasoning is still valid for a 1-cycle (x_1) where $f(x_1) = x_1$ since in this case $g(x_1) = f(x_1) - x_1 = 0$.) Since $\mathbb{Z}_n$ is a disjoint union of such cycles, it is clear that $\sum_{x\in\mathbb{Z}_n} g(x) = 0$. On the other hand, since g is an injective function from $\mathbb{Z}_n$ to itself, it is also a permutation of $\mathbb{Z}_n$, so we have

$$\sum_{x\in\mathbb{Z}_n} g(x) = \sum_{x\in\mathbb{Z}_n} x = \frac{n(n+1)}{2}.$$

Hence, $n(n+1)/2$ must be the zero element of $\mathbb{Z}_n$, so $(n+1)/2$ must be an integer; thus, n must be odd.

Also solved by Paul Budney, Robin Chapman (UK), Joseph DiMuro, Robert L. Doucette, Dmitry Fleischmann, George Washington University Jump Problems Team, Eugene A. Herman, Tom Jager, Daniel López-Aguayo (Mexico), Peter McPolin (UK), Michael Reid, Edward White, and the proposer.

Separable polynomials in a "reverse arithmetic" sequence **February 2017**

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

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

Solution by Michael Reid, University of Central Florida, Orlando, FL.
The statement *"$R(X)$ has simple zeros"* will be interpreted in the strong sense that $R(X)$ is *separable*, i.e., it has no repeated roots in an algebraic closure $\overline{K}$ of K. There is no loss of generality in supposing that P is monic, so we will do so. Let $d = \deg P$ and write $P(X) = X^d + f(X)$, with $f(X)$ a polynomial with coefficients in K and degree at most $d-1$.

Assume first that K is infinite. Given $m \geq 1$, let $R(X) = X^m P(X) - t$. The discriminant of $R(X)$ is a polynomial $D(t)$ in the variable t with (possibly zero) leading term $\pm(d+m)^{d+m} t^{d+m-1}$. Choose $m \geq 1$ such that this leading term is nonzero, so $D(t)$ is a nonzero polynomial (if $d+1$ is not a multiple of the characteristic of K, let $m = 1$; otherwise, let $m = 2$). Since D has finitely many roots while K is infinite, there is $c \in K$ such that $D(c) \neq 0$, hence $\operatorname{disc}(X^m P(X) - c) \neq 0$ and the polynomial $X^m P(X) - c$ is separable, so it suffices to take $Q(X) = -c$ (constant) in this case.

Next, assume that K is finite of prime characteristic p and order equal to some power q of p. Choose $n = 1$ or 2 so that $d + n$ is coprime to p. Then, as above, $D(t) = \operatorname{disc}(X^n P(X) - t)$ is a nonzero polynomial of degree $\deg D = d + n - 1$. Choose $r \geq 1$ such that $\ell := 1 + q^r > d + n$. By elementary properties of finite fields, K has a degree-ℓ extension field L; moreover, L is Galois over K, and also primitive over K, i.e., $L = K(\alpha)$ for some $\alpha \in L$. In particular, the monic minimal polynomial $g(X)$ for α over K is separable and splits in L: $g(X) = \prod_{i=1}^{\ell}(X - \alpha_i)$ with $\alpha_1, \ldots, \alpha_\ell$ distinct elements of L (the Galois conjugates of α, one of which is α itself). The polynomial $g(X)$ is of the form $g(X) = X^\ell + h(X)$ where $\deg h \leq \ell - 1$ and $h(X)$ has coefficients in K. For $i = 1, \ldots, \ell$, the degree of α_i over K is equal to $\ell > d + n > \deg D$, so it follows that α_i is not a zero of the polynomial $D(t)$; therefore, $R_i(X) := X^n P(X) - \alpha_i$ is a polynomial with coefficients in L and separable. Moreover, $R_1(X), \ldots, R_\ell(X)$ are pairwise

relatively prime because their differences $R_j(X) - R_i(X) = \alpha_i - \alpha_j$ are nonzero constants for $i \neq j$. The polynomial $R(X) := g(X^n P(X))$ clearly has coefficients in K and factors as $R(X) = \prod_{i=1}^{\ell}(X^n P(X) - \alpha_i) = \prod_{i=1}^{\ell} R_i(X)$; hence, $R(X)$ is also separable being a product of pairwise coprime separable factors. We will show that $R(X)$ has the form required by the problem.

Since q is a power of $p = \operatorname{char} K = \operatorname{char} L$, the identity $(A+B)^{q^r} = A^{q^r} + B^{q^r}$ holds for arbitrary polynomials A, B with coefficients in L, and we obtain

$$\begin{aligned} R(X) &= g(X^n P(X)) = (X^n P(X))^{\ell} + h(X^n P(X)) = X^{\ell n} P(X)^{1+q^r} + h(X^n P(X)) \\ &= X^{\ell n} P(X) P(X)^{q^r} + h(X^n P(X)) = X^{\ell n} P(X)(X^d + f(X))^{q^r} + h(X^n P(X)) \\ &= X^{\ell n} P(X)(X^{dq^r} + f(X)^{q^r}) + h(X^n P(X)) \\ &= X^{\ell n + dq^r} P(X) + [X^{\ell n} f(X)^{q^r} P(X) + h(X^n P(X))] \\ &= X^m P(X) + Q(X), \end{aligned}$$

where we let $m = \ell n + dq^r$ and $Q(X) = X^{\ell n} f(X)^{q^r} P(X) + h(X^n P(X))$. On the one hand, $X^{\ell n} f(X)^{q^r} P(X)$ has degree at most $\ell n + (d-1)q^r + d = m - (q^r - d) < m$; on the other hand, $h(X^n P(X))$ has degree at most $(\ell - 1)(d + n) = m - n < m$; therefore, $\deg Q < m$. Since $P(X)$, $f(X)$, and $h(X)$ have coefficients in K so does $Q(X)$, completing the solution.

Editor's Note. The solution above shows that the condition that $\deg P$ be coprime to $\operatorname{char} K$ is not necessary. Professor Michael Reid remarks that, at least when K is finite, one may add the requirement that $R(X)$ be irreducible over K as follows. An analog for irreducible polynomials over finite fields of Dirichlet's Theorem on primes in arithmetic progressions was proved by H. Kornblum, Über die Primfunktionen in einer arithmetischen Progression, *Math. Zeitschrift* **5** (1919) 100–111, http://www.digizeitschriften.de/dms/resolveppn/?PID=GDZPPN002364972. The reverse polynomial $\widetilde{P}(X) = X^d P(1/X)$ of P has degree at most d and nonzero constant term; therefore, $\widetilde{P}(X)$ is coprime to X^{d+1}, so there is an irreducible polynomial of the form $\widetilde{R}(X) = X^{d+1} S(X) + \widetilde{P}(X)$ with coefficients in K whose reversal $R(X)$ is also irreducible and has the required form; moreover, $R(X)$ is necessarily separable since every finite field K is perfect.

Also solved by Isaac Garfinkle and the proposer. There was 1 incomplete or incorrect solution.

Answers

Solutions to the Quickies from page 72.

A1077. We use congruences modulo 3 in the ring $\mathbb{Z}[X]$ of formal univariate polynomials with integer coefficients. Explicitly, $p(X) \equiv 0 \pmod 3$ means that all coefficients of p are multiples of 3; equivalently, that $p(X) = 3q(X)$ for some $q \in \mathbb{Z}[X]$. By the binomial formula, we have

$$\sum_{k=0}^{99} \binom{99}{k} X^k = (1+X)^{99}.$$

We evaluate $(1+X)^{99}$ modulo 3 as follows. First, note that $(A+B)^3 - (A^3 + B^3) = 3AB(A+B) \equiv 0 \pmod 3$ holds for all $A, B \in \mathbb{Z}[X]$, hence $(A+B)^3 \equiv A^3 + B^3 \pmod 3$ (the "freshman dream"). By induction, we have $(A+B)^{3^l} = A^{3^l} + B^{3^l}$ for any

integer $l \geq 0$, hence

$$(1+X)^{99} = (1+X)^{81}(1+X)^{9\cdot 2} \equiv (1+X^{81})(1+X^9)^2 \equiv (1+X^{81})(1+2X^9+X^{18})$$
$$\equiv 1+2X^9+X^{18}+X^{81}+2X^{90}+X^{99} \pmod 3.$$

Thus, $\binom{99}{k}$ is not divisible by 3 exactly when $k = 0, 9, 18, 81, 90, 99$.

A1078. The answer is no. We prove that every convergent Ω-sequence is eventually constant. Assume that (x_α) converges to some real number r. For every positive integer n, choose α_n such that $|x_\beta - r| \leq 1/n$ for $\beta \geq \alpha_n$. The countable union $\Gamma := \bigcup_n \Omega_{<\alpha_n}$ is countable since each $\Omega_{<\alpha_n}$ is countable by assumption. Since Ω is uncountable, it has an element γ not in Γ. For all n we obviously have $\gamma \geq \alpha_n$ (since $\gamma \notin \Gamma \supseteq \Omega_{<\alpha_n}$), hence any $\beta \in \Omega$ such that $\beta \geq \gamma$ must satisfy $\beta \geq \alpha_n$, so $|x_\beta - r| \leq 1/n$ holds. We conclude that $|x_\beta - r| = 0$ for $\beta \geq \gamma$, so (x_α) is eventually equal to the constant r.

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.

Inglis, Matthew, and Nina Attridge, *Does Mathematical Study Develop Logical Thinking? Testing the Theory of Formal Discipline*, World Scientific, 2017; xvii + 185 pp, $102. ISBN 978-1-78634-068-9.

Johnson, Peter, Does algebraic reasoning enhance reasoning in general? A response to Dudley. *Notices of the American Mathematical Society* 59 (9) (October 2012): 1270–1271, http://www.ams.org/notices/201209/rtx120901270p.pdf.

Inglis and Attridge ask a question that is absolutely crucial to the place of mathematics in education (apart from its usefulness and its cultural importance): Does mathematical study develop logical thinking? For millennia, the answer—from Plato to Locke and onward, promoted by educators, presumed by parents, and more or less accepted by students—has been *yes*. Of course, the same logic, based on the Theory of Formal Discipline (TFD), was applied also to learning Latin and (by Benjamin Franklin) to learning chess. Research by psychologists—of which mathematicians are almost completely ignorant—uniformly rejects TFD. The leading alternative explanation is the "filtering hypothesis": Better reasoners do better in mathematics (and then continue on to take more mathematics). However, the authors argue (weakly) that studying "advanced" mathematics is "associated with development of reasoning skills," particularly "the ability to reject invalid inferences." Psychologists are highly skeptical of transfer of learning from one domain to another, as the article by Johnson notes particularly in connection with algebra: "There appears to be no research whatsoever that would indicate that the kind of reasoning skills a student is expected to gain from learning algebra would transfer to other domains of thinking or to problem solving or critical thinking in general." That is absolutely damning for TFD as far as algebra goes, despite Johnson's hedge: "The lack of such research evidence does not mean that such transfer does not occur or that algebraic reasoning might not have positive effects on problem solving and critical thinking." Curiously, there is null intersection of the references in the book (whose authors are from the United Kingdom) with those in the article by Johnson (in Connecticut). A key question: If studying mathematics (or Latin, for that matter) does not help develop reasoning, what does?

Diaconis, Persi, and Brian Skyrms, *Ten Great Ideas about Chance*, Princeton University Press, 2017; xi + 246 pp, $27.95. ISBN 978-0-691-17416-7.

"This is a history book, a probability book, and a philosophy book." It is also a terrific book. The authors explain 10 great ideas in probability, starting from their history and pursuing their philosophical implications. They pithily summarize each idea near the start of each chapter: chance can be measured; judgments can be measured in terms of probabilities; the psychology of chance and the logic of probability are different subjects; the law of large numbers; probability through measure theory; Bayes' theorem; exchangeability; randomness; physical chance; and induction. Appendices to some chapters give more detail and depth. The authors assume that the reader has taken an undergraduate course in probability or statistics; an appendix contains a tutorial on basic ideas in probability.

Math. Mag. **91** (2018) 78–79. doi:10.1080/0025570X.2017.1411696

O'Shea, Owen, *The Call of the Primes: Surprising Patterns, Peculiar Puzzles, and Other Marvels of Mathematics*, Prometheus Books, 2016; 330 pp, $19. ISBN 978-1-63388-148-8.
Today, thanks in part to growth in the mathematics profession, there is a vast cornucopia of popular works on mathematics, appealing to varying levels of mathematical experience. Their proliferation and variety has increased interest, curiosity, and support for mathematics. Many such books come to my desk for potential review. Most are inspiring for some readers and are worthy efforts; but there are too many for me to distinguish carefully their strengths and draw your attention to the merits of each, and they compete for attention here with other works. Particularly worthy of attention, however, is O'Shea's *The Call of the Primes*. As a nonmathematician, he approaches mathematics from the point of view of recreation: "[E]njoyment breeds the desire to explore and to seek out new challenges...." He concentrates on asking questions and leading the reader to conjecture patterns; I learned something new in every chapter.

Stewart, Ian, *The Beauty of Number in Nature: Mathematical Patterns and Principles from the Natural World*, MIT Press, 2017; 224 pp, $24.95(P). ISBN 978-0-262-53428-4.
This book, an update and revision of *What Shape Is a Snowflake?* (2001), contains astonishingly beautiful photos of mathematics in nature. It begins with the puzzle of the shape of a snowflake, examines what a pattern is, catalogs various kinds of patterns in different dimensions, considers fractal geometry and chaos, and finally returns to snowflakes. "I do not believe that the beauty of a snowflake can be spoiled by an awareness of what makes it....What shape is a snowflake? Snowflake-shaped."

Hayes, Brian, *Foolproof and Other Mathematical Meditations*, MIT Press, 2017; x + 234 pp, $24.95. ISBN 978-0-262-03686-3.
These 13 excellent essays appeared in their original form in the "Computing Science" column of *American Scientist*, between 1998 and 2014; the versions here are extensively revised and updated (and some are retitled). Author Hayes wrote many other exciting columns, so this collection is just a sample. The first essay examines the history of the legend of Gauss summing the first 100 integers. The last, "Foolproof," begins with "I was a [professional] teenage angle trisector" and goes on to explore how proofs compel belief. In between are the explorations of a nonmathematician who writes extremely well about his struggles and journeys to satisfy his curiosity about mathematical puzzles, concepts, and calculations.

Stillwell, John, *Reverse Mathematics: Proofs from the Inside Out*, Princeton University Press, 2018; vii + 182 pp, $29.95. ISBN 978-0-691-17717-5.
Proofs and Refutations (1977) by Imre Lakatos promoted the idea that mathematical theorems begin from desired conclusions and reason back to sufficient conditions. John Stillwell takes that premise further, to ask what axioms are needed to prove a theorem. Proceeding beyond the parallel postulate and the axiom of choice, Stillwell identifies three levels of axiom systems that (between them) prove most of the basic theorems of analysis. In each case, the axioms can be proved from the theorem. The result is a hierarchy of "deepness": intermediate value theorem $\prec$ Heine–Borel theorem and extreme value theorem $\prec$ Cauchy convergence criterion and Bolzano–Weierstrass theorem. Readers will encounter gentle introductions to mathematical logic, computability, definability, and Σ^0_1 sets.

Newton, Isaac, *The Mathematical Principles of Natural Philosophy*: reissue of 1st American edition published in 1846, translated by Andrew Motte; KroneckerWallis, 2017; 688 pp, € 45.
"Books are something to touch and look at. Not just read." That is the major principle of the publisher of this collector's edition of Newton's *Principia*, with its emphasis on design. Each of the three main chapters is bound separately as a fascicule. The binding, page size (15 cm by 21 cm), typeface (The Serif), type size, amount of matter on a page, and ink colors ("petrol blue and coral orange") were all carefully chosen. The result is indeed a beautiful commemoration of the 330th anniversary of publication of Newton's original work. The contrast between "petrol blue" and "coral orange" is welcome, except that the "coral orange" is so light in thin type as to make paragraphs and entire pages in it hard to read.