LETTER FROM THE EDITOR

After the summer heat, THE MAGAZINE has returned for the cooler weather and fall classes. The October issue begins with a big picture of numerical analysis that describes the four fundamental spaces of numerical analysis in a paper by Omid Khanmohamadi and Ehssan Khanmohammadi. This article should provide a good resource for students (and faculty) to learn about the ideas behind numerical analysis. Back to classes does not mean that summer fun is over. In the second article, Jay Cummings provides a mathematical look at a card trick (called a spelling shuffle) and answers a question about the trick asked previously by well-known mathemagician Colm Mulcahy.

The next two articles focus on real analysis. Stefan Steinerberger provides an amusing sequence of real-valued functions. The functions are amusing in part because every rational point is eventually the location of a strict local minimum of a function in the sequence. Marc Frantz considers a two-sided sequence of functions with interesting limiting behavior. The article touches on topics that include pointwise convergence, uniform convergence, Dini's theorem, the uniform limit theorem, and the Hausdorff metric.

A result by Cauchy bounding the moduli of (scalar) polynomial zeros can be derived by using a localization result for matrix eigenvalues, namely Gershgorin's theorem. In the next article, Aaron Melman shows how Gershgorin's theorem can be used in the extension of Cauchy's result from scalar to matrix polynomials.

The next three articles focus on geometry. T.R. Mukundan starts off the set of articles by providing a simplified approach to the classical Fermat–Torricelli problem (of finding the point in a triangle that minimizes the sum of the distances to the vertices) and its weighted version using only basic calculus and geometry. Stephen Su and C.S. Lee provide a new generalization of theorems by Menelaus, Ceva, and Routh about coincidences of points and lines in triangles. I.P.D. De Silva offers a look at Hofstadter points, one of many possible centers of a triangle.

As with every issue in 2018, there is a Partiti puzzle (and its solution). Look for an interval programming approach to solving Partiti puzzles in the December issue. The art interview returns in this issue, as Allison Henrich interviews Veronika Irvine about patterns of lace. There are new Problems and Solutions and Reviews, as well as two proofs without words (by Charles Marion and by Yajun An and Tom Edgar). The issue concludes with the announcement of the Allendoerfer award: congratulations to Fumiko Futamura and Robert Lehr who received the Allendoerfer award at MathFest for their article A New Perspective on Finding the Viewpoint from October 2017.

Michael A. Jones, Editor

Four Fundamental Spaces of Numerical Analysis

OMID KHANMOHAMADI
Seattle, WA 98164
omidmath@protonmail.com

EHSSAN KHANMOHAMMADI
Lancaster, PA 17603
ehssan@protonmail.com

For a given positive integer n, let us look at two ways to solve the integration problem

$$I_n = \int_0^1 x^{n-1} e^{x-1}\, dx, \tag{1}$$

(An "explicit solution formula" for I_n may be found in terms of the Gamma function and incomplete Gamma function, which may only be evaluated approximately.) For $n = 1$, we have $I_1 = \int_0^1 e^{x-1}\, dx = 1 - 1/e$. We may use integration by parts, $I_n = x^{n-1}e^{x-1}\big|_0^1 - \int_0^1 (n-1)x^{n-2}e^{x-1}\, dx$, to get the recursion relation

$$I_n = \begin{cases} 1 - (n-1)I_{n-1} & \text{if } n > 1, \\ 1 - 1/e & \text{if } n = 1. \end{cases}$$

Let us try to find I_{20} by solving this recursion relation. This problem has the exact solution

$$\begin{aligned} I_{20} &= 121645100408832000a + 76894368849186894 \\ &= \frac{121645100408832000}{e} - 44750731559645106 \\ &\approx 0.047722755796209, \end{aligned}$$

as may be seen by invoking Algorithm 1 with $n = 20$ and $a = 1 - 1/e$.

Algorithm 1. Solve integration problem (1) using recursion with different initial values a.

```
1 function Integral (n, a)
2   if n > 1
3     return 1 − (n − 1)Integral (n − 1, a)
4   else
5     return a
```

Math. Mag. **91** (2018) 243–253. doi:10.1080/0025570X.2018.1494458
MSC: Primary 65J05, Secondary 65D30
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

Now, suppose we instead solve the approximate recursion problem

$$\widetilde{I}_n = \begin{cases} 1 - (n-1)\widetilde{I}_{n-1} & \text{if } n > 1, \\ 0.6321205588285577 & \text{if } n = 1, \end{cases} \tag{2}$$

where $\widetilde{I}_1 := 0.6321205588285577$ is an approximation to $I_1 := 1 - 1/e$ to 16 decimal digits. Let us try to find $\widetilde{I}_{20}$. This approximate problem has the solution

$$\widetilde{I}_{20} = 121645100408832000a + 76894368849186894 = -2.5792669016064000,$$

as may be seen by invoking Algorithm 1 with $n = 20$ and $a = \widetilde{I}_1$. Not even the signs agree for $\widetilde{I}_{20}$ and I_{20}!

Exercise. Suppose Algorithm 1 is implemented in floating-point arithmetic. What would the result of `Integral`$(20, 0.6321205588285577)$ be?

Did you expect the answer to be equal to $\widetilde{I}_{20}$? What is cooking? Try answering this question yourself, before looking at the answer at the end of this article.

Why did a small error (of order 10^{-16}) in approximating $\widetilde{I}_1$ from I_1 get amplified under the recursive relation in equation (2) to a huge error in $\widetilde{I}_{20}$? It turns out this method is an unstable approximation method for solving the integral. In the next section, we precisely define the notions of stability and convergence (together with consistency) and investigate their interplay.

Four fundamental spaces of numerical analysis

The equivalence of stability and convergence of a consistent approximation method for solving an exact problem has been called the fundamental theorem of numerical analysis (see, e.g., [**5**]). This is a general statement and we present it in a unified framework, illustrated in Figure 1. Each box is one of the fundamental spaces of numerical analysis. The top two and the mapping between them define an exact problem

$$(A, U, F) \tag{3}$$

using the linear, exact problem operator A going from the exact solution space U to the exact data space F, both of which we take to be normed vector spaces. The domain and range of A are $D \subset U$ and $R \subset F$. The solid arrow in the top row shows an instance of the exact problem corresponding to some particular data f

$$Au = f, \quad u \in D, \quad f \in R. \tag{4}$$

If for any data $f \in R$ this exact problem has a unique solution u that depends continuously on f, we say the problem is well-posed. This implies that A has a continuous left inverse called a solution operator $S\colon R \to U$, with $SAu = u$, for all u in D. To keep the notation easier to follow, we write A^{-1} for S, keeping in mind that we can only apply A^{-1} on the left of A to get the identity. So, the unique solution to the exact problem is $u = A^{-1}f$. But usually the exact problem is too difficult to solve using A^{-1} directly. In particular, it may be too difficult to construct A^{-1} or its action. So, instead, we try to solve an approximate problem.

This brings us to the lower part of Figure 1. We pick an approximation parameter h in an index set H, a measure of how "refined" our approximation is going to be. For example, h could measure the spatial spacing of the grid in a finite difference approximation method, element diameter in a finite-element approximation method, etc. In these cases, the index set H is typically $\mathbb{R}^+$.

Then we construct an approximate-data bridge, namely a bounded linear operator r_h ("r" for restriction) to get a particular approximation $r_h f$ to any given exact data f.

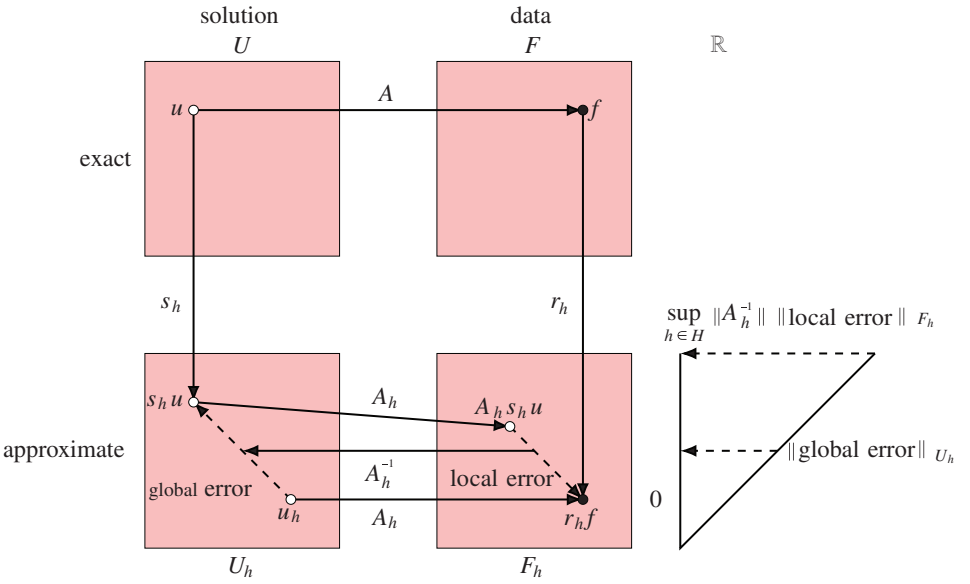


Figure 1 Fundamental theorem of numerical analysis ("stability is equivalent to convergence") and its four associated spaces. Hollow and filled circles denote unknown and known quantities. Solid arrows denote mappings. Dashed arrows denote limits. The top two boxes depict our exact solution and data spaces, and the lower two the respective approximate ones. The operators s_h and r_h are approximation bridges, connecting the exact spaces to approximate ones. The "triangle" on the right proves that for a consistent approximation (i.e., when $\|\text{local error}\|_{F_h} \to 0$), stability (i.e., when $\sup_{h\in H} \|A_h^{-1}\| < \infty$) implies convergence (i.e., $\|\text{global error}\|_{U_h} \to 0$). In other words, $\|\text{global error}\|_{U_h}$ is squeezed between 0 and $\sup_{h\in H} \|A_h^{-1}\| \, \|\text{local error}\|_{F_h}$.

We want our approximate-data bridges $\{r_h\}_{h\in H}$ to be "stable" as a whole, which means that if any "head" f of any bridge r_h is shaken slightly (this may happen due to randomness, measurement error, etc., in f), the "tail" $r_h f$ is shaken only slightly too. This stability is ensured by the uniform boundedness of the family $\{r_h\}_{h\in H}$ and in symbols is written as $\sup_{h\in H} \|r_h\| < \infty$.

We also need to get from the exact problem operator A to an approximate problem operator A_h, a bounded linear operator going from the approximate solution space U_h to the approximate data space F_h. It has domain $D_h \subset U_h$ and range $R_h \subset F_h$. How to get A_h from A is at the heart of the approximation we choose. The lower-most solid arrow in Figure 1 shows an instance

$$A_h u_h = r_h f, \quad u_h \in D_h, \quad r_h f \in R_h, \tag{5}$$

of an approximate problem.

Assuming the approximate problem is well-posed, equation (5) has the unique solution

$$u_h = A_h^{-1} r_h f.$$

We call A_h^{-1} the approximate solution operator (bridge) (the middle solid arrow in the lower part of Figure 1). We want to be able to measure how far the approximate solution u_h is from the exact solution u. But u_h and u may be in different spaces (e.g., U_h may not be a subspace of U). Using a representative-solution bridge, namely a bounded linear operator s_h from U to U_h, we get a representative $s_h u$ (in U_h) of the

exact solution u (which is in U). We want our representative-solution bridges to be stable as a whole too, so that $\sup_{h\in H} \|s_h\| < \infty$. With this bridge in place, we have all the constituents of an approximation method

$$\left\{(A_h, U_h, F_h, r_h, s_h)\right\}_{h\in H} \tag{6}$$

for solving the exact problem. The bridges $\{r_h\}_{h\in H}$ and $\{s_h\}_{h\in H}$ connect the exact problem to the approximation method.

Now we can measure the difference between the approximate solution u_h and the representative $s_h u$ of the exact solution u. This difference $u_h - s_h u$ is known as the global approximation error. An approximation method is convergent for the exact problem if the global approximation error vanishes for all instances of the exact problem in the limit as the approximation is refined; that is, for all $f \in R$

$$\text{(convergence)} \qquad \lim_{h\to 0} \|u_h - s_h u\|_{U_h} = \lim_{h\to 0} \left\| A_h^{-1} r_h f - s_h u \right\|_{U_h} = 0. \tag{7}$$

This limiting process is denoted by the dashed arrow in U_h in Figure 1. Convergence is an analytic condition (it involves a lim, an infinite process) and that usually makes it difficult to prove directly. To prove convergence for an approximation method, we hope to find a bridge between this analytic condition and an algebraic one that is easier to prove. To do this, we need to introduce two new concepts, consistency and stability. Both of them are desired properties in a useful approximation method.

Consistency ensures that the approximation method is at least approximating the correct exact problem. Roughly speaking, consistency means that the exact solution (eventually) solves the approximate problem. More precisely, an approximation method (given in equation (6)) is consistent with the exact problem (from equation (3)) if the approximate representative $s_h u$ of the solution u of any instance of equation (4) of the exact problem solves the corresponding approximate problem (given in equation (5)) in the limit as the approximation is refined; that is,

$$\text{(consistency)} \qquad \lim_{h\to 0} \left\| A_h s_h u - r_h f \right\|_{F_h} = 0, \qquad \text{for all } f \in R, \tag{8}$$

where the quantity inside the norm is called the local approximation error. This limiting process is denoted by the dashed arrow in F_h in Figure 1. As with convergence, consistency depends on the choice of its corresponding norm, namely $\|\cdot\|_{F_h}$.

Stability ensures that the approximation method does not amplify the local approximation errors unboundedly. In other words, a stable method gives nearly the right answer to nearly the right problem. Precisely, an approximation method from equation (6) is stable if our approximate solution bridges $\left\{A_h^{-1}\right\}_{h\in H}$ are stable as a whole. That is,

$$\text{(stability)} \qquad \sup_{h\in H} \left\| A_h^{-1} \right\| < \infty, \tag{9}$$

where

$$\left\| A_h^{-1} \right\| := \sup_{0\neq f_h \in F_h} \frac{\left\| A_h^{-1} f_h \right\|_{U_h}}{\left\| f_h \right\|_{F_h}}. \tag{10}$$

Stability of an approximation method is an intrinsic property of the method and is not related to the exact problem (i.e., r_h and s_h are not involved in the definition of stability). Note that stability depends on both of the norms $\|\cdot\|_{U_h}$ and $\|\cdot\|_{F_h}$, whereas consistency or convergence each depends on only one of these norms, the former on $\|\cdot\|_{F_h}$ and the latter on $\|\cdot\|_{U_h}$.

Proving that the sufficiency of stability of a method for its convergence (i.e., stability implies convergence) is elementary: It just makes use of the definitions from equations (7)–(9). Here is the proof:

$$\lim_{h\to 0} \|u_h - s_h u\|_{U_h} = \lim_{h\to 0} \left\| A_h^{-1} r_h f - s_h u \right\|_{U_h}$$
$$\leq \sup_{h\in H} \left\| A_h^{-1} \right\| \lim_{h\to 0} \left\| r_h f - A_h s_h u \right\|_{F_h} = 0.$$

In other words, the global approximation error of a method comes from the local approximation error that is amplified as it is propagated by the method. Assuming that the local approximation error vanishes as the approximation is refined (consistency), if this amplification is bounded independently of the refinements of the approximation (stability), then the global approximation error also vanishes in the limit (convergence). Note that in the above proof of "stability implies convergence" we did not invoke the uniform boundedness principle, and it is in this sense that this direction of the proof is elementary. The uniform boundedness of the family of approximate solution operators A_h^{-1} was assumed (stability), and it was not concluded from their pointwise boundedness. On the other hand, proving necessity ("no stability implies no convergence") is nonelementary. If we wanted to prove it using elementary methods, we could take an unstable approximation method and try to construct a data point $f \in R$, which is blown up when fed to the method and then we would be done. This "counterexample approach," however, is more elusive than it seems, since finding an f that blows up is not always easy. Thanks to the uniform boundedness principle, in the particular case of U_h being fixed and equal to U for all h, Lax and Richtmyer [**1**] did not have to construct this "bad" f. Thanks to an "extension of the uniform boundedness principle" by Palencia and Sanz-Serna [**2**], in our case of U_h varying with h and not necessarily equal to U, we do not have to either, since without stability we know it exists (and in fact there are infinitely many such f's). Note that an extension of the uniform boundedness principle is necessary, since without it this principle is applicable only to a family of bounded linear operators going from a fixed Banach space to another fixed normed space, whereas here our approximate solution operators A_h^{-1} go from U_h to F_h, which vary when the approximation parameter h is varied.

In the next section, we will look at a well-known approximation method, namely Euler's forward difference approximation method, for an evolution ordinary differential equation and see how to apply the concepts discussed so far to the analysis of the stability and convergence of this method.

Consistency, stability, and convergence of Euler's forward difference approximation method in a unified framework

Consider the prototypical evolution ordinary differential equation (ODE)

$$u'(t) = au(t), \quad t \in (0, \tau], \quad u(0) =: f, \tag{11}$$

for some a and f in $\mathbb{R}$, which has the exact solution

$$u(t) = e^{at} f. \tag{12}$$

We can cast equation (11) as an instance of an exact problem (A, U, F) in the form of equation (4) by choosing U to be the space of twice continuously differentiable functions from $[0, \tau]$ to $\mathbb{R}$ with the sup-norm, F to be $\mathbb{R}$ (with its standard absolute-value

norm), and the exact problem operator A, connecting the solution u to the initial condition f, to be

$$Au = f, \tag{13}$$

with the domain $D := \{u \in U : u'(t) = au(t), t \in (0, \tau]\}$.

Picking a time step $h \in H := \mathbb{R}^+$, we can approximate the interval $[0, \tau]$ by the $(N_h + 1)$-point grid

$$I_h := \{nh : n = 0, 1, \ldots, N_h\},$$

where $N_h := \lfloor \tau / h \rfloor$. Denoting by $u^{n,h}$ our approximation to $u(nh)$, Euler's forward difference method (Euler's method henceforth) for solving equation (11) approximates the values of both the derivative and the function at the current time point nh, replacing the former with the slope $(u^{n+1,h} - u^{n,h})/h$ of the secant and the latter with $au^{n,h}$. To keep the notation clean, for the rest of this section, we suppress the subscript h in N_h and $u^{n,h}$ and instead write N and u^n, respectively, keeping in mind that these quantities are not constants (which is especially important when it comes to taking limits as $h \to 0$). Then, $u'(t) = au(t)$ becomes $(u^{n+1} - u^n)/h = au^n$, i.e., $u^{n+1} = (1 + ah)u^n$, and the differential equation from equation (11) becomes the finite difference equation

$$-\alpha_h u^n + 1u^{n+1} = 0, \quad n = 0, 1, \ldots, N - 1, \quad u^0 =: f,$$

where $\alpha_h := 1 + ah$. We can cast this difference equation in the form of equation (5), as an approximate problem arising from a method $(A_h, U_h, F_h, r_h, s_h)$ for solving the ODE of equation (11), by choosing U_h to be $\mathbb{R}^{N+1}$ with the sup-norm, i.e., $\|u_h\|_{\sup} := \sup_{j=0:N} |(u_h)_j|$, F_h to be $\mathbb{R}^{N+1}$ with the 1-norm, i.e., $\|f_h\|_1 := \sum_{j=0}^{N} |(f_h)_j|$, s_h to be the operator that restricts the solution u to the grid I_h,

$$s_h u := u|_{I_h} \equiv \left[u(0), u(h), \ldots, u(Nh)\right]^{\mathrm{T}}, \tag{14}$$

and the approximate problem operator A_h, connecting the approximate solution u_h to the approximate data $r_h f$, to be

$$A_h u_h = r_h f, \tag{15}$$

where

$$A_h := \begin{bmatrix} 1 & & & \\ -\alpha_h & 1 & & \\ & \ddots & \ddots & \\ & & -\alpha_h & 1 \end{bmatrix}, \quad u_h := \begin{bmatrix} u^0 \\ u^1 \\ \vdots \\ u^N \end{bmatrix}, \quad r_h f := \begin{bmatrix} f \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{16}$$

Recall that for an operator T from a normed space $(V, \|\cdot\|_\alpha)$ to normed space $(W, \|\cdot\|_\beta)$, the operator norm is defined as

$$\|T\| \equiv \|T\|_{\alpha,\beta} := \sup_{0 \neq v \in V} \frac{\|Tv\|_\beta}{\|v\|_\alpha}.$$

Note that both families $\{r_h\}_{h \in H}$ and $\{s_h\}_{h \in H}$ are bounded uniformly in h, as needed by the fundamental theorem of numerical analysis, since

$$\|r_h\| \equiv \|r_h\|_{1,1} := \sup_{0 \neq f \in F} \|r_h f\|_1 / |f| = \sup_{0 \neq f \in F} |f| / |f| = 1$$

and

$$\|s_h\| \equiv \|s_h\|_{\sup,\sup} := \sup_{0 \neq u \in U} \|s_h u\|_{\sup} / \|u\|_{\sup} \leq \sup_{0 \neq u \in U} \|u\|_{\sup} / \|u\|_{\sup} = 1,$$

both for all h.

The approximate problem from equation (15) has the solution

$$u_h = A_h^{-1} r_h f,$$

where

$$A_h^{-1} := \begin{bmatrix} 1 & & & & & & \\ \alpha_h & 1 & & & & & \\ \alpha_h^2 & \alpha_h & \ddots & & & & \\ \vdots & \alpha_h^2 & \ddots & 1 & & & \\ \alpha_h^{N-2} & \vdots & \ddots & \alpha_h & 1 & & \\ \alpha_h^{N-1} & \alpha_h^{N-2} & & \alpha_h^2 & \alpha_h & 1 & \\ \alpha_h^N & \alpha_h^{N-1} & \alpha_h^{N-2} & \cdots & \alpha_h^2 & \alpha_h & 1 \end{bmatrix}.$$

It is easy to see that the norm from equation (10) of an operator (defined by matrix multiplication) from a space with 1-norm to a space with sup-norm is the sup-norm of the matrix, given by the supremum of the absolute values of the entries of the matrix; i.e.,

$$\left\| A_h^{-1} \right\| \equiv \left\| A_h^{-1} \right\|_{1,\sup} = \sup_{\substack{i=0:N,\\ j=0:N}} \left| \left(A_h^{-1} \right)_{i,j} \right| = \sup_{k=0:N} \left\{ 1, \left| \alpha_h^k \right| \right\}. \tag{17}$$

We have stability if the family $\left\{ A_h^{-1} \right\}_{h \in H}$ of the solution operators is uniformly bounded; i.e., $\sup_{h \in H} \left\| A_h^{-1} \right\| < \infty$. Combining this with equation (17), we see that stability requires

$$\sup_{h \in H} \sup_{k=0:N} \left\{ 1, \left| \alpha_h^k \right| \right\} \equiv \sup_{\substack{k=0:N,\\ h \in H}} \left\{ 1, \left| \alpha_h^k \right| \right\} = \sup_{\substack{k=0:N,\\ h \in H}} \left\{ 1, \left| (1+ah)^k \right| \right\} < \infty,$$

which implies the requirement $|1 + ah| < 1$. Of course, if $a > 0$, the exact solution from equation (12) grows exponentially and our approximation method reproduces this growth (and instability), since in that case $1 + ah > 1$. But if $a < 0$, the exact solution decays exponentially, and for the approximate solution to try and follow it stably we need $-1 < 1 + ah < 1$, which gives the stability condition $-2 < ah < 0$.

It is now left to show that Euler's method from equation (15) is consistent with the exact problem from (13), which amounts to checking whether equation (15) is satisfied (as $h \to 0$) for the representative $s_h u$ of the exact solution. The definition of consistency (equation (8)) combined with equations (14) and (16) gives

$$\begin{aligned} \left\| A_h s_h u - r_h f \right\|_1 &= \left\| \begin{bmatrix} 1 & & & \\ -\alpha_h & 1 & & \\ & \ddots & \ddots & \\ & & -\alpha_h & 1 \end{bmatrix} \begin{bmatrix} u(0) \\ u(h) \\ \vdots \\ u(Nh) \end{bmatrix} - \begin{bmatrix} f \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_1 \\ &= \left| u(0) - f \right| + \sum_{j=1}^{N} \left| -\alpha_h u((j-1)h) + u(jh) - 0 \right| \\ &= \sum_{j=1}^{N} \left| u(jh) - (1+ah) u(jh-h) \right|, \end{aligned} \tag{18}$$

where the first summand in the middle equality is zero due to the initial condition of the ODE from equation (11).

Exploiting the twice continuous differentiability of u, we can replace $u(jh-h)$ in the last equality with its Taylor series representation in Lagrange remainder form $u(jh)-u'(jh)h+u''(t_j)h^2/2$, for some $t_j \in [jh-h,\, jh]$ (see, e.g., [3, p. 188]). Then $u(jh)-(1+ah)u(jh-h)$ in equation (18) becomes

$$\begin{aligned} u(jh) &- \Big(u(jh)-u'(jh)h+ahu(jh)+C(j,h)h^2\Big) \\ &= h\left(u'(jh)-au(jh)\right)-C(j,h)h^2 \\ &= -C(j,h)h^2, \end{aligned} \tag{19}$$

where $C(j,h) := (1+ah)u''(t_j)/2 - au'(jh)$ and the parenthesized difference in the middle equality vanished due to the ODE from (11). It is easy to see that $\left|C(j,h)\right|$ is bounded uniformly for all $j = 1:N$ and $h \in [0,1]$. Denoting by M such a uniform bound for $\left|C(j,h)\right|$ and using equation (19) in equation (18) gives

$$\left\|A_h s_h u - r_h f\right\|_1 = \sum_{j=1}^{N}\left|C(j,h)h^2\right| \le NMh^2 = \left\lfloor\frac{\tau}{h}\right\rfloor Mh^2 \le \frac{\tau}{h}Mh^2 = M\tau h. \tag{20}$$

Since the right-hand side of the relation above vanishes as h goes to 0, by the squeeze theorem, so does its left-hand side, which proves consistency. With its consistency and stability (for $-2 < ah < 0$) established, Euler's method (equation (15)) is now concluded to be convergent for the exact problem from equation (13), thanks to the fundamental theorem of numerical analysis (i.e., with consistency in place, "stability is equivalent to convergence").

For this significant example, it pays off to see the convergence directly from the definition too. Since $r_h f$, in equation (16), is nonzero only in its first entry (which is f), u_h is just the first column of A_h^{-1} scaled by f, i.e.,

$$u_h = \left[1, \alpha_h, \ldots, \alpha_h^j, \ldots, \alpha_h^N\right]^{\mathrm{T}} f.$$

For convenience, we copy equation (14) here:

$$s_h u = \left[u(0), u(h), \ldots, u(jh), \ldots, u(Nh)\right]^{\mathrm{T}}.$$

The entry j of the approximation error $s_h u - u_h$ is

$$E_{j,h} := \Big(u(jh) - (1+ah)^j f\Big), \quad j = 0:N, \tag{21}$$

from which it follows that

$$E_{j-1,h} = \Big(u((j-1)h) - (1+ah)^{j-1} f\Big).$$

Subtracting $(1+ah)E_{j-1,h}$ from $E_{j,h}$ in equation (21), we get

$$E_{j,h} - (1+ah)E_{j-1,h} = \Big(u(jh) - (1+ah)u((j-1)h)\Big) = -C(j,h)h^2, \tag{22}$$

where $C(j,h)$ is defined in the intermediate consistency result (equation (19)). Applying the triangle inequality to equation (22), we get

$$\left|E_{j,h}\right| \le |1+ah|\left|E_{j-1,h}\right| + \left|C(j,h)\right|h^2 \le |1+ah|\left|E_{j-1,h}\right| + Mh^2, \tag{23}$$

where M is a uniform bound for $\left|C(j,h)\right|$, as in equation (20). Notice how we used consistency to get equation (22). We "unroll" the recursive inequality $\left|E_{j,h}\right| \leq |1+ah|\left|E_{j-1,h}\right| + Mh^2$ in equation (23), starting with $E_{0,h} = 0$, to find a bound for $\left|E_{j,h}\right|$:

$$\begin{aligned}
\left|E_{1,h}\right| &\leq 0 + Mh^2 \\
\left|E_{2,h}\right| &\leq (|1+ah|+1)Mh^2 \\
\left|E_{3,h}\right| &\leq (|1+ah|^2 + |1+ah| + 1)Mh^2 \\
&\vdots \\
\left|E_{j,h}\right| &\leq (|1+ah|^{j-1} + \cdots + |1+ah| + 1)Mh^2 \\
&= \frac{|1+ah|^j - 1}{|1+ah| - 1} Mh^2, \quad \text{equality holds if } |1+ah| \neq 1. \qquad (24)
\end{aligned}$$

Recalling that $j = 0 : N$ and $N := \lfloor \tau/h \rfloor$, we see that

$$|1+ah|^j \leq (1+|ah|)^j \leq \left(e^{|ah|}\right)^j = e^{|ah|j} \leq e^{|a|hN} \leq e^{|a|h\lfloor \tau/h \rfloor} \leq e^{|a|\tau},$$

which together with the bound from equation (24) gives

$$\left|E_{j,h}\right| \leq \left(e^{|a|\tau} - 1\right) \frac{Mh^2}{\left||1+ah| - 1\right|},$$

and hence

$$\sup_j \left|E_{j,h}\right| \leq \left(e^{|a|\tau} - 1\right) \frac{Mh^2}{\left||1+ah| - 1\right|}, \qquad (25)$$

since the bound is independent of j. Finally, since $\|u_h - s_h u\|_{\sup} = \sup_j \left|E_{j,h}\right|$, we get

$$\lim_{h\to 0} \|u_h - s_h u\|_{\sup} = 0,$$

by applying the squeeze theorem to equation (25). This proves Euler's method (from equation (15)) is convergent for the exact problem of equation (13), by the definition from equation (7).

Equipped with the concept and application of stability from the last two sections, now we can see why equation (2) is an unstable approximation method for solving the problem in equation (1). Recall that $g_n := ng_{n-1}$, $g_0 := 1$, $n > 1$, is another way of writing $g_n := n!$, the factorial of n. Looking at the recursive definition $\widetilde{I}_n = 1 - (n-1)\widetilde{I}_{n-1}$, $n > 1$, with $\widetilde{I}_1$ being an approximation to I_1, we see that $\widetilde{I}_n$ must involve $(n-1)!\widetilde{I}_1$. In other words, the approximate solution operator given by equation (2) involves a multiplication by $(n-1)!$ of the initial condition $\widetilde{I}_1$. The family of these operators (consisting of different operators for different approximation levels) is, of course, not uniformly bounded (the operator norms grow factorially), which means that the method from equation (2) is unstable for solving equation 1. In our particular example, $\widetilde{I}_{20}$ involves a multiplication of $\widetilde{I}_1$ by $(20-1)! = 19! = 121645100408832000$, which is exactly the coefficient of a in the equation defining $\widetilde{I}_{20}$ that we found using symbolic algebra in Algorithm 1. This is how much any error in the approximate initial condition $\widetilde{I}_1$ is amplified by the method from equation (2) to get to $\widetilde{I}_{20}$.

Answer (To the exercise)**.** When implemented in floating-point arithmetic, the algorithm `Integral`$(20, 0.6321205588285577)$ gives 1.559619744279189, as may be seen by running the sample implementation at https://goo.gl/Sg3mJ9 in any web browser. This answer involves roundoff error accumulation over 20 steps, whereas the way we calculated $\widetilde{I}_{20}$ involves roundoff error only in the last step of its evaluation.

It may help the reader to reiterate that a "problem" $Au = f$ may be well-posed (i.e., A has a continuous left inverse) or not, whereas an "approximation method" $A_h u_h = r_h f$ used for solving that problem may be stable ($\sup_{h\in H} \left\| A_h^{-1} \right\|$ is finite) or not. Well-posedness (or lack thereof) is independent of stability (or lack thereof). However, these notions all pertain to the perturbation behavior (of the problem or method), e.g., in the presence of measurement, truncation, or roundoff errors. A well-posed problem may in turn be well-conditioned or not, depending on whether $\sup_{h\in H} \left\| A_h^{-1} \right\|$ is "small enough" or not. Obviously, the notion of well-conditioning is dependent on the context. To give a presentation applicable in different contexts, in this article, we have used the notion of well-posedness, and not well-conditioning.

Conclusion

We have presented a unifying picture for the "fundamental theorem of numerical analysis," illustrating that if we consider an approximation method as a family of solution operators acting on data, then a consistent approximation method is convergent if and only if it is stable. Since it requires some knowledge of the exact solution, convergence is harder to prove in many applications than consistency or stability is, which is why this equivalence theorem is fundamental in numerical analysis. As a concrete example of the application of this equivalence theorem, we have presented in our unified framework, the consistency, stability, and convergence (and their interplay) of Euler's method for the approximate solution of an ODE of evolution. This a simple and significant example: All multistep and Runge–Kutta methods are in a profound sense nothing but a generalization of Euler's method; also, this seemingly mundane ODE approximation method is at the heart of PDE approximation methods. In a future work, we present in this unified framework a more advanced application of this equivalence theorem to the stability and convergence of a combination of two variants of Euler's method for the approximate solution of an evolution PDE, with emphasis on the importance of checking the stability by looking at the norms of the solution operators rather than their spectral radii (which miss transient growth for non-normal operators).

Acknowledgment The title of this article was inspired by "Four Fundamental Subspaces" of Strang [4, Section 2.4], whom we thank for his extraordinary teaching, expository textbooks, and articles.

REFERENCES

[1] Lax, P.D., and Richtmyer, R.D. (1956). Survey of stability of linear finite difference equations. *Comm. Pure Appl. Math.*, 9: 267–293.

[2] Palencia, C., and Sanz-Serna, J.M. (1984). An extension of the Lax–Richtmyer theory. *Numer. Math.*, 44(2): 279–283.

[3] Strichartz, R.S. (2000). *The Way of Analysis*. Burlington, MA: Jones & Bartlett Learning.

[4] Strang, G. (2006). *Linear Algebra and Its Applications*. Belmont, CA: Thomson, Brooks/Cole.

[5] Trefethen, L.N. (2008). Numerical analysis. In: Gowers, T., Barrow-Green, J., Leader, I., eds. *The Princeton Companion to Mathematics*. Princeton, NJ: Princeton University Press, § IV.21, pp. 604–614.

Summary. Every field of mathematics seems to have a fundamental theorem to its name—two of the most well-known are those of calculus and algebra—and numerical analysis is no exception. The precise statement of the equivalence of stability and convergence of a consistent approximation method used for solving a well-posed problem has been called the fundamental theorem of numerical analysis. In this article, we first motivate the importance of stability by looking at an example of a definite integral. A recursive method and a corresponding algorithm for implementing it are given for computing the integral. The algorithm fails miserably due to the instability of the method! We then give a picture that in a unified way introduces the reader to the fundamental theorem of numerical analysis and the operators and spaces associated with it. This picture also illustrates the proof of the theorem in the forward direction (stability plus consistency implies convergence, given well-posedness). For the converse, we present the idea of a proof. We close with a careful application of this fundamental theorem to Euler's forward difference approximation method for an ordinary differential equation, presented in our unified framework.

OMID KHANMOHAMADI (MR Author ID: 861820) has earned a Master's in Mechanical & Aerospace Engineering as well as Master's and Ph.D. in Applied & Computational Mathematics. His research has spanned a wide range of topics, including numerical integration on manifolds, spectral methods for PDEs and inverse problems, numerical analysis of nonnormal operators, and high performance computational software design.

EHSSAN KHANMOHAMMADI (MR Author ID: 916345) received his Ph.D. in Mathematics from The Pennsylvania State University. He works in the overlap of operator algebras, harmonic analysis, and representation theory of Lie groups.

Partiti Puzzle

15	11	20	13	5	16
11	8	6	6	13	11
14	12	11	3	9	12
13	6	10	6	12	9
11	6	13	10	8	6
21	7	16	6	13	18

How to play. In each cell, place one or more distinct integers from 1 to 9 so that they sum to the value in the top left corner. No integer can be used more than once in horizontally, vertically, or diagonally adjacent cells. For an introduction to the Partiti puzzle, see Caicedo, A. E., Shelton, B. (2018). Of puzzles and partitions: Introducing Partiti. *Math. Mag.* 91(1): 20–23. The solution is on page 306.

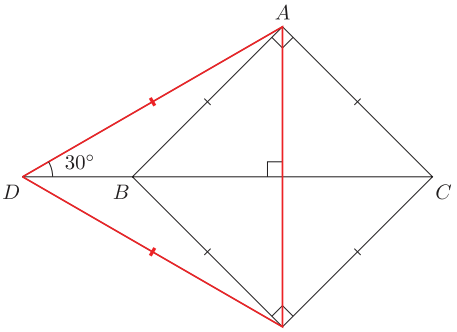
—contributed by Lai Van Duc Thinh,
Vietnam; fibona2cis@gmail.com

Proof Without Words: An Elegant Property of an Isosceles Right Triangle

VICTOR OXMAN
Western Galilee College
Acre, 24121, Israel
victor.oxman@gmail.com

MOSHE STUPEL
Shaanan College
Haifa, 26109, Israel
stupel@bezeqint.net

Given an isosceles right triangle BAC, $\angle A = 90°$, and a point D on the extension of side BC for which $AD = BC$. Then $\angle ADB = 30°$.



Summary. Using a picture, we show an elegant property of an isosceles right triangle.

VICTOR OXMAN (MR Author ID: 756975) teaches at Western Galilee College. His research interests are in geometry and mathematical education.

MOSHE STUPEL (MR Author ID: 1028564) teaches at Shanaan College. His research interests are also in geometry and mathematical education.

Math. Mag. **91** (2018) 254. doi:10.1080/0025570X.2018.1503035
MSC: Primary 51M04
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

The Order of the k-Letter Spelling Shuffle

JAY CUMMINGS
California State University
Sacramento, CA 95819
Jay.Cummings@csus.edu

A mathemagician gives a deck of cards to an audience member and asks her to shuffle it thoroughly. The spectator is then asked to name her favorite mathematician—suppose "Paul Erdős" is chosen—and to cut off roughly the top quarter of the deck. The mathemagician takes both stacks back (we are done with the larger stack) and demonstrates a *spelling deal* on the smaller deck. That is, he spells out PAUL ERDOS one letter at a time, and for each letter he places the top card from the quarter deck onto the table into a growing stack, and then he drops the remaining cards from the quarter deck on top of the dealt stack. In essence, this takes the top nine cards, reverses their order, and moves them to the bottom of the deck.

Then, "to make sure the deck is thoroughly messed up," this quarter deck is handed back to the audience member who performs the same spelling deal twice more.

The mathematician then claims to be able to sense what the top card in the deck is. After making his claim, the audience member dramatically holds up the correct guess to thunderous applause.

There are two small dynamic elements to this trick which are easy to miss. First, the fact that a quarter of the deck was asked for. If the chosen mathematician has a name with k letters, then we need to ensure that deck has between k and $2k$ cards in it. With a nine-letter name like Paul Erdős, a quarter of the deck works well. If Srinivasa Ramanujan is selected, then spell carefully and ask for a half deck.

Second, when the mathemagician takes the two decks from the spectator, he inconspicuously glanced at the card on the bottom of the small stack—this is the card that will wind up on top, due to the following principle.

Principle 1. (Mulcahy [1]) If a k-letter spelling shuffle is performed three times to a deck of n cards where $k \leq n \leq 2k$, then the card that was initially on the bottom of the deck will end up on top of the deck.

In fact, much more structure is maintained. The following more general principle describes this, and contains Principle 1 as a corollary.

Principle 2. (Mulcahy [1]) If a k-letter spelling shuffle is performed four times with a deck of n cards where $k \leq n \leq 2k$, then the entire deck will end up in the same order in which it began.

This principle has a somewhat strict limit on the size of the deck. With the goal to further understand the theory of repeated spelling shuffles, and to develop new magic tricks on larger decks, in 2003 Colm Mulcahy asked how many times one must perform a k-letter spelling shuffle with an n card deck in order to return the deck to its original ordering. We denote this quantity by SPELL(n, k) and in this note we find the answer, of which the above principles will be corollaries.

We will prove the following. Let n be a positive integer, $k \in \{1, 2, \ldots, n\}$, and write $n = qk + b$ where $b \in \{0, 1, \ldots, k-1\}$. Then

Math. Mag. **91** (2018) 255–259. doi:10.1080/0025570X.2018.1485400
MSC: Primary 05A05, Secondary 00A08
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

$$\text{SPELL}(n,k) = \begin{cases} n & \text{if } k = 1; \\ 2q & \text{if } k > 1 \text{ and } b = 0; \\ q(q+1) & \text{if } k > 1 \text{ and either } b = 1 \text{ and } q \text{ is odd} \\ & \qquad\qquad \text{or } b = k-1 \text{ and } q \text{ is even}; \\ 2q(q+1) & \text{otherwise.} \end{cases}$$

If $k = 1$, then trivially $\text{SPELL}(n,k) = n$, so assume $k > 1$. Label the cards $1, 2, \ldots, n$ from top to bottom, so that the ith card initially occupies position i. The easiest case is when k divides n; that is, when $b = 0$. Suppose that $n = qk$ for some integer q. After the first spelling shuffle the first k cards will be in the final k positions, but in the reverse order (the ith card in position $n - i + 1$). After q spelling shuffles the first k cards will have returned to the first k positions, the next k will have returned to positions $k+1, k+2, \ldots, 2k$, and so on; however, each set of k cards will be in the reverse order from which they began. After another q spelling shuffles the deck will, for the first time, return to its original order.

Thus, we currently have

$$\text{SPELL}(n,k) = \begin{cases} n & \text{if } k = 1; \\ 2q & \text{if } k > 1 \text{ and } b = 0. \end{cases}$$

Example

To build intuition for the rest of the proof, let us take a macroscopic look at the case when $n = 13$ and $k = 5$ (Figure 1).

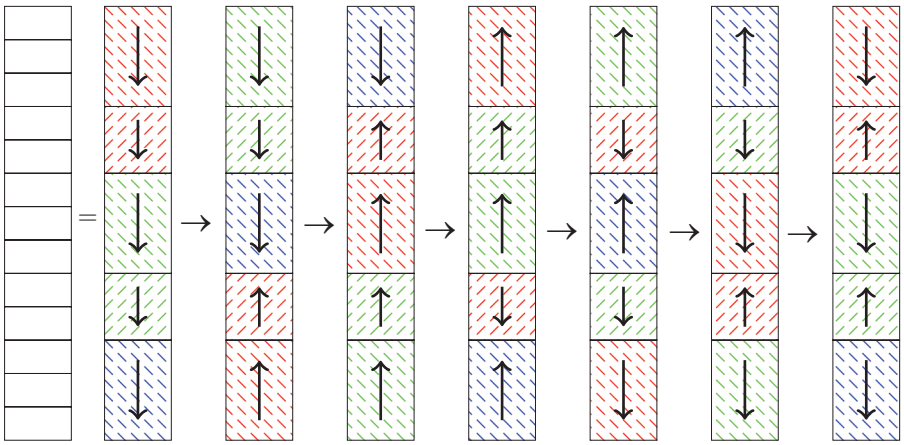


Figure 1 Example when $n = 13$ and $k = 5$. SPELL(13, 5) = 12.

The left-most rectangle is our stack of 13 cards, where the top of the stack is card 1 in position 1. The second rectangle is exactly the same as the first, except that we grouped some of the cards together with colors, and the arrows represent the fact that the cards' labels are currently in increasing order within each group. Each subsequent rectangle represents the result of a five-letter spelling shuffle. Notice how our initial partition is never broken by a shuffle. The last rectangle is the result of six spelling shuffles, and it is the first time that each group of cards has returned to their initial positions.

This is not a complete cycle, though, since cards 4 and 5 are in the wrong order (card 4 is in position 5 and card 5 is in position 4); likewise for cards 9 and 10. Thus, it can be

seen that after an additional six shuffles all the card groups will again be in the correct positions, and moreover all will be in increasing order. Thus SPELL$(13, 5) = 12$.

The general case is an extension of this example. If $n = qk + b$, then we will break up each set of k cards into the first b and the next $k - b$; then we will show that these blocks of cards cycle with a period of $q(q + 1)$; finally, based on parity considerations, we might have to apply another $q(q + 1)$ shuffles to return to our initial ordering. We now go over those details.

Proof and discussion

We will look at the positions that a fixed card cycles through and find the length of that cycle. The least common multiple of the cycle lengths of all the cards will give us our answer. We already have the answer if $k = 1$ or k divides n, so assume that neither of these holds. Let $n = qk + b$, where $b \in \{1, 2, \ldots, k - 1\}$.

As mentioned, to analyze the first set of k cards, we divide them into two blocks: the first b cards and the next $k - b$.

The first b cards. Consider the card that is initially in position i for some $i \in \{1, 2, \ldots, b\}$. We want to know how many moves it takes for this card to return to one of the first k positions, and in which position it will occupy at this point.

After the first shuffle, it will be in position $n + 1 - i = (n - b) + (b + 1 - i)$. Since $i \leq b$, we see that the card will return to one of the first k positions in an additional $(n - b)/k = q$ shuffles, and it will land on position $b + 1 - i \in \{1, 2, \ldots, b\}$. Moreover, by repeating the above it is easy to see that, from position $b + 1 - i$, the next time the card returns to one of the first k positions it will again occupy position i. Therefore, if $i = b + 1 - i$, then card i will cycle with period $q + 1$. If not, then its cycle has length $2(q + 1)$.

The next k − b cards. If instead $i \in \{b + 1, b + 2, \ldots, k\}$, then after the first shuffle card i will be in position $n + 1 - i = (n - b - k) + (k + b + 1 - i)$, and after an additional $(n - b - k)/k = q - 1$ moves will land in position $k + b + 1 - i \in \{b + 1, b + 2, \ldots, k\}$. So these cards cycle with period q if $i = b + 1 - i + k$, and $2q$ otherwise.

The rest. Every card will eventually land among the top k positions. Suppose card j's first appearance in the top k positions is to position i. Then, card j and card i are in the same cycle and hence have the same cycle length. Thus, when taking the least common multiple of all cycle lengths, it is unnecessary to consider any cards except for the initial k; that is, we may focus on just the above two cases.

We wish now to take the least common multiple of all of these k cycle lengths. If we assume that $b \in \{2, 3, \ldots, k - 2\}$ and $k \geq 4$, then each of the first two cases above will consist of at least two values of i, and hence at least one value in the first case must have period $2(q + 1)$, and at least one in the second must have period $2q$.

Therefore, if $b \in \{2, 3, \ldots, k - 2\}$, then we may conclude that the whole deck cycles with period

$$\text{LCM}\{2(q + 1), 2q\} = 2q(q + 1).$$

The cases when $b = 1$ or $b = k - 1$ can be different because a block of size 1 has only one possible ordering. In Figure 1, if the blocks of size 2 were actually of size 1, then after six spelling shuffles the cards would have returned to their original ordering. More explicitly, Figure 2 below shows an such an example, in which $n = 8$ and $k = 3$.

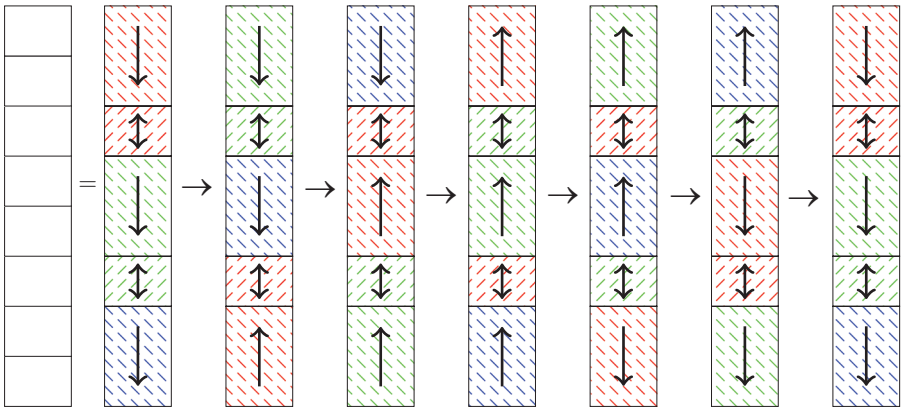


Figure 2 Example when $n = 8$ and $k = 3$. SPELL$(8, 3) = 6$.

Now consider the case that $b = 1$, and refer back to our previous discussion on the first b cards and the next $k - b$. The first b cards then have $i = 1$ as the only possibility and hence this card cycles with period $q + 1$. If $k = 2$, then the next $k - b$ have $i = 2$ as the only possibility, and hence this card cycles with period q. Therefore, the whole deck cycles with period $q(q + 1)$.

If, however, $b = 1$ but $k > 2$, then by the above reasoning we readily see that the deck cycles with period

$$\text{LCM}\{q + 1, 2q\} = \begin{cases} q(q + 1) & \text{if } q \text{ is odd;} \\ 2q(q + 1) & \text{if } q \text{ is even.} \end{cases}$$

Similarly, if $b = k - 1$ and $k > 2$, then the deck has period

$$\text{LCM}\{2(q + 1), q\} = \begin{cases} q(q + 1) & \text{if } q \text{ is even;} \\ 2q(q + 1) & \text{if } q \text{ is odd.} \end{cases}$$

By combining cases, we get our result

$$\text{SPELL}(n, k) = \begin{cases} n & \text{if } k = 1; \\ 2q & \text{if } k > 1 \text{ and } b = 0; \\ q(q + 1) & \text{if } k > 1 \text{ and either } b = 1 \text{ and } q \text{ is odd} \\ & \qquad \text{or } b = k - 1 \text{ and } q \text{ is even;} \\ 2q(q + 1) & \text{otherwise.} \end{cases}$$

Note that Principle 2 follows from this theorem. If $k = 1$, it follows trivially, so assume $k > 1$. If $n = k$ or $n = 2k$, then SPELL$(n, k) = 2$ or SPELL$(n, k) = 4$, respectively, and so in either case the deck will be in its initial order after four shuffles. Lastly, if $k < n < 2k$, then $q = 1$ and either SPELL$(n, k) = q(q + 1) = 2$ or SPELL$(n, k) = 2q(q + 1) = 4$, and again we see that in either case the deck will be in its initial order after four shuffles.

One could also use the above theorem to create analogous principles. For example, if $2k \le n \le 3k$, then either k divides n or $q = 2$. The possible periods are 2, 3, 4, 6 or 12, and the least common multiple of all of these is 12. So with 12 shuffles one is guaranteed for the deck to cycle. Likewise, if $3k \le n \le 4k$, then the deck will certainly cycle with 24 shuffles.

Indeed, when $qk \leq n \leq (q+1)k$, the theorem says that the possible periods are $n, 2q, q(q+1), 2q(q+1)$ or $2(q+1)$ (the last in the event that $n = (q+1)k$). Moreover, the answer is n only if $k = 1$, which would imply that $n = q$. The deck, therefore, certainly cycles after

$$\text{LCM}\{q, 2q, q(q+1), 2q(q+1)\} = 2q(q+1)$$

shuffles. Thus, we have deduced a more general principle from which the previous two follow.

Principle 3. If a k-letter spelling shuffle is performed $2q(q+1)$ times with a deck of n cards where $qk \leq n \leq (q+1)k$, then the entire deck will end up in the same order in which it began. If you instead stop after $2q(q+1) - 1$ shuffles, then the card that began at the bottom of the deck will now be at the top.

Further directions

When a spelling shuffle is performed a number of times using the same word each time, the cycle structure is easier to study. In this note's trick, for example, "Paul Erdős" was treated as a single word of length nine, and we did a nine-letter spelling shuffle with a deck a number of times. A similar trick can be done by performing a spelling shuffle for each word in a sentence, provided each such word has the same length and the number of words in the sentence is chosen just right.

What if we relax the condition that all the words have the same length? Suppose, for instance, we executed a four-letter spelling shuffle for "Paul" and then a five-letter spelling shuffle for "Erdős," and then repeated this a number of times? This suggests the following question.

Open problem. Define SPELL(n, j, k) to be the minimum number of spelling shuffles of an n card deck for the deck to return to its original ordering, where you alternately perform a j-letter shuffle and then a k-letter shuffle. Determine SPELL(n, j, k).

Further work in the cycle structure of the top card or the bottom card would also be interesting, since those cards are most often used in card tricks.

Acknowledgments The author thanks Ron Graham and Colm Mulcahy for introducing him to both this problem and to mathemagics, in general. He also thanks the anonymous referees whose excellent suggestions improved the quality of this article.

REFERENCE

[1] Mulcahy, C. (2013). *Mathematical Card Magic: Fifty-Two New Effects*, 1st ed. Natick, MA: A K Peters/CRC Press.

Summary. Most magic tricks amaze by the subtle use of small lies: a well-practiced sleight of hand, or equipment with a deceptive property. Others, though, require no lies at all–they are self-working tricks, relying only on the magic and mystery that is inherent in mathematics. In this article, we study the mathematics behind the k-letter spelling shuffle, on which several mathemagical tricks are based.

JAY CUMMINGS (MR Author ID: 1090648) is an Assistant Professor of Mathematics at the California State University, Sacramento, CA, USA. He received his Ph.D. in 2016 from the University of California, San Diego studying combinatorics under the guidance of Ron Graham. His research in combinatorics continues, and he particularly enjoys involving students in his work. He is passionate about his teaching and enjoys designing new courses, in part in an effort to add mathemagics to the curriculum.

Proof Without Words: Filling in Pythagorean Gaps

CHARLES F. MARION
Yorktown Heights, NY 10598
charliemath@optonline.net

The following patterns were considered in [2]:

$$3^2 + 4^2 = 5^2 \qquad 6^2 + 7^2 = 2^2 + 9^2$$
$$10^2 + 11^2 + 12^2 = 13^2 + 14^2 \qquad 15^2 + 16^2 + 17^2 = 3^2 + 19^2 + 20^2$$
$$21^2 + 22^2 + 23^2 + 24^2 = 25^2 + 26^2 + 27^2 \qquad 28^2 + 29^2 + 30^2 + 31^2 = 4^2 + 33^2 + 34^2 + 35^2$$
$$\vdots \qquad \vdots$$

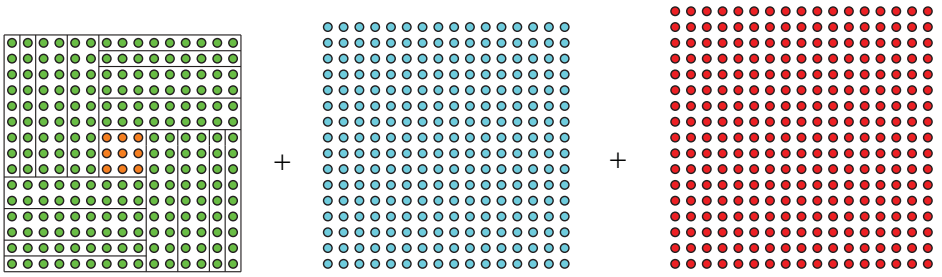
The left equalities can be written as, for $n \geq 1$,

$$\left(2n^2 + n\right)^2 + \left(2n^2 + n + 1\right)^2 + \cdots + \left(2n^2 + 2n\right)^2 = \left(2n^2 + 2n + 1\right)^2 + \cdots + \left(2n^2 + 3n\right)^2.$$

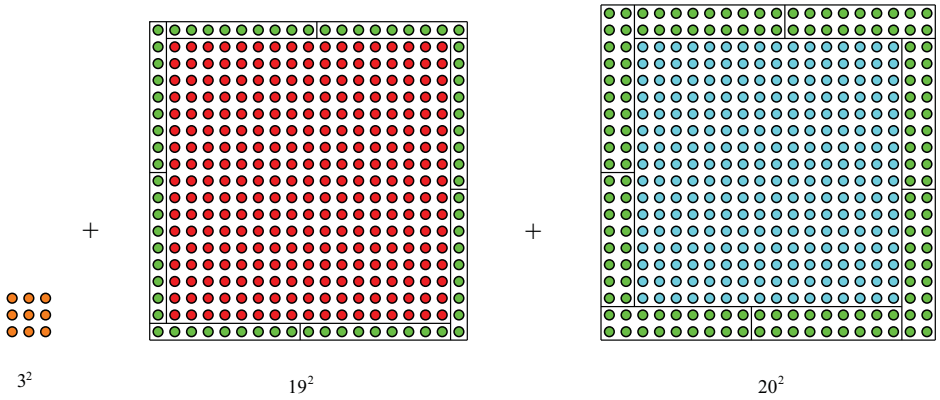
Boardman [1] proved this general equality without words. The right equalities generalize as, for $n \geq 1$,

$$\left(2n^2 + 3n + 1\right)^2 + \left(2n^2 + 3n + 2\right)^2 + \cdots + \left(2n^2 + 4n + 1\right)^2 = \left(n + 1\right)^2 + \left(2n^2 + 4n + 3\right)^2 + \cdots + \left(2n^2 + 5n + 2\right)^2.$$

To complement [1], I prove the second equality without words. For example, for $n = 2$:



Math. Mag. **91** (2018) 260–261. doi:10.1080/0025570X.2018.1466585
MSC: Primary 11D09
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.



REFERENCES

[1] Boardman, M. (2009). Proof without words: Pythagorean Runs. *Math. Mag.* 73(1): 59. doi.org/10.2307/2691496.

[2] Marion, C. F. (2008). Filling in the Gaps. *Math. Teacher.* 101(8): 618–620.

Summary. Boardman proved, without words, a quadratic identity featuring Pythagorean runs. We prove a similar quadratic identity featuring Pythagorean runs using most of the positive integers he *did not* use.

CHARLES F. MARION (MR Author ID: 1211614, ORCID 0000-0001-6620-3896) taught mathematics at Lakeland High School (Shrub Oak, NY) for 32 years. He is thrilled that his daughter, Kristin, has transitioned to teaching mathematics at Eastview Middle School (White Plains, NY).

Artist Spotlight: Veronika Irvine

Delle Caustiche, Veronika Irvine; cotton thread and copper wire, 2017. The pattern features clusters of six sided stars arranged in a honeycomb. The 3D shape and the diaphanous nature of the lace suggested a cloud. The piece is named after the star-cloud Delle Caustiche because of the words Agnus Mary Clerke used to describe the nebula: “a peculiar arrangement of stars in rays, arches, caustic curves and intertwined spirals”.

See interview on page 307–309.

An Amusing Sequence of Functions

STEFAN STEINERBERGER
Yale University
New Haven, CT 06520
stefan.steinerberger@gmail.com

We introduce an amusing sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ given by

$$f_n(x) = \sum_{k=1}^{n} \frac{|\sin(k\pi x)|}{k}.$$

A quick look at the graph (see Figure 1) shows that the function has many local minima: we will prove that for every rational number $x = p/q \in \mathbb{Q}$ there exists an $n_0 \in \mathbb{N}$ such that the function f_n has a strict local minimum at $x = p/q$ for all $n \geq n_0$.

Theorem. *Let $p, q \in \mathbb{Z}$ and $q \neq 0$. The function $f_n(x)$ has a strict local minimum at the point $x = p/q \in \mathbb{Q}$ for all $n \geq q^2$.*

We believe that this curious result is a good indicator that this sequence might have many other unusual properties and could be of some interest. A natural question would be whether anything can be said about the location of local maxima: it is tempting to conjecture that they cannot be too well approximated by rationals with small denominators (because that is where the minima are). Another natural question is whether it is possible to determine the enveloping curve so clearly seen in Figure 1.

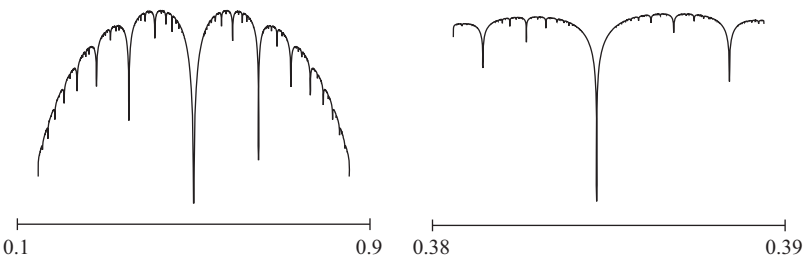


Figure 1 The function f_{50} on $[0.1, 0.9]$ and zoomed in (right). The big cusp in the right picture is located at $x = 5/13$, the two smaller cusps are at $x = 8/21$ and $x = 7/18$.

Both the definition of the function f_n as well as its graph (see Figures 1 and 2) are reminiscent of the Takagi function τ. The Takagi function is continuous but nowhere differentiable and was first considered by Takagi [**15**] in 1903 (with independent rediscoveries by van der Waerden [**16**] in 1930 and de Rham [**12**] in 1957): If $d : \mathbb{R} \to [0, 1/2]$ is the distance-to-the-nearest integer function defined by

$$d(x) = \min\{x - \lfloor x \rfloor, \lceil x \rceil - x\}, \quad \text{then} \quad \tau(x) = \sum_{k=0}^{\infty} \frac{d(2^k x)}{2^k}.$$

Math. Mag. **91** (2018) 262–266. doi:10.1080/0025570X.2018.1491768
MSC: Primary 26A06, Secondary 26A27

The Takagi function τ has since appeared in connection to inequalities for digit sums [1], the Riemann hypothesis [3] and extremal combinatorics [6] (many more results can be found in the surveys [2, 10]). We emphasize a 1959 result of Kahane [9] who proved that the set of points where τ has local minima is exactly the set of dyadic rational numbers.

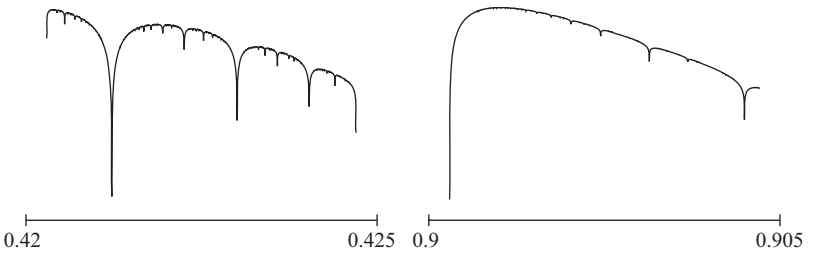


Figure 2 The function f_{50} in two other locations.

The study of functions with intriguing properties that are given by a series representation is classical. An example from Riemann's 'Habilitationsschrift' [13] is

$$\sum_{n=1}^{\infty} \frac{(nx + 1/2) - \lfloor nx + 1/2 \rfloor}{n}.$$

Series in this spirit also appeared in the classical work of Weierstrass (see [14] for a recent variation), Hecke [7], Davenport [4], Davenport [5], and Levy [11]. We refer to a article of Jaffard [8] for a more extensive survey. None of these series seem to have the property of having minima at rational points, which distinguishes the sequence f_n. Another noteworthy difference is that f_n is given as the sum of n relatively simple continuous functions as opposed to the limit of an infinite series. It remains to be seen whether a more detailed and systematic investigation of f_n as n tends to infinity could benefit from ideas developed in this more classical setting.

The proof

We observe that for $x \in \mathbb{R}$ and $\varepsilon \to 0$

$$|\sin(x+\varepsilon)| - |\sin(x)| = \begin{cases} |\varepsilon| + \mathcal{O}(\varepsilon^2) & \text{if } x/\pi \in \mathbb{Z}, \\ \varepsilon \operatorname{sgn}(\sin(x)) \cos(x) + \mathcal{O}(\varepsilon^2) & \text{otherwise.} \end{cases}$$

Here, sgn denotes the signum function

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

Let us now consider the function at $x = p/q$ with $p, q \in \mathbb{Z}$ and $\gcd(p, q) = 1$. We want to understand

$$f_n\left(\frac{p}{q} + \varepsilon\right) - f_n\left(\frac{p}{q}\right) = \sum_{k=1}^{n} \left(\frac{|\sin(k\pi(p/q + \varepsilon))|}{k} - \frac{|\sin(k\pi(p/q))|}{k} \right)$$

for small values of ε. Using the formula above, we get

$$f_n\left(\frac{p}{q}+\varepsilon\right)-f_n\left(\frac{p}{q}\right)=\pi\varepsilon\sum_{k=1}^{n}\operatorname{sgn}\left(\sin\left(\frac{k\pi p}{q}\right)\right)\cos\left(\frac{k\pi p}{q}\right)$$
$$+\pi|\varepsilon|\#\{1\le k\le n : k(p/q)\in\mathbb{Z}\}+\mathcal{O}(\varepsilon^2).$$

We analyze these two sums and show that the first one is bounded in n while the second term is clearly unbounded. The function $\operatorname{sgn}(\sin(x))\cos(x)$ has period π and the map $k\to k\cdot p$ is a permutation on $\mathbb{Z}_q$. It is then easy to see that the symmetries of sine and cosine imply

$$\sum_{k=1}^{q}\operatorname{sgn}\left(\sin\left(\frac{k\pi p}{q}\right)\right)\cos\left(\frac{k\pi p}{q}\right)=0$$

and thus

$$\sum_{k=m+1}^{m+q}\operatorname{sgn}\left(\sin\left(\frac{k\pi p}{q}\right)\right)\cos\left(\frac{k\pi p}{q}\right)=0\qquad\text{for all } m\in\mathbb{N}.$$

Since $k\to k\cdot p$ is a permutation, we have

$$\inf_{n\in\mathbb{N}}\sum_{k=1}^{n}\operatorname{sgn}\left(\sin\left(\frac{k\pi p}{q}\right)\right)\cos\left(\frac{k\pi p}{q}\right)$$
$$=\min_{1\le n\le q}\sum_{k=1}^{n}\operatorname{sgn}\left(\sin\left(\frac{k\pi p}{q}\right)\right)\cos\left(\frac{k\pi p}{q}\right)\ge-\max_{1\le n\le q}\sum_{k=1}^{n}\cos\left(\frac{k\pi}{q}\right)\ge-\frac{q}{2},$$

while, at the same time, we obviously have

$$\#\{1\le k\le n : k(p/q)\in\mathbb{Z}\}\ge\left\lfloor\frac{n}{q}\right\rfloor.$$

This means that

$$f_n\left(\frac{p}{q}+\varepsilon\right)-f_n\left(\frac{p}{q}\right)\ge\left(-\frac{q}{2}+\left\lfloor\frac{n}{q}\right\rfloor\right)\pi|\varepsilon|+\mathcal{O}(\varepsilon^2),$$

which shows that f_n has a strict local minimum at p/q for $n\ge q^2$. ■

A sharpened form A slightly more careful analysis shows that

$$-\max_{1\le n\le q}\sum_{k=1}^{n}\cos\left(\frac{k\pi}{q}\right)=-(1+o(1))\frac{q}{2}\frac{2}{\pi}\int_0^{\pi/2}\cos x\,dx=-(1+o(1))\frac{q}{\pi}$$

from which we get that, asymptotically, $n\ge(1+o(1))q^2/\pi$ suffices. We observe that this is optimal: if $p=q-1$ and $1\le k\le q/2$, then

$$\operatorname{sgn}\left(\sin\left(\frac{k\pi p}{q}\right)\right)\cos\left(\frac{k\pi p}{q}\right)=-\cos\left(\frac{k\pi}{q}\right)$$

and

$$\sum_{k=1}^{\lfloor q/2 \rfloor} -\cos\left(\frac{k\pi}{q}\right) \sim -\frac{q}{\pi}.$$

Some variations

The main argument only used some fairly elementary symmetry properties of the trigonometric functions and also extends to various other functions. A particularly nice example comes from replacing the sine by the cosine

$$g_n(x) = \sum_{k=1}^{n} \frac{|\cos(k\pi x)|}{k}.$$

Whether $x = p/q$ is the location of a local minimum or maximum now depends on whether q is even or odd (see Figure 3, left).

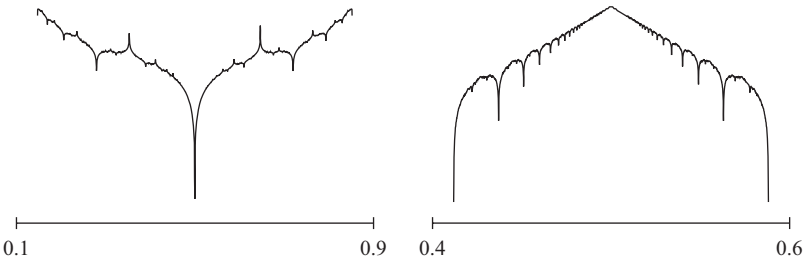


Figure 3 The functions g_{50000} and h_{500}.

Another fascinating example (see Figure 3, right) is given by

$$h_n(x) = \sum_{k=1}^{n} \frac{d(kx)}{k},$$

where d is again the distance to the nearest integer. While sharing similar overall characteristics, there is now also a new peculiar behavior around $x = 1/2$. Is it possible to understand how this shape arises? We observe that all these sequences of functions, f_n, g_n, h_n, and other variations of this type, do not have a limit as $n \to \infty$ and this may be a reason why they have been overlooked. However, clearly they are highly structured as n becomes large—maybe their asymptotic behavior can be understood?

REFERENCES

[1] Allaart, P. C. (2011). An inequality for sums of binary digits, with application to Takagi functions, *J. Math. Anal. Appl.* 381(2): 689–694.
[2] Allaart, P. C., Kawamura, K. (2011/12). The Takagi function: a survey. *Real Anal. Exchange* 37(1): 1–54.
[3] Balasubramanian, R., Kanemitsu, S., Yoshimoto, M. (2006). Euler products, Farey series, and the Riemann hypothesis. II. *Publ. Math. Debracen* 69(1–2): 1–16.
[4] Davenport, H. (1937). On some infinite series involving arithmetical functions. *Quart. J. Math.* 8: 8–13.
[5] Davenport, H. (1937). On some infinite series involving arithmetical functions II. *Quart. J. Math.* 8: 313–320.

[6] Frankl, P., Matsumoto, M., Rusza, I. Z., Tokushige, N. (1995). Minimum shadows in uniform hypergraphs and a generalization of the Takagi function. *J. Comb. Theory A* 69: 125–148.
[7] Hecke, E. (1921). Uber analytische funktionen und die Verteilung von Zahlen mod. Eins, Hamburg. *Math. Abh.* 1: 54–76.
[8] Jaffard, S. (2004). On Davenport expansions. (English summary) Fractal geometry and applications: a jubilee of Benoit Mandelbrot. *Part 1. Proc. Sympos. Pure Math.*, Providence, RI: Amer. Math. Soc., 273–303.
[9] Kahane, J.-P. (1959). Sur l'exemple, donne par M. de Rham, d'une fonction continue sans derivee. *Enseign. Math.* 2:5, 53–57.
[10] Lagarias, J. (2012). The Takagi function and its properties. Functions in number theory and their probabilistic aspects. *RIMS Kokyuroku Bessatsu* 153–189.
[11] Levy, P. (1965). Processus stochastiques et mouvement Brownien. Paris, France: Gauthier-Villars.
[12] de Rham, G. (1957). Sur un exemple de fonction continue sans derivee. *Enseign. Math.* 3:71–72.
[13] Riemann, B. (1953). Uber die darstellbarkeit einer funktion durch eine trigonometrische reihe, Habilitation thesis (1854). in: Collected works of Bernard Riemann. Mineola, NY: Dover.
[14] Stoica, G. (2016). Generating continuous nowhere differentiable functions. *Math. Mag.* 89: 371–372.
[15] Takagi, T. (1903). A simple example of the continuous function without derivative. *Proc. Phys. Math. Japan* 1: 176–177.
[16] van der Waerden, B. L. (1930). Ein einfaches Beispiel einer nichtdifferenzierbaren stetigen funktion *Math. Z.* 32: 474–475.

Summary. We consider the amusing sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ given by

$$f_n(x) = \sum_{k=1}^{n} \frac{|\sin(k\pi x)|}{k}.$$

Every rational point is eventually the location of a strict local minimum of f_n: more precisely, f_n has a strict local minimum in all rational points $x = p/q \in \mathbb{Q}$ with $|q| \leq \sqrt{n}$.

STEFAN STEINERBERGER (MR Author ID: 869041) obtained an undergraduate and master's degree from the Johannes Kepler University in Linz, Austria, a PhD from the University of Bonn, Germany and has since been at Yale. He is very fond of both $\sin x$ and $\cos x$ and does not have the same kind of feelings about $\tan x$.

Speculations by Veronika Irvine

The patterns on pages 261, 285, 287, and 320 were part of a study of the planar crystallographic symmetries possible in bobbin lace. The patterns were generated using an exhaustive combinatorial search on a fixed set of generators selected based on the desired symmetry group. Patterns from all 17 wallpaper symmetry groups were generated in the study.

Save This Family! An Instructive Sequence of Curves

MARC FRANTZ
Indiana University
Bloomington, IN 47405
mfrantz@indiana.edu

This is the story of a family (a two-sided sequence of curves $\mathbf{F}_{2^n}$) that suffered a seemingly irreconcilable breakup, only to be saved by a new point of view that showed their differences were not so extreme after all. One branch of the family migrated in one direction, achieving an existence that was conventional and well-rounded (converging to a circle as n goes to $-\infty$). The "wild" branch of the family migrated in another direction, in the final moment seeming to disintegrate into a handful of dust specks (converging to a finite set as n goes to ∞). But a kindly Dr. Hausdorff showed that, seen properly, the destiny of the wild ones was as familiar and conventional as that of the others (convergence to a square in the Hausdorff metric).

Our eccentric family is a two-sided sequence $\{\mathbf{F}_{2^n}\}_{n=-\infty}^{\infty}$ of functions from the interval $I = [0, 2\pi)$ to $\mathbb{R}^2$, whose behavior yields a variety of examples suitable for an undergraduate class in real analysis. For each $n \in \mathbb{Z}$ and $t \in I$, we define

$$\mathbf{F}_{2^n}(t) = (x_{2^n}(t), y_{2^n}(t)) = \left(\frac{\tanh(2^n \cos t)}{\tanh(2^n)}, \frac{\tanh(2^n \sin t)}{\tanh(2^n)} \right), \tag{1}$$

where tanh denotes the hyperbolic tangent function. Some examples of these curves are illustrated in Figure 1.

The appearance of the curves ranges from circular as n goes to $-\infty$, to square as n goes to ∞. More precisely, we will show that as n goes to $-\infty$, the sequence $\{\mathbf{F}_{2^n}\}_{n=-\infty}^{\infty}$ converges pointwise to the function $\mathbf{C} : I \to \mathbb{R}^2$ given by

$$\mathbf{C}(t) = (\cos t, \sin t). \tag{2}$$

For large positive values of n, depictions of curves such as $\mathbf{F}_{2^7}$ in Figure 1(b) resemble the square S with vertices $\pm(1, 1)$, $\pm(1, -1)$. However, we will see that as n goes to ∞, the pointwise limit of the sequence $\{\mathbf{F}_{2^n}\}_{n=-\infty}^{\infty}$ is the discontinuous function $\mathbf{D} : I \to \mathbb{R}^2$ given by

$$\mathbf{D}(t) = (\text{sign}(\cos t), \text{sign}(\sin t)), \tag{3}$$

where sign : $\mathbb{R} \to \mathbb{R}$ is the *sign* or *signum* function,

$$\text{sign}(x) = -1 \text{ if } x < 0; \quad 0 \text{ if } x = 0; \quad 1 \text{ if } x > 0.$$

The set $\mathbf{D}(I)$ is the discrete set of eight points labeled in Figure 1(a):

$$\mathbf{D}(I) = \{(1, 0), (1, 1), (0, 1), (-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1)\}.$$

It will be helpful to consider a superset of the sequence $\{\mathbf{F}_{2^n}\}_{n=-\infty}^{\infty}$, namely, $\{\mathbf{F}_\alpha\}_{\alpha>0}$, where, for each $\alpha \in (0, \infty)$, the curve $\mathbf{F}_\alpha : I \to \mathbb{R}^2$ is given by

$$\mathbf{F}_\alpha(t) = (x_\alpha(t), y_\alpha(t)) = \left(\frac{\tanh(\alpha \cos t)}{\tanh(\alpha)}, \frac{\tanh(\alpha \sin t)}{\tanh(\alpha)} \right). \tag{4}$$

Math. Mag. **91** (2018) 267–273. doi:10.1080/0025570X.2018.1495456
MSC: Primary 26-01

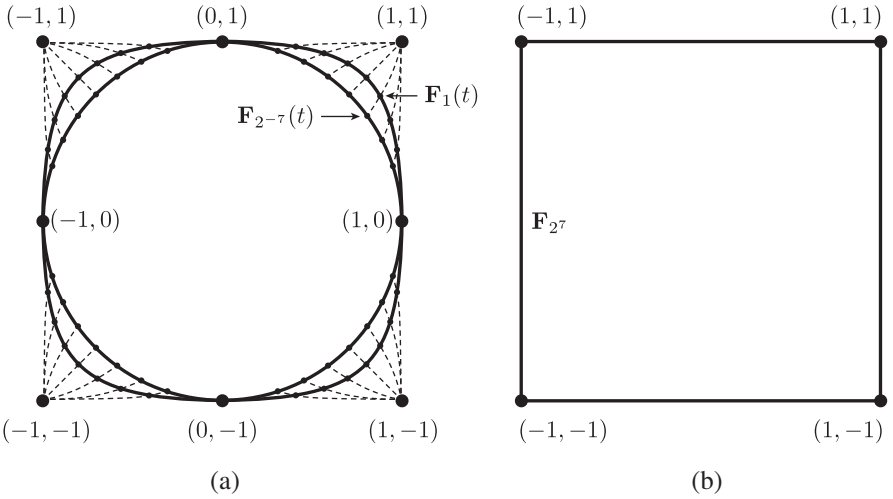


Figure 1 Examples of the curves $\mathbf{F}_{2^n}$ defined in equation (1). Each dashed line in (a) is the path of $\mathbf{F}_\alpha(t)$ for a fixed value of t, as α varies continuously over $(0, \infty)$. In (b) the curve $\mathbf{F}_{2^7}$ does not quite go through the points $\pm(1, 1)$ and $\pm(1, -1)$, the vertices of the square S which $\mathbf{F}_{2^7}$ closely resembles.

Each dashed line in Figure 1(a) is the path of $\mathbf{F}_\alpha(t)$ for a fixed value of t, as α varies continuously over $(0, \infty)$. This shows that in a sense, as n goes to ∞, the points of the curves $\mathbf{F}_{2^n}$ try to crowd into the corners of the square S, except for the four points of $\mathbf{D}(I)$ fixed on the coordinate axes. In the rest of the article, we verify and further explain these claims to show how they can be of value in an introductory real analysis class.

Properties of the curves

We use the term *curve* to denote a continuous function $\mathbf{H} : J \to \mathbb{R}^2$, where J is an interval in $\mathbb{R}$. Following convention, we illustrate such a curve by depicting the set $\mathbf{H}(J)$ in $\mathbb{R}^2$, and labeling it $\mathbf{H}$ (for example, the label $\mathbf{F}_{2^7}$ in Figure 1(b), as opposed to $\mathbf{F}_{2^7}(I)$). However, from a set-theoretic point of view, the function $\mathbf{H}$ is technically equivalent to the set $\{(t, \mathbf{H}(t)) : t \in J\}$, also called the *graph* of $\mathbf{H}$. This subset of $\mathbb{R} \times \mathbb{R}^2$ is impractical to depict in two dimensions, so one typically "graphs" a curve by sketching something other than its graph!

This technicality is typically (and wisely) downplayed when introducing the topic of parametric curves, but in more advanced classes it can be used to segue into other topics. In particular, the conventional depiction of $\mathbf{F}_{2^7}$ in Figure 1(b)—which again is actually $\mathbf{F}_{2^7}(I)$—so closely resembles the square S that it motivates the topic of the Hausdorff metric as a way of measuring the apparent convergence of smooth curves to a square. We explain this in detail later.

Pointwise convergence As n goes to $-\infty$, $\mathbf{F}_{2^n}$ converges pointwise to $\mathbf{C}$, as defined in equation (2). To verify this, consider $\mathbf{F}_\alpha(t)$ in equation (4), with t fixed and $\alpha \in (0, \infty)$ being a continuous real variable. In the limit as $\alpha \to 0$, L'Hôpital's rule applies to each component of $\mathbf{F}_\alpha(t)$ in equation (4), hence

$$\lim_{\alpha \to 0} \mathbf{F}_\alpha(t) = \lim_{\alpha \to 0} \left(\frac{\operatorname{sech}^2(\alpha \cos t) \cos t}{\operatorname{sech}^2(\alpha)}, \frac{\operatorname{sech}^2(\alpha \sin t) \sin t}{\operatorname{sech}^2(\alpha)} \right) = (\cos t, \sin t).$$

It follows that $\lim_{n \to -\infty} \mathbf{F}_{2^n}(t) = \mathbf{C}(t)$ for each $t \in I$.

As n goes to ∞, $\mathbf{F}_{2^n}$ converges pointwise to $\mathbf{D}$, where $\mathbf{D}$ is given by equation (3). To verify this, we use the fact that for a real constant c, $\lim_{x\to\infty} \tanh(cx) = \operatorname{sign}(c)$. It follows from equation (4) that

$$\lim_{\alpha\to\infty} \mathbf{F}_\alpha(t) = (\operatorname{sign}(\cos t), \operatorname{sign}(\sin t)),$$

and hence $\lim_{n\to\infty} \mathbf{F}_{2^n}(t) = \mathbf{D}(t)$ for each $t \in I$.

Uniform convergence As n goes to $-\infty$, $\{\mathbf{F}_{2^n}\}_{n=-\infty}^{\infty}$ converges uniformly to $\mathbf{C}$. To prove this, we use the following theorem by Dini (see for example [1, p. 60] or [2, p. 425]).

Theorem (Dini)**.** *Let X be a compact subset of $\mathbb{R}$. If a monotone sequence of continuous functions from X to $\mathbb{R}$ converges pointwise to a continuous function, then it also converges uniformly.*

To use Dini's theorem, we extend the domain of each function $\mathbf{F}_{2^n}$ to the compact set $[0, 2\pi]$ in the natural way, and show that the corresponding sequences $\{x_{2^n}\}_{n=-\infty}^{\infty}$ and $\{y_{2^n}\}_{n=-\infty}^{\infty}$ in (1) are monotone. Applying the identity

$$\tanh(2u) = \frac{2\tanh(u)}{1+\tanh^2(u)}$$

to the numerator and denominator of $x_{2^{n+1}}(t)$, we obtain

$$x_{2^{n+1}}(t) = \frac{\tanh(2^{n+1}\cos t)}{\tanh(2^{n+1})} = \frac{2\tanh(2^n\cos t)}{1+\tanh^2(2^n\cos t)} \cdot \frac{1+\tanh^2(2^n)}{2\tanh(2^n)}$$

$$= x_{2^n}(t) \cdot \frac{1+\tanh^2(2^n)}{1+\tanh^2(2^n\cos t)} \geq x_{2^n}(t),$$

the final inequality following from the fact that $\tanh^2(u)$ decreases as $|u|$ decreases. We can use this result and the easily verified identity $y_{2^n}(t) = x_{2^n}(t - \pi/2)$ to obtain

$$y_{2^{n+1}}(t) = x_{2^{n+1}}(t - \pi/2) \geq x_{2^n}(t - \pi/2) = y_{2^n}(t).$$

Since $\{x_{2^n}\}_{n=-\infty}^{\infty}$ and $\{y_{2^n}\}_{n=-\infty}^{\infty}$ are therefore monotone sequences, Dini's theorem implies that the convergence $\lim_{n\to-\infty} \mathbf{F}_{2^n} = \mathbf{C}$ is uniform.

Remark. The plots $(t, x_{2^n}(t))$ and $(t, y_{2^n}(t))$ (see Figure 2) are interesting in their own right. As n goes to $-\infty$ they approximate cosine or sine waves, and as n goes to ∞ they are continuous approximations of square waves.

Non-uniform convergence The *uniform limit theorem* (see for example [6, p. 130]), says that the uniform limit of a sequence of continuous functions is continuous. Since $\mathbf{D}$ is obviously discontinuous on $[0, 2\pi)$, the pointwise convergence $\lim_{n\to\infty} \mathbf{F}_{2^n} = \mathbf{D}$ is not uniform. In an intuitive sense, there is something visually jarring about this non-uniform convergence, as shown in Figure 1. At no stage do the sets $\mathbf{F}_{2^n}(I)$ resemble the eight-point set $\mathbf{D}(I)$. Rather, for increasingly large values of n, as shown by $\mathbf{F}_{2^7}$ in Figure 1(b), the sets $\mathbf{F}_{2^n}(I)$ increasingly resemble the square S with vertices $\pm(1, 1)$, $\pm(1, -1)$. A mathematical way to capture this "convergence of appearance" is provided by the *Hausdorff metric*.

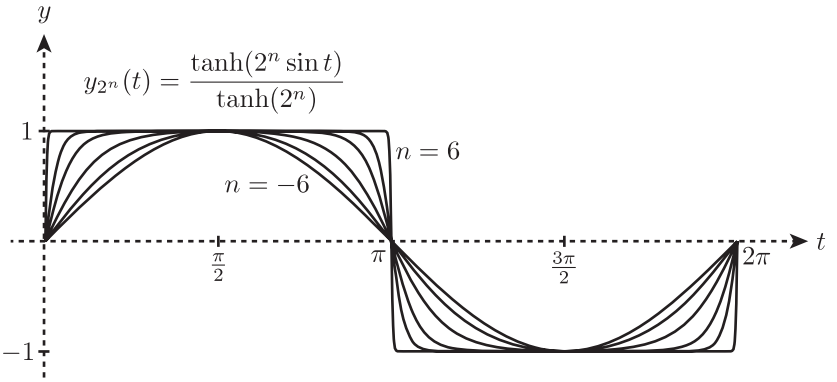


Figure 2 The curves $(t, y_{2^n}(t))$ for six values of n.

The Hausdorff metric

Let $\mathbb{H}$ denote the set of all nonempty, compact subsets of $\mathbb{R}^2$, and define the function $d_{\mathrm{H}} : \mathbb{H} \times \mathbb{H} \to [0, \infty)$ for any $A, B \in \mathbb{H}$ by

$$d_{\mathrm{H}}(A, B) = \max\left\{\max_{\mathbf{a}\in A} \operatorname{dist}(\mathbf{a}, B), \max_{\mathbf{b}\in B} \operatorname{dist}(\mathbf{b}, A)\right\},$$

where $\operatorname{dist}(\mathbf{x}, Y) = \min_{\mathbf{y}\in Y} \|\mathbf{x} - \mathbf{y}\|$ for $\mathbf{x} \in \mathbb{R}^2$ and $Y \in \mathbb{H}$. Intuitively, $\operatorname{dist} : \mathbb{R}^2 \times \mathbb{H} \to [0, \infty)$ measures the distance from a given point to the nearest point(s) in a given compact set. As illustrated in Figure 3, the *Hausdorff distance* $d_{\mathrm{H}}(A, B)$, for $A, B \in \mathbb{H}$, searches A for the farthest point $\mathbf{a}$ from B, searches B for the farthest point $\mathbf{b}$ from A, and outputs the larger of the two distances as $d_{\mathrm{H}}(A, B)$. For $\varepsilon > 0$, we have $d_{\mathrm{H}}(A, B) < \varepsilon$ if and only if every point of A is within ε units of some point of B, and vice versa. It can be shown that $(\mathbb{H}, d_{\mathrm{H}})$ is a metric space.

We note that since $\mathbf{F}_{2^n}(I) = \mathbf{F}_{2^n}([0, 2\pi)) = \mathbf{F}_{2^n}([0, 2\pi])$, each set $\mathbf{F}_{2^n}(I)$ is compact, effectively being the continuous image of a compact set. Therefore $\mathbf{F}_{2^n}(I) \in \mathbb{H}$, and likewise, $\mathbf{C}(I), S \in \mathbb{H}$. We will show that $\lim_{n\to\infty} d_{\mathrm{H}}(\mathbf{F}_{2^n}(I), S) = 0$, and leave the convergence $\lim_{n\to-\infty} d_{\mathrm{H}}(\mathbf{F}_{2^n}(I), \mathbf{C}(I)) = 0$ as an exercise.

Figure 4 shows two squares centered at the origin. One is the set S, and the other is the square S_n with vertices $\pm(a_n, a_n), \pm(a_n, -a_n)$ where, for a given $n \in \mathbb{N}$,

$$a_n = \frac{\tanh(2^n/\sqrt{2})}{\tanh(2^n)}. \tag{5}$$

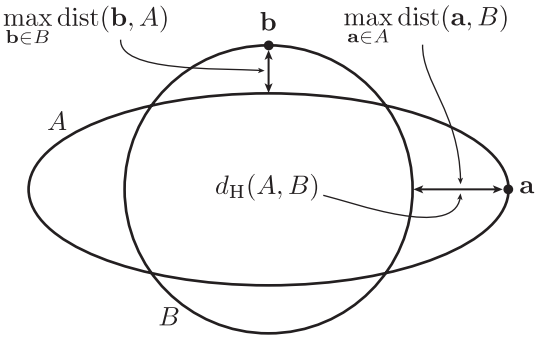


Figure 3 Hausdorff distance between an ellipse A and a circle B.

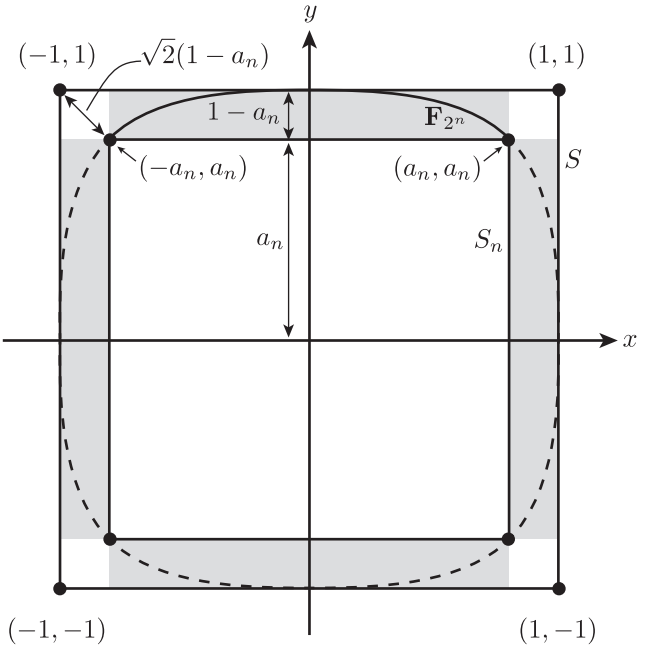


Figure 4 Sketch of the proof that the sets $\mathbf{F}_{2^n}(I)$ converge to the square S in the Hausdorff metric.

The solid curve above S_n is the restriction $\mathbf{F}_{2^n}|_{[\pi/4,3\pi/4]}$. It is straightforward to check that the endpoints of $\mathbf{F}_{2^n}|_{[\pi/4,3\pi/4]}$ are $\mathbf{F}_{2^n}(\pi/4) = (a_n, a_n)$ and $\mathbf{F}_{2^n}(3\pi/4) = (-a_n, a_n)$. It can similarly be verified that for $t \in (\pi/4, 3\pi/4)$, we have

$$-a_n < x_{2^n}(t) < a_n \quad \text{and} \quad a_n < y_{2^n}(t) \le 1.$$

It follows that

$$\mathbf{F}_{2^n}([\pi/4, 3\pi/4]) \subset [-a_n, a_n] \times [a_n, 1]. \tag{6}$$

That is, the solid arc $\mathbf{F}_{2^n}([\pi/4, 3\pi/4])$ in Figure 4 belongs to the shaded gray rectangle above S_n.

By a similar analysis, we can show that the other parts of $\mathbf{F}_{2^n}(I)$ are also contained in the gray rectangles in Figure 4. Specifically, we have

$$\begin{aligned}
\mathbf{F}_{2^n}([3\pi/4, 5\pi/4]) &\subset [-1, -a_n] \times [-a_n, a_n],\\
\mathbf{F}_{2^n}([5\pi/4, 7\pi/4]) &\subset [-a_n, a_n] \times [-1, -a_n],\\
\mathbf{F}_{2^n}([-\pi/4, \pi/4]) &\subset [a_n, 1] \times [-a_n, a_n],
\end{aligned}$$

where in the last formula we have extended the domain of $\mathbf{F}_{2^n}$ to include the interval $[-\pi/4, \pi/4]$. Since $\mathbf{F}_{2^n}(I)$ and $\mathbf{F}_{2^n}([-\pi/4, 7\pi/4])$ are equal as sets, this shows that for each $n \in \mathbb{N}$, the set $\mathbf{F}_{2^n}(I)$ is circuminscribed between the squares S and S_n. Referring to Figure 4, this means that

$$\begin{aligned}
d_{\mathrm{H}}(\mathbf{F}_{2^n}(I), S) &= \max\left\{ \max_{\mathbf{x}\in\mathbf{F}_{2^n}(I)} \operatorname{dist}(\mathbf{x}, S), \max_{\mathbf{x}\in S} \operatorname{dist}(\mathbf{x}, \mathbf{F}_{2^n}(I)) \right\}\\
&\le \max\{(1-a_n), \sqrt{2}(1-a_n)\}\\
&\le \sqrt{2}(1-a_n).
\end{aligned}$$

Using the formulation of a_n in equation (5), we then have

$$0 \le \lim_{n\to\infty} d_{\mathrm{H}}(\mathbf{F}_{2^n}(I), S) \le \lim_{n\to\infty} \sqrt{2}\left(1 - \frac{\tanh(2^n/\sqrt{2})}{\tanh(2^n)}\right) = 0.$$

We therefore conclude that the sets $\mathbf{F}_{2^n}(I)$ converge to the square S in the Hausdorff metric.

Conclusion

Although the family $\{\mathbf{F}_{2^n}\}_{n=-\infty}^{\infty}$ may not need saving in the sense of rescuing, we feel that it is worth saving as an example for the classroom. This two-sided sequence contains lessons in pointwise convergence, uniform convergence, Dini's theorem, the uniform limit theorem, and the Hausdorff metric, as well as practice with parametric curves and the hyperbolic functions. We should note that analogous practice with the inverse trigonometric functions is afforded by a similar sequence $\{\mathbf{G}_{2^n}\}_{n=-\infty}^{\infty}$, where

$$\mathbf{G}_{2^n} = \left(\frac{\arctan(2^n \cos t)}{\arctan(2^n)}, \frac{\arctan(2^n \sin t)}{\arctan(2^n)}\right).$$

Having the range space be $\mathbb{R}^2$ allows for some helpful graphical illustrations, and in particular it shows the rather surprising convergence of a sequence of smooth curves to a finite set of points, underscoring the sometimes counterintuitive aspects of pointwise convergence.

The Hausdorff metric comes to the rescue of our strained intuitions, reassuring us that $\mathbf{F}_{2^n}(I) \to S$ as n goes to ∞. As we mentioned earlier, for $\varepsilon > 0$, we have $d_{\mathrm{H}}(A, B) < \varepsilon$, for $A, B \in \mathbb{H}$, if and only if every point of A is within ε units of some point of B, and vice versa. Thus, if ε is sufficiently small, it is tempting to conclude that graphical representations of A and B would appear identical. Indeed, the Hausdorff distance is used as a measure of visual resemblance in areas such as fractal image compression, pattern analysis, and face recognition (see, e.g., [3–5]). However, like all mathematical concepts, the Hausdorff metric has its nonintuitive aspects. We illustrate this with a final example.

Each of the compact sets $N, D \in \mathbb{H}$ in Figure 5 consists of a word in black letters on a rectangular field of narrow, black, vertical rectangles (stripes) of width ε, separated by gaps also of width ε. The points of each set belong to the black regions. Each rectangular field has the same corners p, q, r, s. When ε is much smaller than shown, the words remain legible against the background (try viewing at a distance), because the white gaps are as plentiful as the black stripes, thus giving a gray tone to the background even when individual stripes cannot be resolved. When one set is laid over the other, we have $d_{\mathrm{H}}(N, D) < \varepsilon$ no matter how small ε is, but despite this, the appearances of N and D remain as different as . . . well, you get the idea.

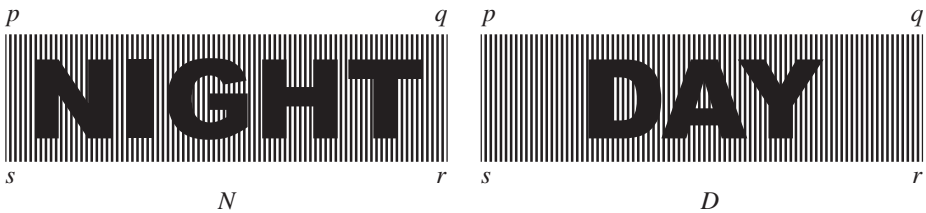


Figure 5 Two sets close in Hausdorff distance, but visually as different as night and day.

REFERENCES

[1] Aliprantis, C. D., Burkinshaw, O. (1981). *Principles of Real Analysis.* New York: North Holland.
[2] Apostol, T. M. (1957). *Mathematical Analysis: A Modern Approach to Advanced Calculus.* Reading, MA: Addison-Wesley,
[3] Barnsley, M. F., Hurd, L. P. (1992). *Fractal Image Compression.* Wellesley, MA: AK Peters, Ltd.
[4] Huttenlocher, D. P., Klanderman, G. A., Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance, *IEEE Trans. Pattern Anal. Mach. Intell.* 15: 850–863.
[5] Jesorsky, O., Kirchberg, K. J., Frischholz, R. (2001). Robust face detection using the Hausdorff distance. In: Bigun, J., Smeraldi, F., eds. *Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA '01).* London: Springer-Verlag, pp. 90–95.
[6] Munkres, J. R. (1975). *Topology: A First Course.* Englewood Cliffs, NJ: Prentice-Hall.

Summary. We investigate a two-sided sequence of plane curves whose surprising behavior motivates and illustrates concepts from real analysis. Topics include pointwise versus uniform convergence, Dini's theorem, the uniform limit theorem, the Hausdorff metric, and practice with parametric curves and the hyperbolic functions.

MARC FRANTZ (MR Author ID: 317801) received a B.F.A. in painting from the Herron School of Art in 1975, followed by a 13-year career with art galleries. He earned an M.S. in mathematics from IUPUI in 1990 and is currently a research associate in mathematics at Indiana University. He loves the visual approach to mathematics, especially links between mathematics and art.

Cauchy, Gershgorin, and Matrix Polynomials

AARON MELMAN
Santa Clara University
Santa Clara, CA 95053
amelman@scu.edu

In an observation from 1829, Cauchy found that all the zeros of the polynomial with complex coefficients $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ lie in the disk defined by $|z| \le s$, where s is the unique positive solution of

$$x^n - |a_{n-1}|x^{n-1} - \cdots - |a_1|x - |a_0| = 0.$$

As an example, consider the polynomial $p(z) = z^5 + (1+i)z^2 - z + 3$, whose largest zero has modulus 1.4653. The unique positive solution of $x^5 - \sqrt{2}x^2 - x - 3 = 0$ is obtained for $x = 1.5044$, which is clearly an upper bound on the moduli of p's zeros.

The (straightforward) proof of this result is usually obtained by direct manipulation of the polynomial ([3], [10, Theorem 27.1, p. 122 and Exercise 1, p. 126]), but an elegant alternative proof was presented in [1] by means of Gershgorin's theorem, a localization result for matrix eigenvalues, that can also be used to derive related results ([11], [13, Section 8.6]).

Polynomials with complex coefficients are *scalar* polynomials, and Cauchy's observation was recently generalized in [2], [8], and [12] to *matrix* polynomials (polynomials whose coefficients are complex matrices) with methods reminiscent of those traditionally used in the scalar case. Our purpose is to show that Gershgorin's theorem can be used here as well, just as in the scalar case, to prove important special cases of the generalized result, and even improve some of them, unlike in the scalar case.

Preliminaries

Matrix polynomials appear in generalized eigenvalue problems where a nonzero complex vector v and a complex number z are sought such that $P(z)v = 0$, with

$$P(z) = A_n z^n + A_{n-1}z^{n-1} + \cdots + A_0, \tag{1}$$

and the coefficients A_j are $m \times m$ complex matrices. We will assume throughout that P is *regular*, namely that $\det P(z)$, which is a polynomial in z of degree at most nm, is not identically zero. One generally prefers to write $P(z)$ (which is, after all, a matrix) in the polynomial form of equation (1) because some properties of its eigenvalues, bounds especially, can be expressed in terms of its (matrix) coefficients.

If A_n is singular, then there are infinite eigenvalues and if A_0 is singular then zero is an eigenvalue. There are nm eigenvalues, including possibly infinite ones. The finite eigenvalues are the solutions of $\det P(z) = 0$. The familiar (linear) eigenvalue problem is obtained as a special case when $n = 1$ and $A_n = I$.

A function $\|\cdot\| : \mathbb{C}^{n\times n} \to \mathbb{R}$ is a *matrix norm* if, for all $A, B \in \mathbb{C}^{n\times n}$, it satisfies the following properties ([9, Section 5.6]):

1. $\|A\| \ge 0$
2. $\|A\| = 0$ if and only if $A = 0$

Math. Mag. **91** (2018) 274–285. doi:10.1080/0025570X.2018.1491767
MSC: Primary 12D10, Secondary 15A18; 30C15

3. $\|cA\| = |c|\|A\|$ for all $c \in \mathbb{C}$
4. $\|A + B\| \leq \|A\| + \|B\|$
5. $\|AB\| \leq \|A\|\|B\|$.

Common norms are the 1-norm, infinity norm, and the 2-norm (or *spectral* norm), defined ([9, pp. 294–295]) for $A \in \mathbb{C}^{n\times n}$ with elements a_{ij} by

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| = \|A^{\mathrm{T}}\|_\infty,$$

$$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| = \|A^{\mathrm{T}}\|_1,$$

$$\|A\|_2 = \max\left\{\sqrt{\mu} \ : \ \mu \text{ is an eigenvalue of } A^*A\right\} = \|A^*\|_2,$$

where * indicates the conjugate transpose. The 1-norm and the infinity norm are very easy to compute, but the 2-norm is costly for large matrices.

We are now ready for the generalized Cauchy result, which states that all the eigenvalues of the regular matrix polynomial P in equation (1), when A_n is the identity matrix, lie in the disk defined by $|z| \leq s$, where s is the unique positive solution of

$$x^n - \|A_{n-1}\|x^{n-1} - \cdots - \|A_1\|x - \|A_0\| = 0,$$

and where the norm (same norm for all matrices) can be any matrix norm. This is similar in form to the classical result for scalar polynomials, with norms replacing absolute values.

We will derive this generalized result for the three above-mentioned matrix norms using two tools: the block companion matrix of P defined in equation (1) and Gershgorin's theorem. A companion matrix of a *scalar* polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ is a matrix whose eigenvalues are the zeros of p, i.e., p is its characteristic polynomial. Such a matrix is given by ([9, 3.3.12])

$$C(p) = \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix}.$$

Blank spaces in matrices indicate zero elements. Likewise, the eigenvalues of a monic matrix polynomial P, i.e., with leading coefficient $A_n = I$, are the eigenvalues of its *$nm \times nm$ block companion matrix* $C(P)$, defined by ([7, (2.4)])

$$C(P) = \begin{pmatrix} 0 & & & -A_0 \\ I & & & -A_1 \\ & \ddots & & \vdots \\ & & I & -A_{n-1} \end{pmatrix}.$$

Since the size of I will usually be clear from the context, we omit it from the notation.

The *Gershgorin column set* of a matrix ([6], [9, Section 6.1]) is a union of disks, centered at the diagonal elements of the matrix, that contains all its eigenvalues. The radii of the disks are the deleted column sums of those elements, namely the sum of the absolute values of all off-diagonal elements in a column. Denoting by $\mathcal{D}(a; r)$ the

closed disk centered at $a \in \mathbb{C}$ with radius $r > 0$, this means that the eigenvalues of a complex $n \times n$ matrix A with elements a_{ij} are contained in the union

$$\bigcup_{j=1}^{n} \mathcal{D}\left(a_{jj}, K_j^{'}\right),$$

where $K_j^{'} = \sum_{i=1, i\neq j}^{n} |a_{ij}|$ is the jth deleted column sum. We illustrate this theorem with the following example.

Example 1. Define the matrix

$$A = \begin{pmatrix} 2i & 0 & 2 \\ 1 & -2+i & 0 \\ -2 & 1+i & 2 \end{pmatrix},$$

for which $K_1^{'} = 3$, $K_2^{'} = \sqrt{2}$, and $K_3^{'} = 2$. Then, by Gershgorin's theorem, all eigenvalues of A lie in the union

$$\mathcal{D}(2i; 3) \cup \mathcal{D}\left(-2+i; \sqrt{2}\right) \cup \mathcal{D}(2; 2).$$

The Gershgorin disks are shown in Figure 1. The eigenvalues are indicated by bullets.

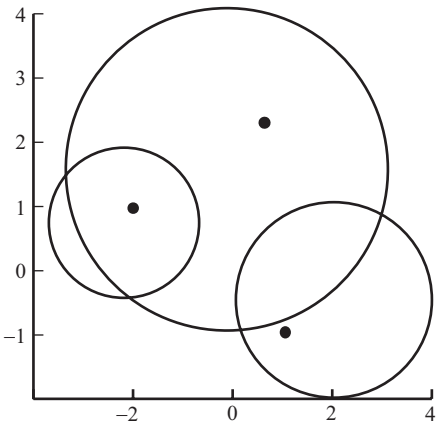


Figure 1 Gershgorin disks for Example 1.

For a historical background of this interesting theorem as well as for similar eigenvalue inclusion regions, we refer to [**14**]. Because the spectra of A and A^T are the same, an analogous result is obtained by replacing the deleted column sums by the deleted row sums; the resulting set is referred to as the *Gershgorin row set*. Moreover, for any nonsingular matrix S, the eigenvalues of A and $S^{-1}AS$ are the same, although their Gershgorin sets are, in general, not. A judiciously chosen S can, therefore, reduce the size of that set, which is precisely our intention.

1-norm and ∞-norm

Our first aim is to obtain the generalized Cauchy result for the 1-norm. For that, we let Δ_x be the block diagonal matrix with diagonal blocks $(x^n I, x^{n-1} I, \ldots, xI)$, where I is the $m \times m$ identity matrix and $x > 0$, and we define

$$C_x(P) = \Delta_x^{-1} C(P) \Delta_x = \begin{pmatrix} 0 & & & -A_0/x^{n-1} \\ xI & & & -A_1/x^{n-2} \\ & \ddots & & \vdots \\ & & xI & -A_{n-1} \end{pmatrix} \tag{2}$$

for a matrix polynomial P as in equation (1) with leading coefficient the identity matrix. We note that $C_x(P)$ has the same eigenvalues as $C(P)$ for any nonzero value of x, although, as we will see, its Gershgorin sets are, in general, different for different values of x.

The Gershgorin column set of $C_x(P)$ is the union of one disk centered at the origin with radius x and m disks centered at the diagonal elements of $(-A_{n-1})$. Of the latter, the jth disk $(1 \le j \le m)$ has a radius given by

$$\sum_{\substack{i=1 \\ i \neq j}}^{m} \left|(A_{n-1})_{ij}\right| + \sum_{k=0}^{n-2} \frac{\sum_{i=1}^{m} \left|(A_k)_{ij}\right|}{x^{n-k-1}}. \tag{3}$$

As an example, consider the cubic matrix polynomial $Q(z)$ given by

$$Q(z) = Iz^3 + A_2 z^2 + A_1 z + A_0 = Iz^3 + \begin{pmatrix} 3 & 1 \\ 0 & -i \end{pmatrix} z^2 + \begin{pmatrix} -i & -1 \\ 0 & -2 \end{pmatrix} z + \begin{pmatrix} -1 & -2 \\ 1 & -1 \end{pmatrix}.$$

Its eigenvalues are all finite and they are equal to those values of z for which

$$\det Q(z) = \det \begin{pmatrix} z^3 + 3z^2 - iz - 1 & z^2 - z - 2 \\ 1 & z^3 - iz^2 - 2z - 1 \end{pmatrix} = 0.$$

For this matrix polynomial Q, we have

$$C(Q) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & i & 1 \\ 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & -3 & -1 \\ 0 & 0 & 0 & 1 & 0 & i \end{pmatrix} \quad \text{and} \quad \Delta_x = \begin{pmatrix} x^3 & 0 & 0 & 0 & 0 & 0 \\ 0 & x^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & x^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & x^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & x & 0 \\ 0 & 0 & 0 & 0 & 0 & x \end{pmatrix},$$

so that

$$C_x(Q) = \Delta_x^{-1} C(Q) \Delta_x = \begin{pmatrix} 0 & 0 & 0 & 0 & 1/x^2 & 2/x^2 \\ 0 & 0 & 0 & 0 & -1/x^2 & 1/x^2 \\ x & 0 & 0 & 0 & i/x & 1/x \\ 0 & x & 0 & 0 & 0 & 2/x \\ 0 & 0 & x & 0 & -3 & -1 \\ 0 & 0 & 0 & x & 0 & i \end{pmatrix}.$$

The Gershgorin column set of $C_x(Q)$ is the union of three disks: one disk centered at the origin with radius x (which is, in fact, the union of four identical disks) and two more disks, one centered at -3 with radius $1/x + 2/x^2$, and one centered at i

with radius $1 + 3/x + 3/x^2$. It is important to remember that the eigenvalues of Q are contained in this set for *any* value of $x > 0$.

The idea now is to vary the parameter x to obtain a single disk containing all the eigenvalues: if we let x increase, then the disk centered at the origin expands monotonically, while the two remaining disks contract monotonically. Consider the first of these disks, centered at -3: as its radius decreases with increasing x, this disk becomes enclosed by (and tangent to) the disk centered at the origin when $x = 3 + 1/x + 2/x^2$, or $x = 3.46$. Reasoning similarly for the disk centered at i, we obtain that it will become enclosed by the disk centered at the origin when $x = 1 + 1 + 3/x + 3/x^2$, or $x = 3.22$. The disk with the larger radius of 3.46 (the first disk) must therefore contain all eigenvalues of Q since $\mathcal{D}(i; 3.22) \subseteq \mathcal{D}(i; 3.46)$. This is illustrated in Figure 2, which shows these three disks with, from left to right, $x = 2.85, 3.22, 3.46$. The value $x = 2.85$ is arbitrarily chosen to be less than the other two values. The eigenvalues are indicated by the bullets. As x increases from 2.85 to 3.22, the disk centered at the origin expands while the other two contract. When $x = 3.22$, the disk centered at the origin encompasses and is tangent to the disk centered at i. As x further increases to 3.46, the same thing happens with the disk centered at -3, while the one centered at i has contracted even more and becomes completely absorbed by the disk centered at the origin, which must now contain all the eigenvalues.

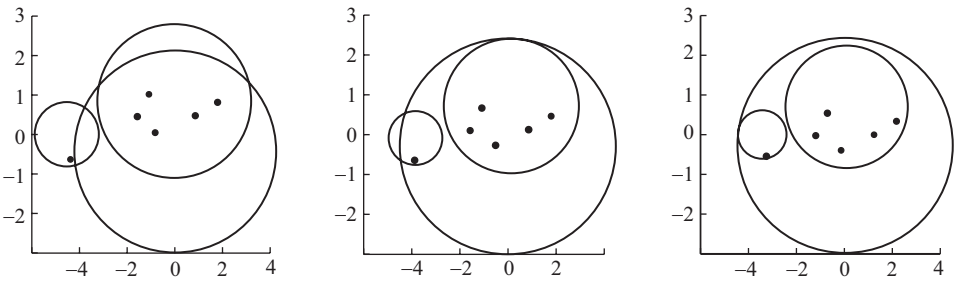


Figure 2 Gershgorin disks for Q with, from left to right, $x = 2.85, 3.22, 3.46$.

We will now carry out this procedure for the general case to show how it leads to Cauchy's result, so let us go back to $C_x(P)$ in equation (2) and the radii of its Gershgorin disks from equation (3). As in the example, if x increases, then the radius of the disk centered at the origin increases monotonically, while the radii of the m remaining disks decrease monotonically. Consider the jth disk centered at $(-A_{n-1})_{jj}$: as its radius decreases with increasing x, this disk becomes enclosed by (and tangent to) the disk centered at the origin with radius x when

$$x = \left|(A_{n-1})_{jj}\right| + \sum_{\substack{i=1 \\ i \neq j}}^{m} \left|(A_{n-1})_{ij}\right| + \sum_{k=0}^{n-2} \frac{\sum_{i=1}^{m} \left|(A_k)_{ij}\right|}{x^{n-k-1}},$$

which is equivalent to

$$x^n - \left(\sum_{i=1}^{m} \left|(A_{n-1})_{ij}\right| \right) x^{n-1} - \sum_{k=0}^{n-2} \left(\sum_{i=1}^{m} \left|(A_k)_{ij}\right| \right) x^k = 0.$$

If we denote the unique positive solution of this equation by ρ_j, then, by Gershgorin's theorem, all the eigenvalues are contained in a disk centered at the origin with radius

$\rho = \max_{1\le j\le m}\{\rho_j\}$. This proves the generalized Cauchy result for the 1-norm because, for *any* j, we have that

$$x^n - \|A_{n-1}\|_1\, x^{n-1} - \sum_{k=0}^{n-2} \|A_k\|_1\, x^k \le x^n - \sum_{i=1}^{m} \left|(A_{n-1})_{ij}\right| x^{n-1} - \sum_{k=0}^{n-2} \left(\sum_{i=1}^{m} \left|(A_k)_{ij}\right| \right) x^k,$$

i.e., for $x > 0$, the polynomial on the left of the inequality lies below that on the right, while both have a unique positive zero and are nonpositive at the origin. This means that the polynomial on the right vanishes for a smaller value of x than that on the left, i.e., $\rho_j \le s$, where $x = s$ is the solution of

$$x^n - \|A_{n-1}\|_1\, x^{n-1} - \sum_{k=0}^{n-2} \|A_k\|_1\, x^k = 0.$$

That in turn means that $\rho \le s$ and thus that any eigenvalue z satisfies $|z| \le \rho \le s$.

We have derived, in fact, an improvement of Cauchy's result for matrix polynomials, since $\rho \le s$. It is also worth mentioning again that different Gershgorin sets are obtained for different values of the parameter x and that the eigenvalues are contained in the intersection of all of them. The following example further illustrates what we just explained.

Example 2. Consider the quadratic matrix polynomial

$$Q(z) = Iz^2 + A_1 z + A_0 = Iz^2 + \begin{pmatrix} -i & -1 \\ 0 & -2 \end{pmatrix} z + \begin{pmatrix} -6 & -1 \\ 0 & -1 \end{pmatrix}.$$

Its (block) companion matrix is

$$C(Q) = \begin{pmatrix} 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & i & 1 \\ 0 & 1 & 0 & 2 \end{pmatrix},$$

so that $\|A_0\|_1 = 6$ and $\|A_1\|_1 = 3$. For this example, Cauchy's result for the 1-norm delivers $s = 4.37$ as an upper bound on the magnitudes of the eigenvalues of Q, namely the positive solution of $x^2 - 3x - 6 = 0$, while $\rho_1 = 3$ and $\rho_2 = 3.56$ are the positive solutions of $x^2 - x - 6 = 0$ and $x^2 - 3x - 2 = 0$, respectively. Clearly, $\rho = \max\{\rho_1, \rho_2\} = \rho_2 < s$, i.e., we have found a better upper bound than the generalized Cauchy bound. With equation (3), we find that the Gershgorin column set for $x = \rho_1$ is given by $\Gamma_1 = \mathcal{D}(0; 3) \cup \mathcal{D}(i; 2) \cup \mathcal{D}(2; 1.67)$, while it is given by the single disk $\Gamma_2 = \mathcal{D}(0; 3.56) = \mathcal{D}(0; 3.56) \cup \mathcal{D}(i; 1.68) \cup \mathcal{D}(2; 1.56)$ when $x = \rho_2$. The sets Γ_1 and Γ_2 are illustrated on the left and right in Figure 3, respectively. The dashed circle on the left is the disk $\mathcal{D}(0; \rho)$, which is included here for reference. It shows that Γ_1 has, in fact, a smaller area than $\mathcal{D}(0; \rho) = \Gamma_2$, which itself lies in the disk $\mathcal{D}(0; s)$ (not shown in the figure). The bullets are the eigenvalues of Q.

Of course, one is not limited to the particular values that we chose for the parameter x, which only served our purpose of proving Cauchy's result. Just as an example, one could choose a value for which the radius of the disk centered at the origin is the arithmetic mean of the radii of the disks centered at the diagonal elements of A_0,

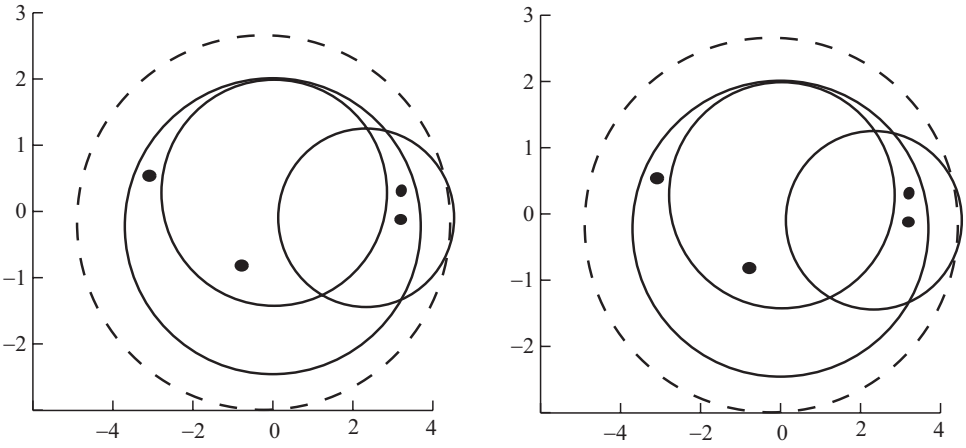


Figure 3 Gershgorin sets for the eigenvalues of Q.

making that disk's radius similar to those of the other disks. In the case of Example 2, e.g., this means that

$$x = \frac{1}{2}\left(\frac{6}{x} + 1 + \frac{2}{x}\right),$$

so that $x^2 - \frac{1}{2}x - 4 = 0$, which gives $x = 2.27$. In this case, the Geshgorin column set is given by $\mathcal{D}(0; 2.27) \cup \mathcal{D}(i; 2.65) \cup \mathcal{D}(2; 1.88)$, which can be seen in Figure 4, where the dashed circle is the same as in Figure 3. It produces an inclusion region with an area that is actually smaller than that of $\mathcal{D}(0; \rho)$.

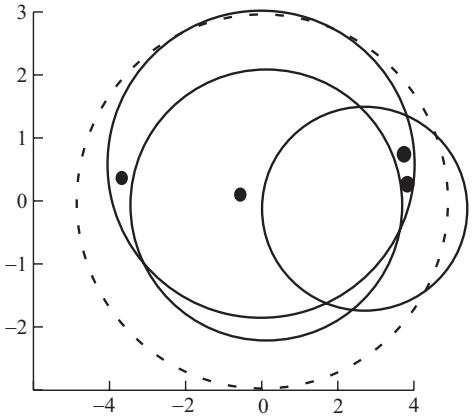


Figure 4 Gershgorin set for the eigenvalues of Q with $x = 2.27$.

Once it is established for the 1-norm, the generalized Cauchy result for the infinity norm follows immediately from the observation that

$$\begin{aligned}\det\left(I\lambda^n + A_{n-1}\lambda^{n-1} + \cdots + A_0\right) &= \det\left(I\lambda^n + A_{n-1}\lambda^{n-1} + \cdots + A_0\right)^{\mathrm{T}} \\ &= \det\left(I\lambda^n + A_{n-1}^{\mathrm{T}}\lambda^{n-1} + \cdots + A_0^{\mathrm{T}}\right),\end{aligned}$$

and from the fact that, for any square complex matrix A, $\|A^{\mathrm{T}}\|_1 = \|A\|_\infty$.

The examples given so far were mainly meant to explain and clarify. The following example is an illustration of the improved bounds for a real-world engineering problem.

Example 3. In [5], a structural dynamics model representing a reinforced concrete machine foundation is formulated as a sparse quadratic eigenvalue problem with $n = 2$ and $m = 3627$. The largest eigenvalue has magnitude 2.12×10^4.

This example was also used in [8] to compare a large number of different bounds on polynomial eigenvalues. Among them, generalized Cauchy bounds were the best by a significant margin: 3.53×10^4 and 3.17×10^4 for the 1-norm and ∞-norm, respectively (other norms would be too computationally costly for matrices of this size). Our refined version delivers even better results: 3.32×10^4 and 2.96×10^4 for the 1-norm and ∞-norm, respectively.

2-norm

Let us now turn to the 2-norm. Here, we require A_{n-1} to be a normal matrix, namely a matrix for which $A^*A = AA^*$. If its eigenvalues are μ_j, then the eigenvalues of A^*A are $|\mu_j|^2$ and $\|A\|_2 = \max_j |\mu_j|$.

We will make use of a generalized version of Gershgorin's theorem ([4]) for matrices that have been partitioned into blocks. It is similar to the classical theorem and its statement is equally straightforward: consider a complex $n \times n$ matrix A, partitioned in the following manner:

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,N} \\ A_{2,1} & A_{2,2} & \dots & A_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N,1} & A_{N,2} & \dots & A_{N,N} \end{pmatrix},$$

where the diagonal submatrices $A_{i,i}$ are square of order n_i, $1 \le i \le N$ and $\sum_{i=1}^{N} n_i = n$, and let I_i be the $n_i \times n_i$ identity matrix. Then each eigenvalue of A lies in a Gershgorin set G_j for at least one j, $1 \le j \le N$, where G_j is the set of all complex numbers z such that

$$\left\| \left(A_{j,j} - zI_i\right)^{-1} \right\|_2^{-1} \le \sum_{\substack{k=1 \\ k \ne j}}^{N} \|A_{k,j}\|_2. \tag{4}$$

When z is an eigenvalue of $A_{j,j}$, the quantity $\left\| \left(A_{j,j} - zI_i\right)^{-1} \right\|_2^{-1}$ is set equal to zero. The set G_j is, in general, complicated, but it turns out to be a simple union of disks when $A_{j,j}$ is a normal matrix and the norm is the 2-norm. This is so because normal matrices can be diagonalized by a unitary matrix, i.e., for any normal matrix A, there exists a matrix U with $U^*U = I = UU^*$, such that

$$U^*AU = \begin{pmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n \end{pmatrix},$$

where $\mu_1, \mu_2, \dots, \mu_n$ are the eigenvalues of A. In addition, we have that $\|A\|_2 = \|UAU^*\|_2$.

As a consequence,

$$\begin{aligned}
\left\|(A-zI)^{-1}\right\|_2^{-1} &= \left\|U\,(A-zI)^{-1}\,U^*\right\|_2^{-1} \\
&= \left\|\left(U^*\,(A-zI)\,U\right)^{-1}\right\|_2^{-1} \\
&= \left\|\begin{pmatrix} (\mu_1-z)^{-1} & & & \\ & (\mu_2-z)^{-1} & & \\ & & \ddots & \\ & & & (\mu_n-z)^{-1} \end{pmatrix}\right\|_2^{-1} \\
&= \min_{1\le k\le n} |\mu_k - z|.
\end{aligned}$$

Therefore, in this case, the set G_j is a union of n_j disks, centered at the eigenvalues of $\left(-A_{j,j}\right)$ and all have the same radius, given by the right-hand side of equation (4).

As for the 1-norm, we now apply this generalized Gershgorin theorem to the companion matrix of a matrix polynomial. The matrix $C_x(P)$, which was defined in equation (2), is by its very nature already partitioned into $m \times m$ blocks. Its first $n-1$ diagonal blocks are all equal to the $m \times m$ zero matrix, while the last diagonal block is $(-A_{n-1})$. If A_{n-1} is normal and the 2-norm is used, then the generalized Gershgorin set that contains all the eigenvalues of P is the union of one disk centered at the origin with radius x and m disks, centered at the eigenvalues of $(-A_{n-1})$ with radius $\sum_{j=0}^{n-2} \|A_j\|_2 / x^{n-1-j}$.

We now proceed similarly as for the 1-norm: as x increases, the disk centered at the origin with radius x increases monotonically in size, while all the others decrease monotonically. If we denote by μ_j $(j = 1, ..., m)$ the eigenvalues of $(-A_{n-1})$, then the disk centered at the origin is tangent to and encloses the disk centered at μ_j when

$$x = |\mu_j| + \sum_{j=0}^{n-2} \frac{\|A_j\|_2}{x^{n-1-j}},$$

or

$$x^n - |\mu_j| x^{n-1} - \|A_{n-2}\|_2\, x^{n-2} - ... - \|A_1\|_2\, x - \|A_0\|_2 = 0. \tag{5}$$

Denote the unique positive solution of equation (5) by ρ_j and set $\rho = \max_{1\le j\le m} \rho_j$. Then, since the m disks centered at the eigenvalues of $(-A_{n-1})$ all have the same radius that decreases with increasing x, the disk $\mathcal{D}(0;\rho)$ necessarily encloses all of them. Consequently, all the eigenvalues of P are contained in $\mathcal{D}(0;\rho)$.

We now note that $|\mu_i| \ge |\mu_j|$ implies that $|\rho_i| \ge |\rho_j|$. Therefore, ρ is the solution of

$$x^n - \left(\max_{1\le j\le m} |\mu_j|\right) x^{n-1} - \|A_{n-2}\|_2\, x^{n-2} - \cdots - \|A_1\|_2\, x - \|A_0\|_2 = 0. \tag{6}$$

Since A_{n-1} is normal, we have $\max_{1\le j\le m} |\mu_j| = \|A_{n-1}\|_2$, so that equation (6) can be written as

$$x^n - \|A_{n-1}\|_2\, x^{n-1} - \|A_{n-2}\|_2\, x^{n-2} - \cdots - \|A_1\|_2\, x - \|A_0\|_2 = 0. \tag{7}$$

We just showed that the unique positive solution of equation (7) is an upper bound on the eigenvalues of P, and this is precisely the generalized Cauchy result for the 2-norm when the second highest matrix coefficient of P is normal. We note that, unlike for the

1-norm and infinity norm, we do not obtain an improvement. On the one hand, we had to impose a restriction on A_{n-1} to prove this result, but on the other hand our approach allows for greater flexibility in the choice of eigenvalue inclusion regions (as for the 1-norm and infinity norm) that we illustrate with the following example.

Example 4. Consider a matrix polynomial T, which differs from Q in Example 2 only in A_1, which is diagonal here (and therefore normal):

$$T(z) = Iz^2 + \begin{pmatrix} -i & 0 \\ 0 & -2 \end{pmatrix} z + \begin{pmatrix} -6 & -1 \\ 0 & -1 \end{pmatrix}.$$

Its (block) companion matrix is

$$C(T) = \begin{pmatrix} 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & i & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}.$$

We have $\|A_0\|_2 = 6.085$, $\|A_1\|_2 = 2$, $\rho_1 = 3.02$ (the positive solution of $x^2 - x - 6.085 = 0$), and $\rho_2 = 3.66$ (the positive solution of $x^2 - 2x - 6.085 = 0$). Cauchy's generalized result for the 2-norm therefore delivers $\rho = \rho_2 = \max\{\rho_1, \rho_2\}$ as the upper bound on the magnitude of the eigenvalues of T. The Gershgorin column set for $x = \rho_1$ is given by $\mathcal{D}(0; 3.02) \cup \mathcal{D}(i; 2.02) \cup \mathcal{D}(2; 2.02)$, while for $x = \rho_2$ it is given by the single disk $\mathcal{D}(0; 3.66) = \mathcal{D}(0; 3.66) \cup \mathcal{D}(i; 1.66) \cup \mathcal{D}(2; 1.66)$. These two sets can be seen on the left and right in Figure 5, respectively. The dashed circle on the left represents the disk $\mathcal{D}(0; \rho)$ and the bullets are the eigenvalues of T. Note that the Gershgorin set on the left in Figure 5 is clearly smaller in area than $\mathcal{D}(0; \rho)$.

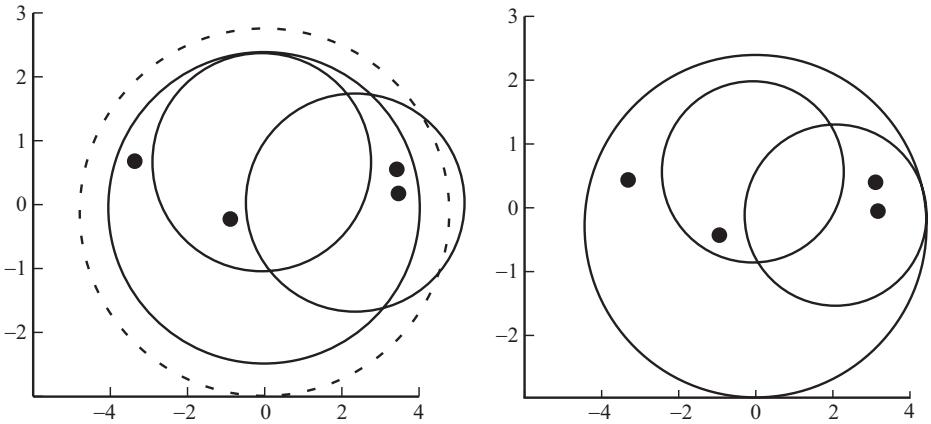


Figure 5 Gershgorin set for the eigenvalues of T.

As an additional example of the flexibility introduced by the parameter x, one could also choose a value for it so that all disks have the same radius, which, in this case, requires

$$x = \frac{\|A_0\|_2}{x},$$

or $x = \sqrt{\|A_0\|_2} = 2.47$. The resulting Gershgorin column set, illustrated in Figure 6, is smaller in area than $\mathcal{D}(0; \rho)$. As in Figure 5, the dashed circle is the disk $\mathcal{D}(0; \rho)$ and the bullets are the eigenvalues of T.

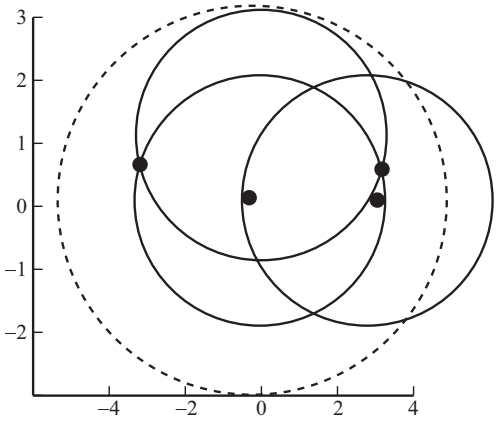


Figure 6 Gershgorin set for the eigenvalues of T with $x = 2.47$.

Conclusion

We have used Gershgorin's theorem to prove the generalization to matrix polynomials of a result by Cauchy for several common matrix norms. We also saw that, by varying a parameter, Gershgorin's theorem allows for many additional eigenvalue inclusion regions. We explored some of those, but by no means all. The interested reader can obtain even more inclusion regions by using a different diagonal matrix Δ_x and by using more general sets such as can be found in [**14**].

Another topic worth investigating is the generalization to matrix polynomials of results for scalar polynomials, other than Cauchy's. Obviously, not every scalar result can be generalized, but there is no doubt that some can. A good source for scalar results, some of which do not appear to be widely known, is [**13**].

Acknowledgment We thank Françoise Tisseur for providing the matrix polynomial in Example 3.

REFERENCES

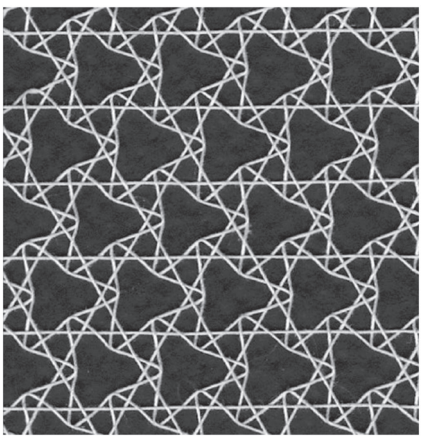
[1] Bell, H. M. (1965). Gershgorin's theorem and the zeros of polynomials. *Amer. Math. Monthly* 72: 292–295.
[2] Bini, D. A., Noferini, V., Sharify, M. (2013). Locating the eigenvalues of matrix polynomials. *SIAM J. Matrix Anal. Appl.* 34: 1709–1727.
[3] Cauchy, A. L. (1829). Sur la résolution des équations numériques et sur la théorie de l'élimination. *Exercices de Mathématiques, Quatrième Année*. Paris. Also in: Oeuvres Complètes, Série 2, Tome 9, pp. 86–161. Gauthiers-Villars et fils, Paris, 1891.
[4] Feingold, D. G., Varga, R. S. (1962). Block diagonally dominant matrices and generalizations of the Gerschgorin circle theorem. *Pacific J. Math.* 12: 1241–1250.
[5] Feriani, A., Perotti, F., Simoncini, V. (2000). Iterative system solvers for the frequency analysis of linear mechanical systems. *Comput. Methods Appl. Mech, Eng.* 19: 1719–1739.
[6] Gershgorin, S. (1931). Über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk SSSR, Ser. Fiz.-Mat.* 6: 749–754.
[7] Gohberg, I. Lancaster, P. Rodman, L. (1982). *Matrix Polynomials* New York: Academic Press.
[8] Higham, N. J., Tisseur, F. (2003). Bounds for eigenvalues of matrix polynomials. *Linear Algebra Appl.* 358:5–2.
[9] Horn, R. A., Johnson, C. R. (1988). *Matrix Analysis.* Cambridge: Cambridge University Press.
[10] Marden, M. (1966). *Geometry of polynomials.* Mathematical Surveys. No. 3. Providence, RI: American Mathematical Society.
[11] Melman, A. (2013). The twin of a theorem by Cauchy. *Amer. Math. Monthly* 120: 164–168.
[12] Melman, A. (2013). Generalization and variations of Pellet's theorem for matrix polynomials. *Linear Algebra Appl.* 439: 1550–1567.

[13] Rahman, Q. I., Schmeisser, G. (2002). *Analytic Theory of Polynomials.* London Mathematical Society Monographs. New Series, 26. Oxford: The Clarendon Press, Oxford University Press.
[14] Varga, R. S. (2004). *Geršgorin and his Circles.* Springer Series in Computational Mathematics, Vol. 36. Berlin: Springer-Verlag.

Summary. A well-known result by Cauchy bounding the moduli of (scalar) polynomial zeros can be derived by using a localization result for matrix eigenvalues, namely Gershgorin's theorem. Cauchy's result was recently generalized from scalar to matrix polynomials and we will show how Gershgorin's theorem can be used here as well to prove important special cases of the generalized result. Although it does not prove all cases, it does—unlike for the scalar result—improve some of those that it does.

AARON MELMAN (MR Author ID: 293268) received his M.Sc. from the Technion—Israel Institute of Technology in 1986 and Ph.D. from Caltech in 1992. He taught at Ben-Gurion University in Israel and is currently teaching at Santa Clara University in California. His research interests are in numerical analysis and linear algebra.

Artist Spotlight: Veronika Irvine



Triaxial arrowheads, Veronika Irvine; White cotton thread, 2015. This pattern has 3*3 orbifold symmetry. Three-fold rotational symmetry is extremely rare in traditional bobbin lace designs.

See interview on page 307–309.

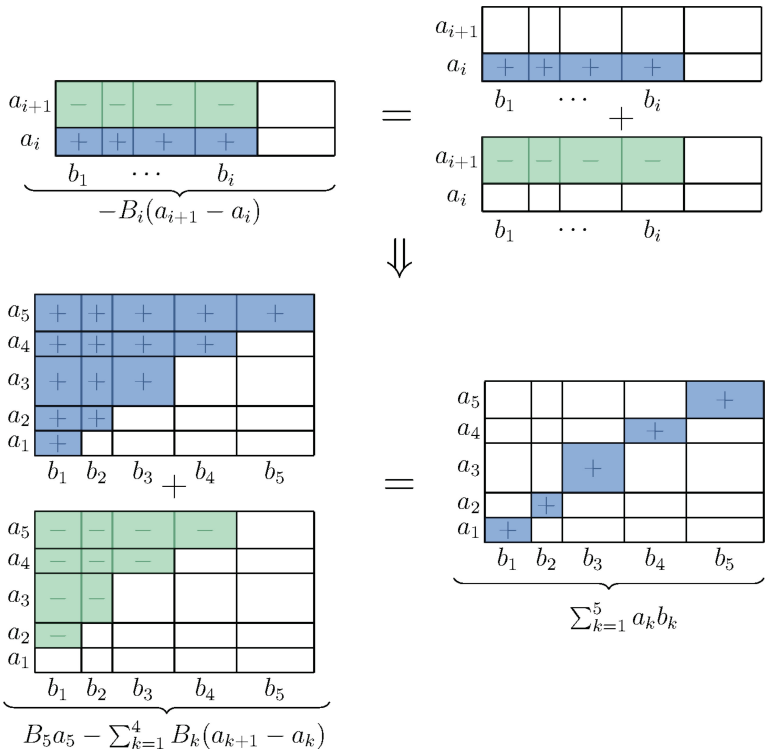
Proof Without Words: Abel's Transformation

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

Theorem. *Let* $(a_1, a_2, a_3, \ldots)$ *and* $(b_1, b_2, b_3, \ldots)$ *be sequences of positive real numbers and let* $n \in \mathbb{N}$. *If* $B_i := b_1 + b_2 + \cdots + b_i$, *then*

$$\sum_{k=1}^{n} a_k b_k = B_n a_n - \sum_{k=1}^{n-1} B_k (a_{k+1} - a_k).$$

Proof. (e.g. for $n = 5$)



■

Exercise. Use $a = (1, 2, 3, 4, \ldots)$ and $a' = (1, 4, 9, 16, \ldots)$ separately in conjunction with $b = (1, 1, 1, 1, \ldots)$ to show that, for all n,

$$\sum_{k=1}^{n} k = \frac{n^2 + n}{2} \qquad \text{and} \qquad \sum_{k=1}^{n} k^2 = \frac{2n^3 + 3n^2 + n}{6}.$$

Math. Mag. **91** (2018) 286–287. doi:10.1080/0025570X.2018.1467175
MSC: Primary 40G10
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

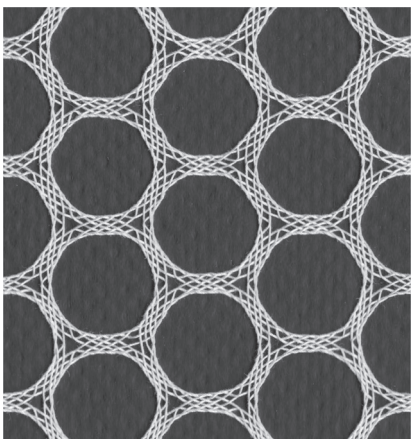
Exercise. Use Abel's transformation to provide closed formulas for $\sum_{k=1}^{n} k^3$ and $\sum_{k=1}^{n} k^4$.

Summary. We provide a wordless computation demonstrating the discrete analog of integration by parts, which is often referred to as summation by parts or Abel's transformation. We mention a few consequences.

YAJUN AN (MR Author ID: 1212929) is a Visiting Assistant Professor of Mathematics and is currently interested in conditionally convergent infinite series and their rearrangements.

TOM EDGAR (MR Author ID: 821633) is an Associate Professor of Mathematics and first became interested in Abel's transformation due to its relationship to Möbius inversion on posets.

Artist Spotlight: Veronika Irvine



Bee prepared, Veronika Irvine; White cotton thread, 2015. This pattern has *632 orbifold symmetry. The honeycomb structure is formed on a hexagonal lattice.

See interview on page 307–309.

Generalized Fermat–Torricelli Problem: An Algorithm

T. R. MUKUNDAN
Mysore 570 022, India
trmukundan@gmail.com

Imagine that we have to build a road network to connect three towns at the corners of a triangle. Since the budget is limited, we are asked to build the most economical network possible. It is reasonable to assume that the cost of building a road is proportional to its length. One obvious way would be to lay the roads along the sides of the triangle, but is this the most economical solution? Is there a way of connecting the towns such that the total length of the roads is minimum? In essence, this is the issue addressed by the classical Fermat–Torricelli problem. The objective of the classical Fermat-point problem is to find a point within a triangle such that the sum of its distances from the three vertices is a minimum. This seemingly simple problem has a long history. It was first proposed by Fermat (1601–1665) to Torricelli (1608–1647) who solved the problem using geometrical construction [1]. The method involves drawing the circumcircles of equilateral triangles constructed on each of three sides of the triangle and lying outside it. The common point of intersection of the circles is the location of the required minimum. Subsequent to this solution over three centuries ago, many new solutions have been proposed, mostly based on some form of geometrical construction and reasoning. Looking back, it is surprising that the problem had not been solved using standard analytical methods. It was only recently that Hajja [2] gave an analytical solution using advanced vector calculus.

Let us now slightly modify the above example and ask a different question. In the above example, we assumed that the cost of building the road is proportional to the distance. Implicit is the assumption that the costs are the same in all the segments. In practice, it is quite probable that, because of the differing terrain along the three segments, the costs are different. In this case, the total cost is the weighted sum of the distances, where the weighting factors are the cost per unit length of the corresponding segments.

The aim of the generalized Fermat–Torricelli problem, also known as the Weber problem, is to find a point within a triangle such that the sum of the weighted distances from the point to the vertices is a minimum. This problem is important in operations research as it has practical applications like optimal facility location. Tellier [3] has approached the problem using the analogy of equilibrium of forces and the geometry of the triangle, but the reasoning is somewhat involved. With the availability of computing power, this problem has also been solved by numerical techniques. Recently, Uteshev [4] has given an analytical solution to this problem.

The algorithm presented in this article also gives an analytical approach of solving the generalized problem using basic calculus and geometry. The solution proposed is believed to be much simpler and, unlike earlier analytical methods, can be easily understood even by undergraduates. Obviously, this method can also be used for the classical Fermat–Torricelli problem that becomes a particular case with equal weighting factors.

Math. Mag. **91** (2018) 288–293. doi:10.1080/0025570X.2018.1488504
MSC: Primary 51M04, Secondary 51M16
Supplemental data for this article can be accessed on the publisher's website.

Proposed method

Referring to Figure 1, let ABC be the triangle under consideration, with the coordinates of the vertices marked as shown. To solve the generalized problem, we have to find a point $P(x, y)$ in the interior of the triangle such that the sum of weighted distances from P to the vertices is a minimum. The conditions to be satisfied for the minimum to fall within the triangle are stated in [4] and reproduced in Appendix 1 from the supplementary material.

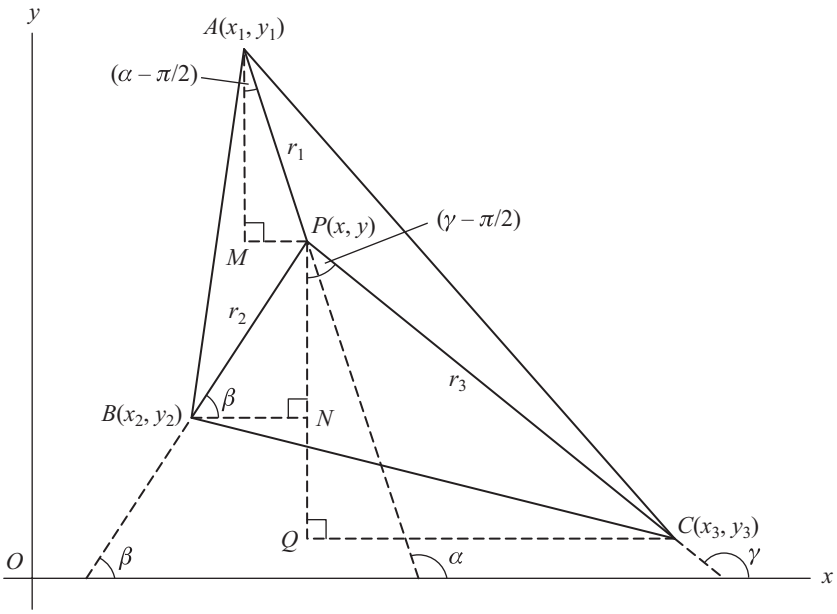


Figure 1 The triangle ABC and the point $P(x, y)$.

Stated formally, we need to minimize the following function

$$S(x, y) = \sum_{i=1}^{3} m_i \sqrt{(x - x_i)^2 + (y - y_i)^2},$$

where m_i, for $i = 1, 2$, and 3, are the weights for the distances from P to A, B, and C, respectively. With respect to Figure 1, it follows that $S = m_1 r_1 + m_2 r_2 + m_3 r_3$, where $r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ for all i. Because the r_i are functions of two variables x and y, the following partial derivative equations have to be satisfied for the sum S to be a minimum:

$$\frac{\partial S}{\partial x} = \sum_{i=1}^{3} m_i \frac{\partial r_i}{\partial x} = 0 \quad \text{and} \quad \frac{\partial S}{\partial y} = \sum_{i=1}^{3} m_i \frac{\partial r_i}{\partial y} = 0.$$

Taking the derivatives and simplifying yields

$$\frac{\partial S}{\partial x} = \sum_{i=1}^{3} \frac{m_i (x - x_i)}{r_i} = 0 \quad \text{and} \quad \frac{\partial S}{\partial y} = \sum_{i=1}^{3} \frac{m_i (y - y_i)}{r_i} = 0. \tag{1}$$

Simultaneously solving the equations in (1) algebraically is difficult, and is probably the reason why standard analytical solutions did not appear for a long time. For this reason, we adopt a hybrid approach by interpreting geometrically the terms in equation (1).

To get started, consider the triangle AMP in Figure 1, where PM and AM are drawn parallel to the x- and y-axes, respectively. It follows that $\angle PAM$ is equal to $\alpha - \pi/2$, so that

$$\frac{x - x_1}{r_1} = \sin(\alpha - \pi/2) = -\cos\alpha \quad \text{and} \quad \frac{y - y_1}{r_1} = -\cos(\alpha - \pi/2) = -\sin\alpha. \tag{2}$$

Similarly, from triangles PNB and PQC in Figure 1, we get the following equations:

$$\frac{x - x_2}{r_2} = \cos\beta \quad \text{and} \quad \frac{y - y_2}{r_2} = \sin\beta; \tag{3}$$

and

$$\frac{x - x_3}{r_3} = -\sin(\gamma - \pi/2) = \cos\gamma \quad \text{and} \quad \frac{y - y_3}{r_3} = \cos(\gamma - \pi/2) = \sin\gamma. \tag{4}$$

Substituting the relations in equations (2)–(4) into equation (1) and simplifying yields

$$m_1\cos\alpha - m_2\cos\beta - m_3\cos\gamma = 0 \quad \text{and} \quad m_1\sin\alpha - m_2\sin\beta - m_3\sin\gamma = 0.$$

Consequently, equation (1) is converted to simpler trigonometric equations.

There are three unknown angles but only two equations. Hence, a unique solution for the angles is not possible. However, by transposing, squaring, and adding, then

$$\cos(\alpha - \beta) = \frac{m_1^2 + m_2^2 - m_3^2}{2m_1m_2}, \quad \cos(\gamma - \beta) = \frac{m_1^2 - m_2^2 - m_3^2}{2m_2m_3}, \quad \text{and}$$
$$\cos(\gamma - \alpha) = \frac{m_1^2 - m_2^2 + m_3^2}{2m_1m_3}. \tag{5}$$

Further inspection of Figure 1 also results in the following equations:

$$\angle APB = \theta_{12} = \pi - (\alpha - \beta), \quad \angle BPC = \theta_{23} = \gamma - \beta, \quad \text{and}$$
$$\angle CPA = \theta_{31} = \pi - (\gamma - \alpha).$$

For clarity, angles θ_{12}, θ_{23}, and θ_{31} are shown in Figure 2.

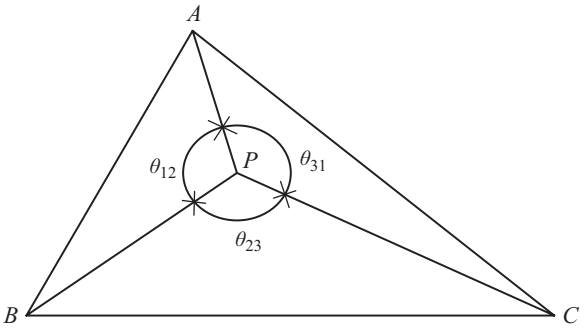


Figure 2 The angles θ_{12}, θ_{23}, and θ_{31}.

Substituting the definitions of the θ_{ij} into equation (5), it follows that

$$\cos\theta_{12} = \frac{m_3^2 - m_1^2 - m_2^2}{2m_1m_2},\ \cos\theta_{23} = \frac{m_1^2 - m_2^2 - m_3^2}{2m_2m_3},\ \text{and}$$
$$\cos\theta_{31} = \frac{m_2^2 - m_3^2 - m_1^2}{2m_3m_1}. \tag{6}$$

These equations determine the angles subtended by the sides AB, BC, and CA of the triangle at the point P where the function $S(x, y)$ attains a minimum. The problem of locating the minimum, therefore, involves solving for triangles PAB, PBC, and PCA with vertical angles θ_{ij} as defined above in equation (6). It is known from basic geometry that the locus of the vertex of a triangle with a base of a given length and vertical angle is a circle. Geometrically, this follows from the theorem that states that "a chord of a circle subtends equal angles at the circumference"; this is proved analytically in Appendix 2 of the supplementary material. Hence, P lies at the intersection of the circumcircles of triangles with AB, BC, and CA as the base and vertical angles θ_{12}, θ_{23}, and θ_{31}, respectively. See Figure 3.

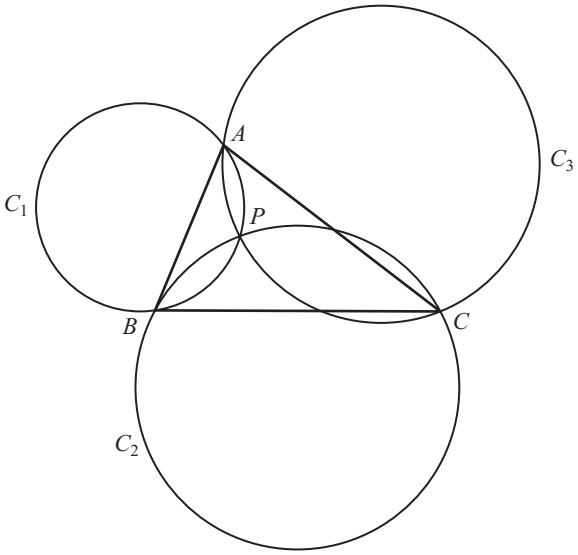


Figure 3 P lies at the intersection of circumcircles.

Indeed, this is the basis of the geometrical construction to locate the minimum point in the classical Fermat–Torricelli problem, which is a particular case of the generalized problem with equal weighting factors. When we substitute $m_1 = m_2 = m_3 = 1$, then

$$\cos\theta_{12} = \cos\theta_{23} = \cos\theta_{31} = -1/2 \text{ so that } \theta_{12} = \theta_{23} = \theta_{31} = 2\pi/3.$$

The procedure of drawing equilateral triangles for locating the minimum in the classical problem is based on this fact. Because all the angles are equal, this point is also known as the isogonal point.

To obtain the coordinates of the generalized Fermat–Torricelli point, we proceed to find the equations of three circles C_1, C_2, and C_3, as in Figure 3. The common point of intersection of the circles is the required minimum. The procedure is based on standard methods of analytical geometry and is given in Appendix 2 of the supplementary material.

The algorithm

The algorithm given here is sequence-sensitive, i.e., it is necessary for the (x_i, y_i) to be in the counter-clockwise direction around the triangle. To ensure this, we carry out the following procedure.

Step 1. Using the Heron formula for the area of a triangle, evaluate the following

$$\Delta = (1/2)\left[x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)\right].$$

If Δ is positive, then the sequence is counter-clockwise and hence acceptable. However, if Δ is negative, then we can interchange the indices of any two vertices (and the corresponding values of the m_i) to make the sequence counter-clockwise and proceed according to Step 2.

Step 2. Using the values of θ_{12}, θ_{23}, and θ_{31} from equation (6), evaluate the following parameters:

$$a_{11} = (x_3 - x_1) + \frac{y_2 - y_3}{\tan\theta_{23}} - \frac{y_1 - y_2}{\tan\theta_{12}}, \qquad a_{21} = (x_1 - x_2) + \frac{y_3 - y_1}{\tan\theta_{31}} - \frac{y_2 - y_3}{\tan\theta_{23}},$$

$$a_{12} = (y_3 - y_1) + \frac{x_3 - x_2}{\tan\theta_{23}} - \frac{x_2 - x_1}{\tan\theta_{12}}, \qquad a_{22} = (y_1 - y_2) + \frac{x_1 - x_3}{\tan\theta_{31}} - \frac{x_3 - x_2}{\tan\theta_{23}},$$

$$b_1 = (x_1x_2 + y_1y_2) - (x_2x_3 + y_2y_3) + \frac{x_2y_3 - x_3y_2}{\tan\theta_{23}} - \frac{x_1y_2 - x_2y_1}{\tan\theta_{12}},$$

$$b_2 = (x_2x_3 + y_2y_3) - (x_3x_1 + y_3y_1) + \frac{x_3y_1 - x_1y_3}{\tan\theta_{31}} - \frac{x_2y_3 - x_3y_2}{\tan\theta_{23}}.$$

From these equations, we can obtain the coordinates (x^*, y^*) of the Fermat-Torricelli point, where

$$x^* = \frac{a_{12}b_2 - a_{22}b_1}{a_{11}a_{22} - a_{21}a_{12}} \quad \text{and} \quad y^* = \frac{a_{21}b_1 - a_{11}b_2}{a_{11}a_{22} - a_{21}a_{12}}.$$

Conclusion

The essential feature of the analytical method described, being based on basic calculus and geometry, is its simplicity. Unlike many previous methods, it does not need any advanced concepts and can be understood by beginners.

Acknowledgments The author wishes to acknowledge with thanks the suggestions received from the referees and the editor in the preparation of the article.

REFERENCES

[1] Courant, R., Robbins, H. (1941). *What is Mathematics?* Oxford: Oxford University Press.
[2] Hajja, M. (1994). An advanced calculus approach to finding the Fermat point. *Math. Mag.* 67: 29–34.
[3] Tellier, L.-N. (1972). The Weber problem: Solution and interpretation. *Geogr. Anal.* 4: 215–233.
[4] Uteshev, A.Y. (2014). Analytical solution for the generalised Fermat–Torricelli problem. *Amer. Math. Monthly* 121: 318–331.

Summary. The classical Fermat–Torricelli problem has a long history spanning over three centuries. It is interesting that, till recently, this problem has been solved only by a variety of geometrical methods. The generalized

version of the problem has practical applications and has received considerable attention through advanced analytical and numerical techniques. This article presents a simplified approach to this problem using only basic calculus and geometry, thus making it understandable to beginners.

T.R. MUKUNDAN (MR Author ID: 1264501) received his master's degree in engineering from the Indian Institute of Technology , Bombay, India. His professional career has been devoted to research and development in the electrical industry. He is interested in the history of mathematics and applications of mathematics to practical problems.

Simultaneous Generalizations of the Theorems of Menelaus, Ceva, Routh, and Klamkin/Liu

STEPHEN SU
Rio American High School
Sacramento, CA 95864
sustephen1@gmail.com

CHENG SHYONG LEE
National Taipei University of Education
Taipei 106, Taiwan
chengshyonglee@gmail.com

Menelaus' and Ceva's theorems provide elegant geometric methods for identifying collinear points and for testing concurrent lines, respectively. However, few are familiar with their varieties of generalizations. We hope that this article draws attention to the subject. In 1891, Routh [21] offered two expressions that are the generalizations of the theorems of Menelaus and Ceva, respectively. (The theorems of Menelaus, Ceva, and Routh have become mainstays in geometry texts, such as in [2–6], and have been the frequent topic for expository mathematics, such as in [7, 10–12, 14, 15], and [16–20].) Recently, in 2013, Bényi and Ćurgus [1] generalized and combined Routh's two theorems in only one expression under the affine transformation of $\mathbb{R}^2$. Previously, in 1992, Klamkin and Liu [13] applied the barycentric coordinate system in the extended Euclidean lane to give simultaneous generalizations of the theorems of Menelaus and Ceva. Later, Houston and Powers [11] in 2009 used a homographic transformation to extend the result of Klamkin and Liu to projective planes. It is also interesting to note that Grünbaum and Shephard [8] in 1995 generalized the theorems of Ceva and Menelaus as well, but with a different approach than Routh and Klamkin and Liu. In this article, for the convenience of the readers, we first restate the theorems of Menelaus, Ceva, Routh, and Klamkin/Liu with adjusted notation, and review some relevant results. We then present our main result of only one expression that includes all the theorems of Menelaus, Ceva, Routh, and Klamkin/Liu as special cases. We prove this theorem by using two standard undergraduate tools: linear algebra and analytical geometry. Finally, we apply our result to offer a solution to a problem recently proposed in 2016 by Su and Lee [22] that gives a very general solution to a problem proposed in 1987 by Herbert Guelicher [9].

Background

In the first century A.D., the ancient Greek astronomer-geometer Menelaus of Alexandria discovered the following elegant result concerning collinear points.

Theorem A (Menelaus' theorem)**.** *Let $A_1A_2A_3$ be a triangle in the Euclidean plane. Let $m_k \neq -1$ be a real number and M_k be a point on the line through A_{k+1} and A_{k+2}, denoted by $L(A_{k+1}A_{k+2})$, such that $A_{k+1}M_k = \frac{1}{1+m_k}A_{k+1}A_{k+2}$, for $k = 1, 2, 3$, as indicated in Figure 1. Then, the three points M_1, M_2, and M_3 are collinear if and only if $m_1m_2m_3 = -1$.*

Math. Mag. **91** (2018) 294–303. doi:10.1080/0025570X.2018.1495435
MSC: Primary 51M04

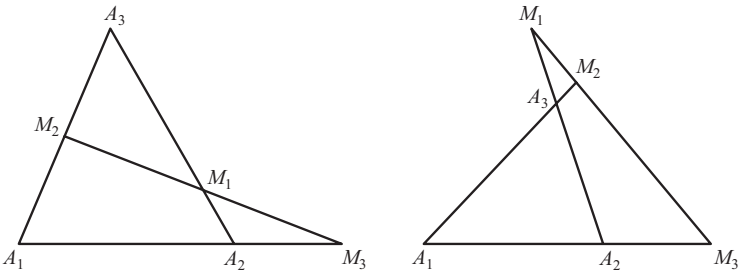


Figure 1 Two possible cases of Menelaus' theorem.

Throughout this article, distance of line segment $\overline{AB}$ is taken to be signed (directed) so that $AB = -BA$. Furthermore, subscripts are taken to be modulo 3, so that A_{k+2} for $k = 2$ is the same as A_1. Finally, triangles $A_1A_2A_3$ are assumed to be nondegenerate so that A_1, A_2, and A_3 are noncollinear.

In 1678, the Italian engineer–mathematician Giovanni Ceva found a laudable result concerning concurrent lines that we reproduce below.

Theorem B (Ceva's theorem)**.** *Let $A_1A_2A_3$ be a triangle in the Euclidean plane. Let $c_k \neq -1$ be a real number and C_k be a point on the line $L(A_{k+1}A_{k+2})$ such that $C_kA_{k+2} = \frac{1}{1+c_k}A_{k+1}A_{k+2}$, for $k = 1, 2, 3$, as indicated in Figure 2. Then, the three lines $L(C_kA_k)$, for $k = 1, 2, 3$, are concurrent if and only if $c_1c_2c_3 = 1$.*

Two centuries later, in 1891, the English mathematician E.J. Routh [**21**] offered, without proof, the following two expressions (Theorems C and D) when evaluating stress–strain and tension in mechanical frameworks. Theorems C and D generalize the theorems of Menelaus and Ceva, respectively.

Theorem C (Routh's first theorem)**.** *Let $A_1A_2A_3$ be a triangle in the Euclidean plane. Let $m_k \neq -1$ be a real number and M_k be a point on the line $L(A_{k+1}A_{k+2})$ such that $A_{k+1}M_k = \frac{1}{1+m_k}A_{k+1}A_{k+2}$, for $k = 1, 2, 3$ (see Figure 3). Then, the ratio between the area $a(M_1M_2M_3)$ of the triangle $M_1M_2M_3$ and the area $a(A_1A_2A_3)$ of the triangle $A_1A_2A_3$ satisfies*

$$\frac{a(M_1M_2M_3)}{a(A_1A_2A_3)} = \frac{1 + m_1m_2m_3}{(1 + m_1)(1 + m_2)(1 + m_3)}.$$

Throughout this article, the area $a(A_1A_2 \ldots A_n)$ of any simple (convex or concave) n-gon $A_1A_2 \ldots A_n$ for $n \geq 3$ is conventionally taken to be positive or negative according to whether the vertices A_k, $1 \leq k \leq n$, in that order, are listed in a counterclockwise order or in a clockwise order. Here, "simple" means that the sides of the n-gon do not cross one another.

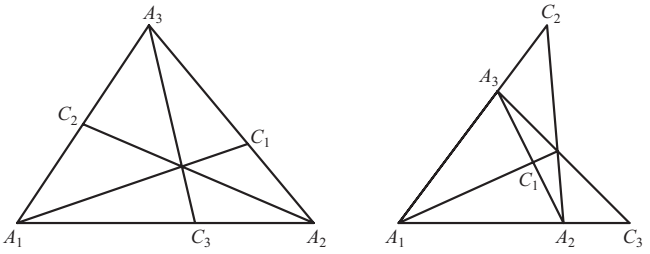


Figure 2 Two possible cases of Ceva's theorem.

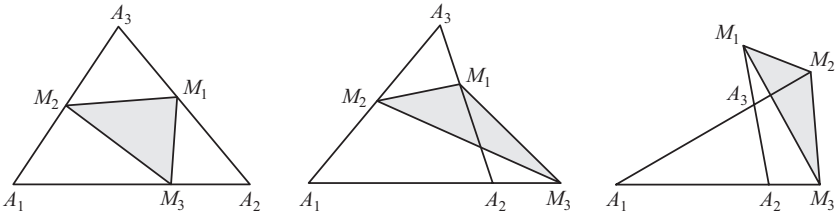


Figure 3 Three possible cases of Routh's first theorem.

To see that Routh's first theorem may be viewed as a generalization of Menelaus' theorem, if $a(M_1M_2M_3) = 0$ in Theorem C, it follows that three points M_1, M_2, and M_3 are collinear if and only if $m_1m_2m_3 = -1$ (the result of Theorem A).

Theorem D (Routh's second theorem)**.** *Let $A_1A_2A_3$ be a triangle in the Euclidean plane. Let $c_k \neq -1$ be a real number and C_k be a point on the line $L(A_{k+1}A_{k+2})$ such that $C_kA_{k+2} = \frac{1}{1+c_k}A_{k+1}A_{k+2}$, for $k = 1, 2, 3$. Suppose that two lines $L(C_kA_k)$ and $L(C_{k+1}A_{k+1})$ intersect at exactly one point R_k, for $k = 1, 2, 3$ (see Figure 4). Then, the ratio between the area of $R_1R_2R_3$ and the area $A_1A_2A_3$ satisfies*

$$\frac{a(R_1R_2R_3)}{a(A_1A_2A_3)} = \frac{(1 - c_1c_2c_3)^2}{(1 + c_1 + c_1c_2)(1 + c_2 + c_2c_3)(1 + c_3 + c_3c_1)}.$$

To see that Routh's second theorem may be viewed as a generalization of Ceva's theorem, if $a(R_1R_2R_3) = 0$ in Theorem D, it follows that the three lines $L(C_kA_k)$, $k = 1, 2, 3$ are concurrent if and only if $c_1c_2c_3 = 1$ (the result of Theorem B). Although the implications of Routh's two theorems look quite different from one another, it is possible to formulate a single statement that contains both of the theorems as special cases. Indeed, in 2013, Bényi and Ćuargus [**1**] constructed and proved such a result using an affine transformation of $\mathbb{R}^2$. Proofs of Theorems C and D are difficult, but there are a number of elegant demonstrations that can be found in the literature (see e.g., [**1**, **4**, **7**, **13**, **16**], and [**19**–**21**]).

Although collinearity and concurrency seem to be related in Menelaus' and Ceva's theorems, it is not immediately obvious how to formulate a single statement containing both theorems as special cases. However, in 1992, Klamkin and Liu [**13**] did just that; their theorem is stated below.

Theorem E (Klamkin and Liu's theorem)**.** *Let $A_1A_2A_3$ be a triangle in the Euclidean plane. Let $c_k \neq -1$ and $m_k \neq -1$ be real numbers, and C_k and M_k be points on the line $L(A_{k+1}A_{k+2})$ such that $A_{k+1}M_k = \frac{1}{1+m_k}A_{k+1}A_{k+2}$ and $C_kA_{k+2} = \frac{1}{1+c_k}A_{k+1}A_{k+2}$, $k = 1, 2, 3$ (see Figure 5). Then, the three lines $L(C_kM_{k+1})$, for $k = 1, 2, 3$, are concurrent*

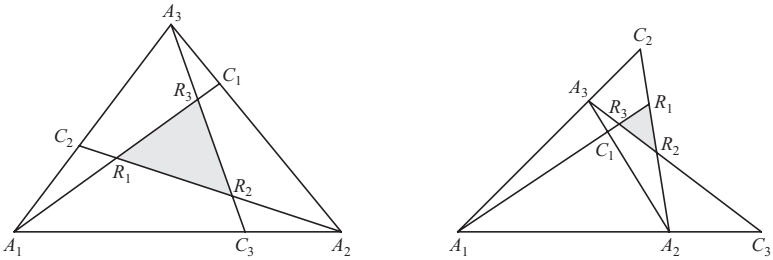


Figure 4 Two possible cases of Routh's second theorem.

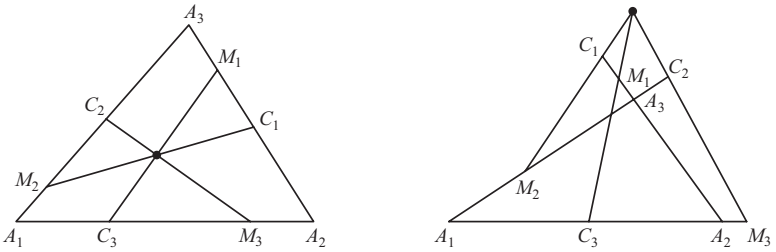


Figure 5 Two possible cases of Klamkin and Liu's theorem.

if and only if

$$c_1c_2c_3 + m_1m_2m_3 + c_1m_1 + c_2m_2 + c_3m_3 = 1.$$

For Klamkin and Liu's theorem, if M_k coincides with A_{k+2}, then $m_k = 0$, for $k = 1, 2, 3$, then Theorem E reduces to $c_1c_2c_3 = 1$, or Ceva's theorem. On the other hand, if M_k coincides with C_k, then $c_km_k = 1$, for $k = 1, 2, 3$, then Theorem E reduces to $m_1m_2m_3 = -1$, or Menelaus' theorem. Furthermore, Klamkin and Liu [**13**] actually proved a somewhat stronger result in which they allowed c_k or m_k to be -1, with the corresponding point C_k or M_k being a "point at infinity" in the extended Euclidean plane. Seventeen years later, in 2009, Houston and Powers [**11**] used the homographic transformation to extend Klamkin and Liu's result to projective planes.

In 1995, Grünbaum and Shephard [**8**] applied the area principle to generalize the theorems of Menelaus and Ceva from triangles to general n-gons, thereby using a different approach than that of Routh [**21**] or Klamkin and Liu [**13**].

Main result

Our main result offers one expression that includes all the theorems of Ceva, Menelaus, Routh and Klamkin, and Liu as special cases.This elaborative generalization is a simple, novel, and powerful illustration of an application of the Cartesian coordinate system.

Theorem 1. *Let $A_1A_2A_3$ be a triangle in the Euclidean plane. Let $c_k \neq -1$ and $m_k \neq -1$ be real numbers and let C_k and M_k be points on the line $L(A_{k+1}A_{k+2})$ such that $C_kA_{k+2} = \frac{1}{1+c_k}A_{k+1}A_{k+2}$and $A_{k+1}M_k = \frac{1}{1+m_k}A_{k+1}A_{k+2}$, for $k = 1, 2, 3$. Suppose that two lines $L(C_kM_{k+1})$ and $L(C_{k+1}M_{k+2})$ intersect at exactly one point R_k, for $k = 1, 2, 3$ (see Figure 6). Then, the ratio of the areas of the triangles $R_1R_2R_3$ and $A_1A_2A_3$ satisfies*

$$\frac{a(R_1R_2R_3)}{a(A_1A_2A_3)} = \frac{N^2}{N_1N_2N_3},$$

where

$$N = c_1c_2c_3 + m_1m_2m_3 + c_1m_1 + c_2m_2 + c_3m_3 - 1 \text{ and}$$

$$N_k = 1 + c_k + c_kc_{k+1} + m_{k+2} + m_{k+2}m_{k+1} - c_{k+1}m_{k+1}, \text{ for } k = 1, 2, 3.$$

For Theorem 1, if M_k coincides with A_{k+2}, then $m_k = 0$, for $k = 1, 2, 3$, and our main result reduces to

$$\frac{a(R_1R_2R_3)}{a(A_1A_2A_3)} = \frac{(1 - c_1c_2c_3)^2}{(1 + c_1 + c_1c_2)(1 + c_2 + c_2c_3)(1 + c_3 + c_3c_1)},$$

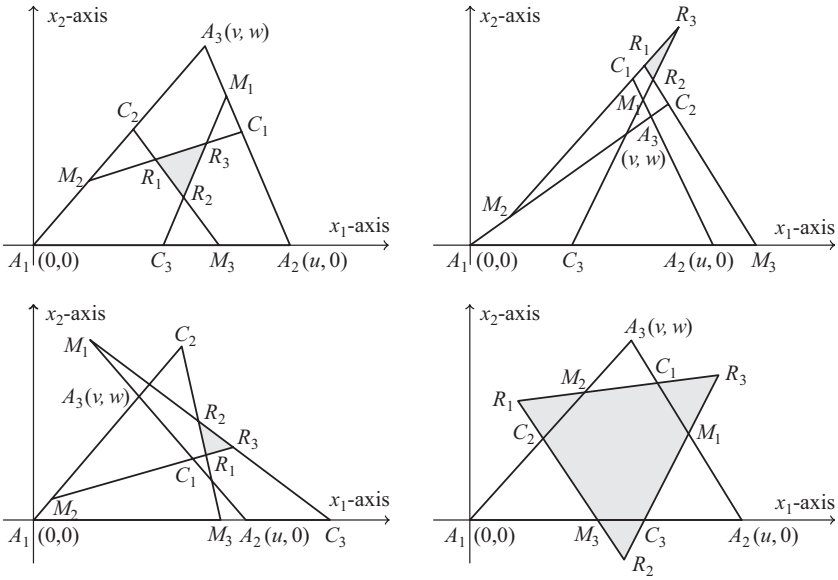


Figure 6 Four possible cases of our main result (Theorem 1).

or Routh's second theorem. On the other hand, if M_k coincides with C_k, then $c_k m_k = 1$ and $M_k = C_k = R_{k+2}$, for $k = 1, 2, 3$. After replacing c_k by m_k and C_k and replacing R_k by M_k, $k = 1, 2, 3$, our main result yields

$$\frac{a(M_1M_2M_3)}{a(A_1A_2A_3)} = \frac{(1 + m_1m_2m_3)^4}{(1 + m_1m_2m_3)^3(1 + m_1)(1 + m_2)(1 + m_3)}. \tag{1}$$

There are two cases to consider.

1. If $1 + m_1m_2m_3 \neq 0$, equation 1 reduces to

$$\frac{a(M_1M_2M_3)}{a(A_1A_2A_3)} = \frac{1 + m_1m_2m_3}{(1 + m_1)(1 + m_2)(1 + m_3)}, \tag{2}$$

   which is Routh's first theorem.

2. If $1 + m_1m_2m_3 = 0$, then we choose $X \in L(A_1A_2)$ and $x \in \mathbb{R}\backslash\{-1\}$ such that $\frac{A_1X}{A_1A_2} = \frac{1}{1+x}$. See Figure 7. As long as $x \neq m_3$ (so that $X \neq M_3$), then $1 + m_1m_2x \neq 0$. From equation 2 and still requiring $x \neq m_3$, we have

$$\frac{a(M_1M_2X)}{a(A_1A_2A_3)} = \frac{1 + m_1m_2x}{(1 + m_1)(1 + m_2)(1 + x)}. \tag{3}$$

   Now, by using limit approximation, we get

$$\begin{aligned}\frac{a(M_1M_2M_3)}{a(A_1A_2A_3)} &= \lim_{X\to M_3} \frac{a(M_1M_2X)}{a(A_1A_2A_3)} = \lim_{x\to m_3} \frac{1 + m_1m_2x}{(1 + m_1)(1 + m_2)(1 + x)} \\ &= \frac{1 + m_1m_2m_3}{(1 + m_1)(1 + m_2)(1 + m_3)}.\end{aligned} \tag{4}$$

   This is also a result of Routh's first theorem even in the case when $1 + m_1m_2m_3 = 0$. In fact, if M_1, M_2, and M_3 are collinear, then the lines $L(M_kM_{k+1})$ and $L(M_{k+1}M_{k+2})$ are the same (since $C_k = M_k$), so they do not intersect in exactly

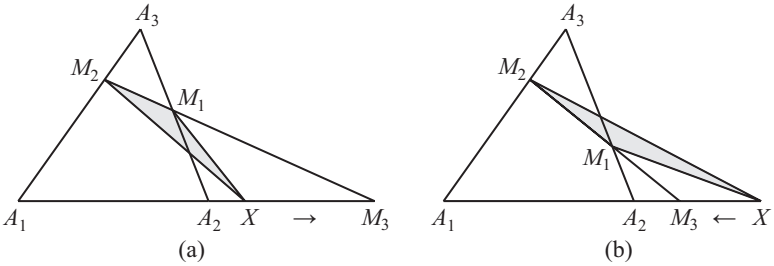


Figure 7 Geometrically representing the limit approximation of equation (4).

one point, and therefore the theorem cannot be applied. To handle this case, one can still obtain Routh's first theorem by using the technique as in the preceding argument.

Moreover, equation (4) provides the information that $1 + m_1m_2m_3 = 0$ if and only if $a(M_1M_2M_3) = 0$ if and only if M_1, M_2, and M_3 are collinear. This is Menelaus' theorem.

Furthermore, by letting $a(R_1R_2R_3) = 0$ in equation (1), three lines $L(C_kM_{k+1})$, for $k = 1, 2, 3$, are concurrent if and only if $c_1c_2c_3 + m_1m_2m_3 + c_1m_1 + c_2m_2 + c_3m_3 = 1$, which is Klamkin and Liu's theorem.

Before proving our main theorem, we introduce the following visual diagram (see Figure 8) that illustrates the hierarchy of the various generalizations of Menelaus' and Ceva's theorems.

Proof of Theorem 1

Without loss of generality, we may set up a Cartesian coordinate system on the Euclidean plane so that vertex A_1 of the given triangle $A_1A_2A_3$ is situated at the origin $(0, 0)$; line $L(A_1A_2)$ is placed along the x_1-axis; the coordinates of A_2 are taken to be $(u, 0)$ with $u \neq 0$; and the coordinates of A_3 are taken to be (v, w) with $w \neq 0$. Figure 6 illustrates four of the possible cases.

Now we use the weighted average formula on a line that states that given two points, $A = (a_1, a_2)$ and $B = (b_1, b_2)$, and two real numbers m and n, then $\forall X(x_1, x_2) \in$

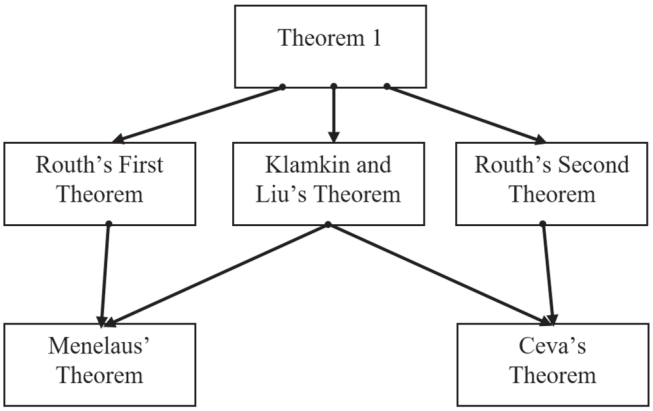


Figure 8 Relationships between Theorem 1, Routh's two theorems, Klamkin and Liu's theorem, and Menelaus' and Ceva's theorems.

$L(AB)$, we have $\frac{AX}{AB} = \frac{m}{m+n}$ if and only if $x_k = \frac{na_k+mb_k}{m+n}$, for $k = 1, 2$. Thus, by the conditions of Theorem 1, we can determine the coordinates of M_1, M_2, and M_3 as follow:

$$M_1 = \left(\frac{m_1u + v}{1+m_1}, \frac{w}{1+m_1}\right), M_2 = \left(\frac{m_2v}{1+m_2}, \frac{m_2w}{1+m_2}\right), \text{ and } M_3 = \left(\frac{u}{1+m_3}, 0\right),$$

and the coordinates of C_1, C_2, and C_3 as follow:

$$C_1 = \left(\frac{u + c_1v}{1+c_1}, \frac{c_1w}{1+c_1}\right), C_2 = \left(\frac{v}{1+c_2}, \frac{w}{1+c_2}\right), \text{ and } C_3 = \left(\frac{c_3u}{1+c_3}, 0\right).$$

Consequently, we can obtain the equations of the three lines $L(C_kM_{k+1})$, for $k = 1, 2, 3$, from the following determinant-form equations:

$$L(C_1M_2): \begin{vmatrix} x_1 & x_2 & 1 \\ \frac{u+c_1v}{1+c_1} & \frac{c_1w}{1+c_1} & 1 \\ \frac{m_2v}{1+m_2} & \frac{m_2w}{1+m_2} & 1 \end{vmatrix} = 0,$$

$$L(C_2M_3): \begin{vmatrix} x_1 & x_2 & 1 \\ \frac{v}{1+c_2} & \frac{w}{1+c_2} & 1 \\ \frac{u}{1+m_3} & 0 & 1 \end{vmatrix} = 0,$$

and

$$L(C_3M_1): \begin{vmatrix} x_1 & x_2 & 1 \\ \frac{c_3u}{1+c_3} & 0 & 1 \\ \frac{m_1u+v}{1+m_1} & \frac{w}{1+m_1} & 1 \end{vmatrix} = 0.$$

The three determinant-form equations simplify to the following:

$$L(C_1M_2): (c_1 - m_2)wx_1 + (-u - m_2u - c_1v + m_2v)x_2 + m_2uw = 0,$$
$$L(C_2M_3): (1 + m_3)wx_1 + (u + c_2u - v - m_3v)x_2 - uw = 0,$$

and

$$L(C_3M_1): (1 + c_3)wx_1 + (c_3u - m_1u - v - c_3v)x_2 - c_3uw = 0.$$

Because we assume that two lines $L(C_kM_{k+1})$ and $L(C_{k+1}M_{k+2})$ intersect at exactly one point R_k, for $k = 1, 2, 3$, we have $N_k \neq 0$, for $k = 1, 2, 3$. Thus, after algebraic calisthenics, the coordinates of three points R_k, $\{R_k\} = L(C_kM_{k+1}) \cap L(C_{k+1}M_{k+2})$, for $k = 1, 2, 3$, can be found so that

$$R_1\left(\frac{u - c_2m_2u + c_1v + m_2m_3v}{N_1}, \frac{(c_1 + m_2m_3)w}{N_1}\right),$$
$$R_2\left(\frac{c_2c_3u + m_1u + v - c_3m_3v}{N_2}, \frac{(1 - c_3m_3)w}{N_2}\right),$$

and

$$R_3\left(\frac{c_3u + m_1m_2u + c_1c_3v + m_2v}{N_3}, \frac{(c_1c_3 + m_2)w}{N_3}\right),$$

where, from the statement of Theorem 1 and for $k = 1, 2, 3$,

$$N_k = 1 + c_k + c_k c_{k+1} + m_{k+2} + m_{k+2} m_{k+1} - c_{k+1} m_{k+1}.$$

The above coordinates of R_k, for $k = 1, 2, 3$, lead us to find the area $a(R_1R_2R_3)$ of $R_1R_2R_3$, as below:

$$\begin{aligned}
a(R_1R_2R_3) &= \frac{1}{2}\begin{vmatrix} \frac{u-c_2m_2u+c_1v+m_2m_3v}{N_1} & \frac{(c_1+m_2m_3)w}{N_1} & 1 \\ \frac{c_2c_3u+m_1u+v-c_3m_3v}{N_2} & \frac{(1-c_3m_3)w}{N_2} & 1 \\ \frac{c_3u+m_1m_2u+c_1c_3v+m_2v}{N_3} & \frac{(c_1c_3+m_2)w}{N_3} & 1 \end{vmatrix} \\
&= \frac{1}{2} \times \frac{w}{N_1N_2N_3} \begin{vmatrix} u - c_2m_2u + c_1v + m_2m_3v & c_1 + m_2m_3 & N_1 \\ c_2c_3u + m_1u + v - c_3m_3v & 1 - c_3m_3 & N_2 \\ c_3u + m_1m_2u + c_1c_3v + m_2v & c_1c_3 + m_2 & N_3 \end{vmatrix} \\
&= \frac{1}{2} \times \frac{w}{N_1N_2N_3}[N_1(c_3uN) - N_2(-m_2uN) + N_3(-uN)] \\
&= \frac{N}{N_1N_2N_3}(c_3N_1 + m_2N_2 - N_3) \times \left(\frac{1}{2}u \times w\right) \\
&= \frac{N^2}{N_1N_2N_3} \times a(A_1A_2A_3), \left(\text{since } a(A_1A_2A_3) = \frac{1}{2}u \times w\right),
\end{aligned}$$

where, from the statement of Theorem 1, $N = c_1c_2c_3 + m_1m_2m_3 + c_1m_1 + c_2m_2 + c_3m_3 - 1$. Dividing each side by $a(A_1A_2A_3)$ gives the desired result and the proof of the main result is complete.

Application

As an application, we use Theorem 1 to offer a solution to a problem proposed in 2016 by Su and Lee [**22**] which gives a general solution to a problem proposed in 1987 by Herbert Guelicher [**9**]. If we add one more condition that $L(C_kM_{k+1})$ is parallel to $L(A_kA_{k+1})$, denoted by $L(C_kM_{k+1}) \parallel L(A_kA_{k+1})$, with $L(C_kM_{k+1}) \neq L(A_kA_{k+1})$, in Theorem 1, then $c_k = m_{k+1} \neq 0$, for $k = 1, 2, 3$, and equation (1) becomes

$$\frac{a(R_1R_2R_3)}{a(A_1A_2A_3)} = \frac{(2m_1m_2m_3 + m_1m_2 + m_2m_3 + m_3m_1 - 1)^2}{(1 + m_1)^2(1 + m_2)^2(1 + m_3)^2}. \tag{5}$$

By taking $l_k = \frac{1}{m_k}$, for $k = 1, 2, 3$, in equation (5), we obtain that

$$\frac{a(R_1R_2R_3)}{a(A_1A_2A_3)} = \frac{[l_1l_2l_3 - (l_1 + l_2 + l_3) - 2]^2}{(1 + l_1)^2(1 + l_2)^2(1 + l_3)^2}. \tag{6}$$

The above synthetic dissection leads us to obtain the following corollary which solves the problem proposed in 2016 by Su and Lee [**22**]. The corollary is demonstrated in Figure 9.

Corollary 1 (Su and Lee's problem [**22**])**.** *Let $A_1A_2A_3$ be a triangle in the Euclidean plane. Let $l_k \neq -1$ be a real number; let C_k and M_k be points on the line $L(A_{k+1}A_{k+2})$ such that $M_kA_{k+2} = \frac{1}{1+l_k}A_{k+1}A_{k+2}$ and $L(C_kM_{k+1}) \parallel L(A_kA_{k+1})$ with $L(C_kM_{k+1}) \neq L(A_kA_{k+1})$, for $k = 1, 2, 3$. Suppose that two lines $L(C_kM_{k+1})$ and $L(C_{k+1}M_{k+2})$ intersect at exactly one point R_k, for $k = 1, 2, 3$ (see Figure 9). Then, the ratio between*

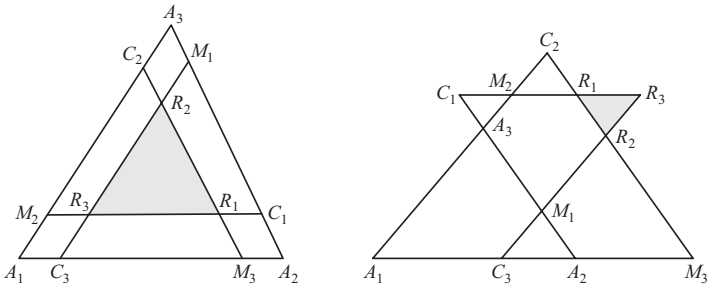


Figure 9 Two of the possible cases of Su and Lee's problem.

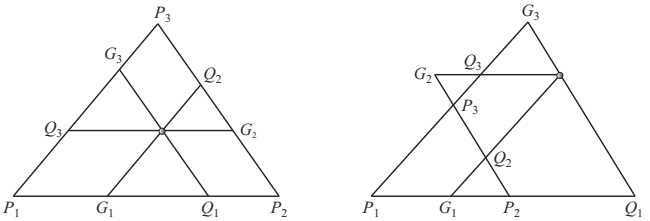


Figure 10 Two of the possible cases of Guelicher's problem.

the area $a(R_1R_2R_3)$ of the triangle $R_1R_2R_3$ and the area $a(A_1A_2A_3)$ of the triangle $A_1A_2A_3$ is given by

$$\frac{a(R_1R_2R_3)}{a(A_1A_2A_3)} = \frac{[l_1l_2l_3 - (l_1 + l_2 + l_3) - 2]^2}{(1 + l_1)^2(1 + l_2)^2(1 + l_3)^2}. \tag{7}$$

The solution to Su and Lee's problem follows from Theorem 1 since equation (7) is the same as equation (6).

As an immediate consequence of Corollary 1, by letting $a(R_1R_2R_3) = 0$ in equation (7), we solve the following problem posed in 1987 by Herbert Guelicher [**9**]; see Figure 10.

Corollary 2 (Guelicher's problem [**9**])**.** *Let $P_1P_2P_3$ be a triangle in the Euclidean plane. Let $\lambda_k \neq -1$ be a real number; let Q_k and G_k be points on the line $L(P_kP_{k+1})$ such that $Q_kP_{k+1} = \frac{1}{1+\lambda_k}P_kP_{k+1}$, and $L(Q_kG_{k+2}) = s_k \parallel p_k = L(P_{k+1}P_{k+2})$ with $s_k \neq p_k$, for $k = 1, 2, 3$, see Figure 10. Then, three lines s_k, for $k = 1, 2, 3$, are concurrent if and only if $\lambda_1\lambda_2\lambda_3 - (\lambda_1 + \lambda_2 + \lambda_3) = 2$.*

Notice that Figure 10 comes from Figure 9 by identifying $A_k = P_k$, $l_k = \lambda_k$ and taking $C_k = G_{k+1}$, $M_k = Q_{k+1}$, for $k = 1, 2, 3$, with $a(R_1R_2R_3) = 0$.

Acknowledgments The authors thank the editors and referees for their valuable suggestions and comments in improving the content of this article.

REFERENCES

[1] Bényi A., Ćurgus, B. (2013). A generalization of Routh's triangle theorem. *Amer. Math. Monthly*. 120: 841–846.
[2] Casse, R. (2006). *Projective Geometry: An Introduction*. New York: Oxford University Press.
[3] Coxeter, H. S. M., Greitzer, S.L. (1967). *Geometry Revisited*. Washington, DC: MAA, pp. 4–5, 66–67.
[4] Coxeter, H. S. M. (1969). *Introduction to Geometry*. New York: John Wiley and Sons Inc., pp. 203–221.
[5] Coxeter, H. S. M. (1987). *Projective Geometry*, 2nd ed. New York: Springer-Verlag.

[6] Eves, H. (1963). *A Survey of Geometry (I)*. Boston, MA: Allyn and Bacon, Inc., pp. 63–80, 86, 90–99, 284–294.
[7] Feng, Z. (2008). Why are the Gergonne and Soddy lines perpendicular? A synthetic approach. *Math. Mag.* 81: 211–214.
[8] Grünbaum, B., Shephard, B. G. (1995). Ceva, Menelaus, and the area principle. *Math. Mag.* 68: 254–268.
[9] Guelicher, H. (1987). Problem E3231. *Amer. Math. Monthly*. 94: 876.
[10] Holland, A. S. B. (1967). Concurrrencies and areas in a triangle. *Elemente der Math.* 22: 49–55.
[11] Houston, K. B., Powers, R.C. (2009). Simultaneous generalizations of the theorems of Ceva and Menelaus for field planes. *Int. J. Math. Educ. Sci. Tech.* 40: 1085–1091.
[12] Klamkin, M. S., Liu, A. (1981). Three more proofs of Routh's theorem, *Crux Math.* 7: 199–203.
[13] Klamkin, M. S., Liu, A. (1992). Simultaneous generalizations of the theorems of Ceva and Menelaus. *Math. Mag.* 65: 48–52.
[14] Klamkin, M. S., Kung, S. H. (1996). Ceva's and Menelaus' theorems and their converses via centroids. *Math. Mag.* 69: 49–51.
[15] Kline, J. S., Velleman, D. J. (1995). Yet another proof of Routh's theorem. *Crux Math.* 21: 37–40.
[16] Lee, C. S. (1997). The integration of math and xcience via centroids. *School Science and Mathematics*. 97: 200–205.
[17] Lipman, J. (1960). A Generalization of Ceva's theorem. *Amer. Math. Monthly*. 67: 162–163.
[18] Nakamura, H., Oguiso, K. (2003). Elementary moduli space of triangles and iterative processes. *J. Math. Sci. Univ. Tokyo*. 10:209–224.
[19] Niven, I. (1976). A new proof of Routh's theorem. *Math. Mag.* 49: 25–27.
[20] Pedoe, D. (1977). The theorems of Ceva and Menelaus. *Crux Math.* 3: 2–4.
[21] Routh, E. J. (1896). *A Treatise on Analytical Statics with Numerous Examples*, Vol. 1, Chapter IV, 2nd ed. London: Cambridge University Press, p. 82.
[22] Su, S., Lee, C.S. (2016). problem 4125, *Crux Math*. 42: 122.

Summary. In this article, we use the Cartesian Coordinate System in the Euclidean plane to establish a very general result of only one statement that comprises all the theorems of Menelaus, Ceva, Routh, and Klamkin/Liu as special cases. As an application, we use our findings to offer a solution to a problem recently proposed in 2016 by Su and Lee, which in turn gives a general solution to a problem proposed by Herbert Guelicher in 1987.

STEPHEN SU (MR Author ID: 1232148) graduated from Rio Americano High School in 2017. He now attends Stanford University. His favorite subject in mathematics is geometry.

CHENG SHYONG LEE (MR Author ID: 210841) received his M.S. and Ph.D. in Mathematics in 1982 and 1984 from Washington State University and Ed.M. in Curriculum and Teaching in 1991 from Boston University. He is a professor emeritus at the National Taipei University of Education, where he taught mathematics for over 20 years.

Hofstadter Points for Exterior Angles

I. P. D. DE SILVA
International School of Engineering
Chulalongkorn University
Pathumwan, Bangkok 10330, Thiland
prabhath.d@chula.ac.th

The law of sines and law of cosines have long been an essential part of mathematics in engineering. While these laws provide relationships between trigonometric functions of the angles (A, B, and C) and the side lengths (a, b, and c) of a triangle ABC, simple laws relating angles in their bare form to distances are not directly evident. So, although the law of sines gives $\sin(B)/\sin(C) = b/c$, a geometric representation of the ratio of the bare angles B/C as a ratio of distances is not quite intuitive. Some evidence can be seen in the definition of the Hofstadter zero point.

The Hofstadter r point ($0 \le r \le 1$) is defined in reference to a triangle ABC as follows (see, e.g., [2]). Let D be the intersection point of the extended segments constructed by rotating side BC about pivot B by an angle rB and side CB about pivot C by an angle rC; see Figure 1. Let points E and F (not shown in the figure) be constructed in a similar manner starting with sides CA and AB. Douglas Hofstadter is credited with the observation that AD, BE, and CF concur and the point of concurrence is called the Hofstadter r point. The limiting cases, as $r \to 0$ and as $r \to 1$, are called the Hofstadter zero point and the Hofstadter one point and are identified as $X(360)$ and $X(359)$, respectively, in Kimberling's encyclopedia of triangle centers [2]. A general treatment that guarantees the existence of the above two points can also be found in [1]. A careful inspection shows that the base point Z corresponding to Hofstadter zero point divides the side length BC into two segments with the ratio of the segments being inversely proportional to the ratio of the adjacent bare angles. Thus, $B/C = BZ/ZC$. This can be shown by observing the limit as r goes to 0, or graphically as illustrated in Figure 1.

Consider the scenario where the two circles ZXX_1 and ZYY_1 with centers at B and C, respectively, roll without slipping. From the initial contact point Z, let the circles roll by angles rB and rC to a state where the initial contact points are now at X and Y. The extended radii BX and CY meet at D. Let the circles continue to roll to a final state where the contact points are at X_1 and Y_1 and the extended radii meet at A. For the reference triangle ABC, as r goes to 0, then the point A goes to Z, so that the angle ratio $B/C = CZ/ZB$, which is the ratio of the radii of the rolling circles. The path of point D from $r = 1$ to $r = 0$ is shown by the curve ADZ.

Trilinear coordinates of a point $\alpha : \beta : \gamma$ are defined as the proportional directed distances from the point to the sides BC, CA, and AB, respectively, and are often used to denote a point in triangle geometry. An extensive treatment of triangle centers and trilinear coordinates can be found in [2]. The trilinear coordinates of Hofstadter zero point and Hofstadter one point, are $A/a : B/b : C/c$ and $a/A : b/B : c/C$, respectively.

Hofstadter points may also be extended to exterior angles. Accordingly, let the swing angles in this case be $r(\pi - B)$ and $r(\pi - C)$, about pivot points B and C, respectively, and let the intersection of the swing segments be D'. We will provide a proof that AD' and the similarly constructed intersection of swing segments for sides CA and AB also concur.

Math. Mag. **91** (2018) 304–306. doi:10.1080/0025570X.2018.1494438
MSC: Primary 51M04
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

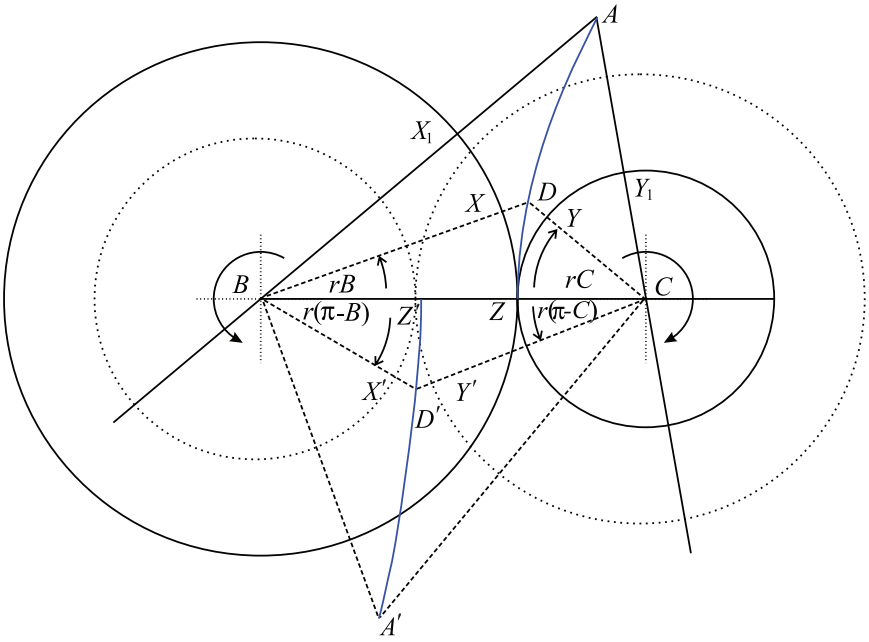


Figure 1 Graphical representation of Hofstadter zero base point and its relation to bare angles.

Graphically this situation is illustrated with dotted rolling circles. The points corresponding to X and Y are represented by X' and Y' for exterior angles. Similar arguments lead to bare external angle ratio $(\pi - B)/(\pi - C) = CZ'/Z'B$.

To prove that the exterior Hofstadter r points concur, we begin the following way. The equation of the straight line AD' in trilinear coordinates, from [3], is given by:

$$[\sin(r\bar{B})\sin(C + r\bar{C})]\gamma - [\sin(r\bar{C})\sin(B + r\bar{B})]\beta = 0,$$

where $\bar{A} = \pi - A$. Similarly, points D' and F' (not shown) are constructed considering external angles of the vertices B and C. After writing the equations for the lines BD' and CF', it can be shown that the determinant of the following matrix is zero:

$$\begin{bmatrix} 0 & \sin(r\bar{C})\sin(B + r\bar{B}) & -\sin(r\bar{B})\sin(C + r\bar{C}) \\ \sin(r\bar{C})\sin(A + r\bar{A}) & 0 & -\sin(r\bar{A})\sin(C + r\bar{C}) \\ -\sin(r\bar{B})\sin(A + r\bar{A}) & -\sin(r\bar{A})\sin(B + r\bar{B}) & 0 \end{bmatrix}.$$

Thus, the lines AD', BE' and CF' concur [4] and the trilinears of the Hofstadter exterior r point can be shown to be $\sin(r\bar{A})/\sin(A + r\bar{A}) : \sin(r\bar{B})/\sin(B + r\bar{B}) : \sin(r\bar{C})/\sin(C + r\bar{C})$. The trilinears for the limiting cases $r \to 0$ and $r \to 1$ can be shown to be $\bar{A}/a : \bar{B}/b : \bar{C}/c$, and $a/\bar{A} : b/\bar{B} : c/\bar{C}$, respectively, and are identified as $X(1115)$ and $X(7021)$ in [2].

Acknowledgments The author thanks Dr. Yan Zhao and Dr. Joshua Staubs for their help in preparing this note and the reviewers for their helpful comments.

REFERENCES

[1] Johnson, R. A. (2007). *Advanced Euclidean Geometry*. Mineola, NY: Dover, pp. 222–223.
[2] Kimberling, C., Encyclopedia of Triangle Centers. Available at: faculty.evansville.edu/ck6/encyclopedia/ETC.html

[3] Kimberling, C. (1994). Central points and central lines in the plane of a triangle. *Math. Mag.* 67: 163–187.
[4] Moses, P. J. C., Kimberling, C. (2011). Perspective isoconjugate triangle pairs, Hofstadter pairs, and cross-sums on the nine-point circle. *Forum Geom.* 11: 83–93.

Summary. The Hofstadter zero point and Hofstadter one point give relationships between bare angles of a triangle to some distances without using trigonometric functions. Graphical illustration of the ideas are presented. As a natural extension Hofstadter points related to exterior angles of a triangle are also investigated.

I. P. D. DE SILVA (MR Author ID: 626800) earned his PhD in Mechanical Engineering from Arizona State University. He is currently teaching automotive engineering at the International School of Engineering, Chulalongkorn University, Thailand.

Partiti Puzzle Solution

15 159	11 38	20 479	13 238	5 5	16 1267
11 47	8 26	6 15	6 6	13 49	11 38
14 158	12 39	11 47	3 3	9 27	12 156
13 247	6 6	10 28	6 15	12 48	9 9
11 38	6 15	13 49	10 37	8 26	6 15
21 2469	7 7	16 268	6 15	13 49	18 378

Veronika Irvine: The Art and Mathematics of Making Bobbin Lace

ALLISON HENRICH
Seattle University
Seattle, WA 98122
henricha@seattleu.edu

Veronika Irvine has a Master of Applied Mathematics from the University of Waterloo and a Ph.D in Computer Science from the University of Victoria, where she studied combinatorial objects. She is currently a postdoctoral fellow at the Cheriton School of Computer Science at the University of Waterloo. In addition to her interest in math and computer science, she is a fiber artist and has been making lace for over 35 years. For her research, she applies ideas from computer science to explore and expand the techniques used in bobbin lace design. Irvine was interviewed for this article from her home in Waterloo on February 23, 2018. Images of her work appear on pages 261, 266, 285, 287, and 320.

Q: *How did you first get interested in making lace?*

VI: As a kid, I enjoyed learning to make fiber crafts such as crochet, knitting, and cross-stitch. My first memorable encounter with lace happened when I was about five years old. On my mom's dresser, there was a small pile of lace fragments that I would later learn were made by a technique called tatting. I spent hours trying to figure out how the pieces connected together. In some places, the knots were falling apart, revealing more about their structure. I wondered how they had been created. The pieces were made by my grandmother—she was a talented lacemaker. I never got to meet her because she died when my mom was very young. The fragments were from the outer ring of the very last doily she ever worked on. While it makes me sad to reflect on the reason why these pieces were never joined together, the fact that they came into my life in this unfinished, disconnected state is what attracted my attention and held my interest, imagining the process and the piecing together the puzzle. I think my interest in designing lace came from that early experience.

Q: *When did it first occur to you that there might be some interesting math behind the lace-making process?*

VI: It really started when I discovered bobbin lace in my mid-20s. Bobbin lace is complicated, but very logical. It is kind of like doing a Sudoku puzzle. Here, the puzzle is to figure out what threads to braid together and in what order, to find a successful path through the pattern.

Q: *What does success mean in this context? Does it mean producing a piece that's aesthetically pleasing or something else?*

VI: Unlike a pattern for knitting or crochet, a bobbin lace pattern is not expressed as a linear sequence of written steps. A bobbin lace pattern is typically a diagram. There are many possible ways different threads can come together and most of those are not successful, meaning you could end up with too many threads in one area and not

Math. Mag. **91** (2018) 307–309. doi:10.1080/0025570X.2018.1503465
MSC: Primary 01A70
Allison Henrich (MR Author ID: 900050)
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

enough threads in another. Making bobbin lace is definitely not something you can do while distracted. You have to focus and think about the flow of threads in an analytical way. Designing a new pattern is even more complicated because you must guarantee that a successful path exists.

Q: *This makes me wonder if you've ever made a mistake as you've tried to make a bobbin lace pattern that has led to a useful insight or discovery about lace patterns.*

VI: Many times, I have come to a spot and realized I didn't have the right number of threads to continue. Fortunately, unlike Sudoku, the threads hold an ordered record of my choices so I can usually backtrack and find my mistake.

Q: *Is that insight about how you can fail helpful for creating algorithms that generate lace patterns?*

VI: Yes, absolutely. Failure was also really helpful in refining the model. I wrote a computer program to generate lace patterns based on the rules of the model. When the program generated patterns that were impossible to make, it helped me realize that there is something else about these patterns that I haven't captured in the model.

Q: *What kind of mathematics is useful for studying lace patterns and creating algorithms to generate new patterns?*

VI: Studying lace has led me to learn about many different kinds of math. It seems every new problem that I look at touches on a new area. The main work that I've done so far has used a lot of graph theory, some group theory, and some combinatorics. For example, I can represent the lace pattern as a graph drawing created by gluing together a bunch of lattice paths. To get an upper limit on the number of graph drawings that can be created this way, I needed to count how many lattice paths of a certain length exist. So I determined their generating function—a process that introduced me to the kernel method, Cauchy's integral formula, and residue calculus. I also found a bijection between my lace paths and other kinds of lattice paths that people have studied for completely different reasons.

More recently, I have been exploring how we can tweak a drawing of a lace graph using an algorithm very similar to the Tutte embedding of a three-connected planar graph, but on a torus. I am interested in altering the maximally symmetric pattern to make it a little more complicated and aesthetically interesting. How far can we modify the shapes of the faces before the algorithm breaks? Studying this question has so far required me to look more deeply into linear programming and also learn about rigidity theory.

Q: *Thinking more about your work from an artist's perspective, is there a specific piece of yours that you think is really representative of the type of art that you do?*

VI: Before 2010, I made lace from patterns that other people had designed, so up until that point, I can't say there is anything that's representative of me. I was learning the craft by copying, using practices that have been followed for hundreds of years. In 2010, I started creating my own patterns, exploring bobbin lace as an open, living medium. I respect the tradition, but I am not restricted by it. I explore the limits by breaking some of the rules, challenge what I have learned, and sometimes I get away with it. So, the piece that represents me best is probably whatever piece I am currently working on.

My most recent finished piece is *Delle Caustiche* (pictured in Figure 1), named after a star cloud in the Sagittarius arm of the Milky Way. For this piece, I designed a 2D pattern with *632 orbifold symmetry. The symmetry was a bit too simple, too saccharine, so I worked with a student to write a computer program that morphs the pattern by mapping it onto a ring. I made the lace by braiding threads around the ring,

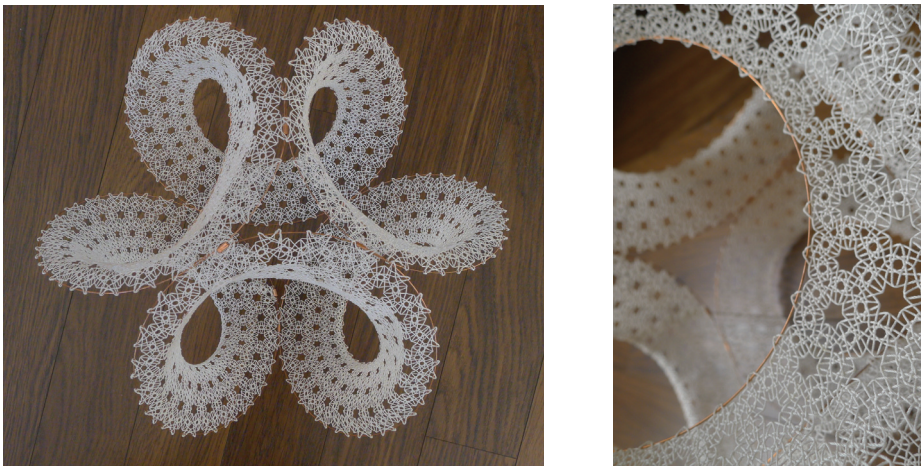


Figure 1 *Delle Caustiche* (2017).

but instead of stopping at one complete rotation, I kept going to create a spiral of 1080 degrees. Finally, I connected the start and end together in a way that had no twists. The finished piece has three full rotations—too much curvature to sit flat in the plane. That's why it has undulations, or waves.

Q: *Can you tell me a bit about your involvement in STEM educational outreach with middle school and high school kids?*

VI: My dream is to see many more women and indigenous people in computer science and math. I was thinking about why there is currently an imbalance and also why I like what I do. I'm not a gamer. I wouldn't classify myself as a nerd, someone who writes programs for the sake of writing programs. But that's often the perception of what a computer scientist is. I was always attracted to the artistic, creative side of computer science. You can make very interesting designs using algorithms and equations. If more people could see the rich diversity of topics that come under the umbrella of computer science, then maybe they'd be interested in trying it out.

Q: *Is there anything else I didn't ask about that you would like people to know about your work?*

VI: There are crafts that are on the verge of extinction. They're not practiced by very many people, and most of the people who practice them are getting much older. They require skills and ideas that will be lost when the craft is forgotten. In my experience, crafts are a great source of problems to explore mathematically, but you need to have some experience with them to bring them within the realm of mathematics. And it's not just about solving math problems for the sake of solving math problems; bringing crafts and math together could open up many new applications. For instance, lace can be more than just an ornament. It combines fibers into textiles in a very sophisticated way. My collaborators and I can computationally generate patterns for fabrics with properties that can't be achieved by weaving or knitting. In addition to making new kinds of art, there may be medical applications or applications in architecture.

This is another reason why I want to see greater diversity in computer science and mathematics. The topics that I study, to the best of my knowledge, haven't been explored before. Primarily, bobbin lace has been a craft practiced by women. Since women are underrepresented in math and computer science, it's not too surprising that bobbin lace hasn't been studied from this perspective before. There are many other crafts that have the same potential.

PROBLEMS

EDUARDO DUEÑEZ, *Editor*
University of Texas at San Antonio

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

JOSÉ A. GÓMEZ, Facultad de Ciencias, UNAM, Mexico; *CODY PATTERSON,* University of Texas at San Antonio; *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 March 1, 2019.

2051. *Proposed by Ángel Plaza, Universidad de Las Palmas de Gran Canaria, Spain.*

For any positive integer n consider the polynomial $P_n(x) = x^4 - nx^3 - nx^2 - nx + 1$ and let a_n be the largest of its real roots. Find

$$\lim_{n\to\infty} \frac{a_1 + a_2 + \cdots + a_n}{n^2}.$$

2052. *Proposed by Michel Bataille, Rouen, France.*

Let $\triangle ABC$ be a scalene triangle. Let D be a variable point on line $\overleftrightarrow{BC}$ such that $D \neq B$ and $D \neq C$. Let E lie on $\overleftrightarrow{BC}$ so $\overleftrightarrow{AE}$ is the reflection of $\overleftrightarrow{AD}$ across the bisector of angle $\angle BAC$. Let O_1, O_2 be the circumcenters of triangles $\triangle ABD$ and $\triangle ACD$, respectively. Prove that there exists a point P, independent of the choice of D, such that line $\overleftrightarrow{O_1O_2}$ passes through P.

2053. *Proposed by Sung Soo Kim, Hanyang University, Korea.*

Let $a = (a_1, a_2, \ldots, a_{2018})$ be a permutation of the integers $1, 2, \ldots, 2018$. For any integer k in the range $1 \leq k \leq 2018$, let $l_k(a)$ be the length of the longest monotone subsequence of $(a_k, a_{k+1}, \ldots, a_{2018})$ whose first term is a_k, and let $L(a) = \sum_{k=1}^{2018} l_k(a)$. Find the minimum value of $L(a)$ as a ranges over all permutations of $1, 2, \ldots, 2018$.

2054. *Proposed by Florin Stanescu, Şerban Cioculescu school, Găeşti, Romania.*

Math. Mag. **91** (2018) 310–316. doi:10.1080/0025570X.2018.1501259.

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.

Let $f : [0, 1] \to \mathbb{R}$ be differentiable with bounded derivative. If $\int_0^1 xf(x)dx = 0$, prove that

$$36 \cdot \left| \int_0^1 x^2 f(x)dx \right| \leq \sup_{x\in[0,1]} |f'(x)|.$$

2055. *Proposed by Ioan Băetu, Botoşani, Romania.*

Let n be a cube-free positive integer. Assume that G is a finite group of order n such that for every subgroup H of G and every automorphism f of H, the equality $K = \{f(x) : x \in K\}$ holds for every subgroup K of H. Prove that G is cyclic.

Quickies

1083. *Proposed by Konstantinos Gaitanas, Volos, Greece.*

Prove that there exist infinitely many quadruples (a, b, c, d) of pairwise different positive integers such that

$$\gcd(a, b, c, d) = 1, \qquad ab = cd, \qquad \text{and} \qquad a + b \mid c + d.$$

1084. *Proposed by George Stoica, New Brunswick, Canada.*

Let a, b be positive real numbers such that $a < b$. Let f be a continuous real-valued function on $[a, b]$. Are there $u, v \in [a, b]$ such that $f(u) = f(v)$ and u/v is irrational?

Solutions

A triangle with integer sides and one angle thrice another **October 2017**

2026. *Proposed by Lokman Gökçe, Turkey.*

The angles of triangle $\triangle ABC$ satisfy $\angle A = 3 \cdot \angle B$. What is the least possible perimeter of $\triangle ABC$ assuming its side lengths are integers?

Solution by Robin Chapman, University of Exeter, UK.

We prove that the least possible perimeter is 21, attained when the side lengths are $a = BC = 10$, $b = AC = 8$ and $c = AB = 3$. Let $p = a + b + c$ be the perimeter of $\triangle ABC$ and let $\theta = \angle B$, so $\angle A = 3\theta$ and $\angle C = \pi - 4\theta$, where evidently $0 < \theta < \pi/4$. By the law of sines and standard trigonometric identities:

$$\frac{a}{b} = \frac{\sin 3\theta}{\sin\theta} = 4\cos^2\theta - 1 \quad \text{and} \quad \frac{c}{b} = \frac{\sin(\pi - 4\theta)}{\sin\theta} = \frac{\sin 4\theta}{\sin\theta} = 8\cos^3\theta - 4\cos\theta.$$

By the law of cosines,

$$t = 2\cos\theta = \frac{a^2 + c^2 - b^2}{ac} \tag{1}$$

is rational; moreover,

$$\frac{a}{b} = t^2 - 1 \qquad \text{and} \qquad \frac{c}{b} = t^3 - 2t, \tag{2}$$

so

$$p = a + b + c = b(t^3 + t^2 - 2t) = bt(t + 2)(t - 1). \tag{3}$$

Write $t = u/v$ in lowest terms with u and v coprime positive integers. Then, we have

$$p = \frac{bu(u + 2v)(u - v)}{v^3}.$$

Since v is coprime to u, it is also coprime to $u + 2v$ and $u - v$; thus, $u(u - v)(u + 2v)/v^3$ is a reduced fraction. The perimeter p is an integer, hence v^3 divides b, and

$$p \geq u(u + 2v)(u - v).$$

Since $\theta \in (0, \pi/4)$, we have $u/v = t = 2\cos\theta \in (\sqrt{2}, 2)$; therefore, $u > v > 1$. It follows that $u + 2v \geq 3 + 2 \cdot 2 = 7$, so

$$p \geq u(u + 2v)(u - v) \geq 3 \cdot 7 \cdot 1 = 21.$$

The minimum perimeter $p = 21$ is achieved when $a = 10$, $b = 8$, $c = 3$ and $t = 3/2$. These values evidently satisfy all conditions of the problem in view of equations (1)–(3).

Also solved by Armstrong Problem Solvers, Herb Bailey, Robert Calcaterra, John Christopher, Joseph DiMuro, Robert L. Doucette, Madison Estabrook, Habib Y. Far & Jeffrey Groah, Russell Gordon, Raymond N. Greenwell, Eugene A. Herman, Megan Kaiser, Elias Lampakis (Greece), Kee-Wai Lau (China), László Lipták, David E. Manes, Peter McPolin (Northern Ireland), Missouri State University Problem Solving Group, Michael Reid, Iason Rusodimos & Barrett Walls, Ioannis D. Sfikas (Greece), Nicholas C. Singer, David Stone & John Hawkins, Michael Vowe (Switzerland), Edward White & Roberta White, and the proposer. There was 1 incomplete or incorrect solution.

Sums of squares and products of consecutive integers **October 2017**

2027. *Proposed by Marian Tetiva, National College "Gheorghe Roşca Codreanu", Bîrlad, Romania.*

Prove that the equation

$$n(n + 1)(n + 2)(n + 3) = a^2 + b^2$$

admits infinitely many solutions in integers a, b, n.

Solution by Nicholas Singer, Annandale, VA.
Call a positive integer k *special* if $k(k + 1)/2$ is a perfect square. (Note that $k(k + 1)/2$ is necessarily an integer when k is an integer.) The first four special numbers are 1, 8, 49, and 288. If k is special, then so is the larger integer $k' = 4k(k + 1)$ since

$$k'(k' + 1)/2 = 8k^4 + 16k^3 + 10k^2 + 2k = (4k + 2)^2 \cdot k(k + 1)/2$$

is a perfect square, being the product of two perfect squares by the hypothesis that k is special. It follows that there exist infinitely many special numbers. For each special number k, the problem has a distinct solution given by $n = 3k$, $a = 9k^2 + 9k$, and $b = 6\sqrt{k(k + 1)/2}$, since

$$n(n + 1)(n + 2)(n + 3) = 3k(3k + 1)(3k + 2)(3k + 3) = 81k^4 + 162k^3 + 99k^2 + 18k$$
$$= (9k^2 + 9k)^2 + 6^2 \cdot k(k + 1)/2 = a^2 + b^2.$$

Also solved by Michael R. Bacon & Charles K.Cook, Jordan Bailey, Brian D. Beasley, Anthony J. Bevelacqua, Robert Calcaterra, Robin Chapman (UK), John Christopher, Joseph DiMuro, Csaba P. Gabor, Abhay Goel (student), Russell Gordon, José Hernández Santiago (student, Mexico), Elias Lampakis (Greece), Missouri State University Problem Solving Group, Michael Reid, Herman Roelants (Belgium), Achilleas Sinefakopoulos (Greece), Ioannis D. Sfikas (Greece), Texas State Problem Solvers Group, Michael Vowe (Switzerland), Catherine Way and Bianca Reilly (students), Edward and Roberta White, and the proposer.

A limit related to the equality case of AM–GM **October 2017**

2028. *Proposed by Oniciuc Gheorghe, Botosani, Romania.*

Is there a nonconstant sequence $\{a_n\}_{n\geq 1}$ in the interval $(0, 1)$ such that

$$\lim_{n\to\infty}(1+a_1-a_2)(1+a_2-a_3)\dots(1+a_{n-1}-a_n)(1+a_n-a_1)=1?$$

Solution by Northwestern University Math Problem Solving Group, Evanston, IL.
We prove that answer is no. Let $L_n=(1+a_1-a_2)(1+a_2-a_3)\cdots(1+a_{n-1}-a_n)(1+a_n-a_1)$. Using a first-order Taylor approximation with Lagrange remainder, for all $x\in(-1,1)$ we have

$$\log(1+x)=x-\frac{x^2}{2(1+\theta)^2}\qquad\text{for some }\theta\text{ strictly between 0 and }x.$$

Applying the above approximation for $x=a_1-a_2$, $x=a_2-a_3$, …, $x=a_{n-1}-a_n$, $x=a_n-a_1$ (which are numbers in $(-1,1)$ by the hypothesis that $0<a_n<1$):

$$\begin{aligned}\log L_n&=\log(1+a_n-a_1)+\sum_{i=1}^{n-1}\log(1+a_i-a_{i+1})\\&=(a_n-a_1)-\frac{(a_n-a_1)^2}{2(1+\theta_n')^2}+\sum_{i=1}^{n-1}\left((a_i-a_{i+1})-\frac{(a_i-a_{i+1})^2}{2(1+\theta_i)^2}\right)\\&=-\frac{(a_n-a_1)^2}{2(1+\theta_n')^2}-\sum_{i=1}^{n-1}\frac{(a_i-a_{i+1})^2}{2(1+\theta_i)^2}\\&\qquad\text{(since }(a_n-a_1)+\textstyle\sum_{i=1}^{n-1}(a_i-a_{i+1})=0\text{),}\end{aligned}$$

where θ_i $(i=1,2,\dots,n-1)$ is some real number—depending on i but not on n—that lies between 0 and a_i-a_{i+1}, while θ_n' is between 0 and a_n-a_1. Note that all terms on the right-hand side of the identity above are nonpositive. If $\{a_n\}$ is nonconstant then $a_k\neq a_{k+1}$ for some k, hence $\log L_n\leq\alpha=-(a_k-a_{k+1})^2/[2(1+\theta_k)^2]<0$ for all n, hence $\lim_{n\to\infty}L_n\leq\exp(\alpha)<1$. We conclude that $\lim_{n\to\infty}L_n=1$ implies that $\{a_n\}$ is constant.

Editor's Note. For fixed n, the positive numbers $b_1=1+a_1-a_2$, $b_2=1+a_2-a_3$, …, $b_{n-1}=1+a_{n-1}-a_n$, $b_n'=1+a_n-a_1$ obviously have arithmetic mean 1. The equality $L_n=1$ for fixed n implies that $b_1,\dots,b_{n-1},b_n'$ have geometric mean $\sqrt[n]{L_n}=1$ equal to their arithmetic mean. The conditions for equality in the AM–GM inequality give $b_1=\cdots=b_{n-1}=b_n'=1$, so $a_1=\cdots=a_n$. The solution above shows that the seemingly weaker assumption $\lim_{n\to\infty}L_n=1$ actually implies that $\{L_n\}$ is the constant sequence with value 1, and thus $\{a_n\}$ is also constant.

Also solved by Robert Calcaterra, Joseph DiMuro, Reiner Martin (Germany), Omarjee Moubinool (France), Nicolas Singer, and the proposer.

A hierarchy of "plenary" collections of natural numbers **October 2017**

2029. *Proposed by Kenneth Schilling, University of Michigan-Flint, MI.*

Let $\mathbb{N}$ be the set of natural numbers. For any function $g : \mathbb{N} \to \mathbb{N}$ and a subset Y of $\mathbb{N}$, we denote by $g[Y] = \{g(y) : y \in \mathbb{N}\}$ the forward image of Y by g. Given a natural number n, a collection $\mathcal{C}$ of subsets of $\mathbb{N}$ is called:

- *n-plenary* if there exist functions $f_1, f_2, \ldots, f_n : \mathbb{N} \to \mathbb{N}$ such that, for all $X \in \mathcal{C}$, $f_1[X] \cup f_2[X] \cup \cdots \cup f_n[X] = \mathbb{N}$.
- *strongly n-plenary* if there exist $f_1, f_2, \ldots, f_n : \mathbb{N} \to \mathbb{N}$ such that, for all $X \in \mathcal{C}$, either $f_1[X] = \mathbb{N}$, or $f_2[X] = \mathbb{N}$, ..., or $f_n[X] = \mathbb{N}$.

For every natural number n:

(a) Prove that a collection $\mathcal{C}$ is n-plenary if and only if it is strongly n-plenary.

(b) Construct an $(n + 1)$-plenary collection $\mathcal{C}$ that is not n-plenary.

Solution by Joseph DiMuro, Biola University, La Mirada, CA.

(a) It suffices to show that all n-plenary collections are strongly n-plenary; the converse is obvious.

Let $\mathcal{C}$ be an n-plenary collection, and let $f_1, \ldots, f_n$ be the corresponding functions. Fix any bijection $g : \mathbb{N} \to \mathbb{N}^n$. For $i = 1, 2, \ldots, n$, let $g_i : \mathbb{N} \to \mathbb{N}$ be the i-th coordinate function of g, so $g(x) = (g_1(x), g_2(x), \ldots, g_n(x))$ for all $x \in \mathbb{N}$. Define functions $h_1, \ldots, h_n : \mathbb{N} \to \mathbb{N}$ by $h_i(x) = g_i(f_i(x))$. We prove that $h_1, \ldots, h_n$ witness the strong n-plenarity of $\mathcal{C}$, i.e., that for all $X \in \mathcal{C}$ there is i such that $h_i[X] = \mathbb{N}$. Assume the contrary, i.e., that there exists $X \in \mathcal{C}$ such that $h_i[X] \neq \mathbb{N}$ for all i. For each i, choose $y_i \in \mathbb{N}$ such that $y_i \notin h_i[X]$. Let $y = g^{-1}(y_1, \ldots, y_n)$. By n-plenarity of $\mathcal{C}$ and the choice of $f_1, \ldots, f_n$, we have $\bigcup_{i=1}^{n} f_i[X] = \mathbb{N}$, so there must exist some i and some $x \in X$ such that $f_i(x) = y$, and thus $y_i = g_i(y) = g_i(f_i(x)) = h_i(x) \in h_i[X]$, contradicting the choice of y_i. This concludes the proof that any n-plenary collection is strongly n-plenary.

(b) Fix any bijection $f : \mathbb{N} \to \mathbb{N}^{n+1}$ and let $f_1, f_2, \ldots, f_{n+1} : \mathbb{N} \to \mathbb{N}$ be the coordinate functions of f. Let $\mathcal{C}$ be the collection consisting of those subsets X of $\mathbb{N}$ such that $f_i[X] = \mathbb{N}$ for some i. By construction, $\mathcal{C}$ is strongly $(n + 1)$-plenary and hence $(n + 1)$-plenary. We show that $\mathcal{C}$ is not n-plenary.

On the contrary, if $\mathcal{C}$ were n-plenary, there would exist functions $g_i : \mathbb{N} \to \mathbb{N}$ $(i = 1, \ldots, n)$ such that $\bigcup_{i=1}^{n} g_i[X] = \mathbb{N}$ for all $X \in \mathcal{C}$. In this case, we claim that there exists $m_1 \in \mathbb{N}$ such that, for all $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{N}^n$, there exists i such that $g_i(f^{-1}(m_1, \mathbf{x})) = 1$. Otherwise, for all $m \in \mathbb{N}$, there exists $\mathbf{x}^{(m)} = (x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, \ldots, x_n^{(m)}) \in \mathbb{N}^n$ such that $g_i(f^{-1}(m, \mathbf{x}^{(m)})) \neq 1$ for all i. It follows that the set $X = \{f^{-1}(m, \mathbf{x}^{(m)}) : m \in \mathbb{N}\}$ satisfies $f_1[X] = \mathbb{N}$, so $X \in \mathcal{C}$, and also, $1 \notin g_i[X]$ for all i, so $1 \notin \bigcup_{i=1}^{n} g_i[X]$. This contradicts the assumption $\bigcup_{i=1}^{n} g_i[X] = \mathbb{N}$. Thus, $m_1 \in \mathbb{N}$ exists with the property claimed. By the same reasoning, for $k = 2, \ldots, n + 1$, there exists $m_k \in \mathbb{N}$ such that for all $(x_1, x_2, \ldots, x_n) \in \mathbb{N}^n$ there is i satisfying

$$g_i(f^{-1}(x_1, x_2, \ldots, x_{k-1}, m_k, x_k, x_{k+1}, \ldots, x_{n-1}, x_n)) = k.$$

Let $\mathbf{m} = (m_1, m_2, \ldots, m_{n+1})$. By construction of $m_1, m_2, \ldots, m_{n+1}$, for all k with $1 \leq k \leq n + 1$, there exists i with $1 \leq i \leq n$ such that $g_i(f^{-1}(\mathbf{m})) = k$. Thus, the image of mapping $i \mapsto g_i(f^{-1}(\mathbf{m}))$ from $\{1, 2, \ldots, n\}$ into $\mathbb{N}$ contains each of the numbers $1, 2, \ldots, n + 1$, which is impossible since $n < n + 1$. This contradiction shows that $\mathcal{C}$ is not n-plenary.

Also solved by the proposer.

Matrices with constant diagonal and repeated eigenvalues **October 2017**

2030. *Proposed by Rajesh Sharma, Himachal Pradesh University, India.*

(a) Prove that a real symmetric 3×3 matrix with equal diagonal entries has a multiple eigenvalue if and only if its nondiagonal entries all have the same absolute value.

(b) Characterize all complex Hermitian 3×3 matrices with equal diagonal entries and a multiple eigenvalue.

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

We answer part (b) first. Consider a general complex Hermitian matrix with constant diagonal, say

$$H = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} = \begin{pmatrix} d & a & \overline{c} \\ \overline{a} & d & b \\ c & \overline{b} & d \end{pmatrix}$$

for arbitrary $a, b, c \in \mathbb{C}$ and $d \in \mathbb{R}$. Its characteristic polynomial is

$$\phi(\lambda) = (\lambda - d)^3 - 3m^2(\lambda - d) - (abc + \overline{abc}),$$

where $m = \sqrt{(|a|^2 + |b|^2 + |c|^2)/3}$. The roots of $\phi'(\lambda) = 3(\lambda - d)^2 - 3m^2$ are $\lambda = d \pm m$. A multiple eigenvalue of H is precisely a common root of ϕ and ϕ', so we must have $\phi(d + m) = 0$ or $\phi(d - m) = 0$, i.e.,

$$abc + \overline{abc} = 2m^3 \qquad \text{or} \qquad abc + \overline{abc} = -2m^3.$$

Thus, the multiple eigenvalue condition is equivalent to

$$\left(\frac{abc + \overline{abc}}{2}\right)^2 = m^6 = \left(\frac{|a|^2 + |b|^2 + |c|^2}{3}\right)^3. \tag{1}$$

Writing a, b, c in polar form as $a = |a| \exp(i\alpha)$, $b = |b| \exp(i\beta)$, $c = |c| \exp(i\gamma)$, equation (1) assumes the form

$$\sqrt{\frac{|a|^2 + |b|^2 + |c|^2}{3}} = \sqrt[3]{|a||b||c|} \cdot \sqrt[3]{|\cos(\alpha + \beta + \gamma)|}. \tag{2}$$

The above equality holds if $a = b = c = 0$. Otherwise, a, b, c must all be nonzero and by the QM–GM (quadratic mean–geometric mean) inequality, equation (2) holds precisely when $|a| = |b| = |c|$ and $|\cos(\alpha + \beta + \gamma)| = 1$. The last condition means that $\arg(abc) = \alpha + \beta + \gamma$ equals a multiple of π, i.e., $abc \in \mathbb{R}$ (which still holds in case $a = b = c = 0$). In conclusion, a complex Hermitian 3×3 matrix H with equal diagonal entries has multiple eigenvalues if and only if all nondiagonal entries have the same absolute value and $H_{12}H_{23}H_{31}$ is real.

The solution to part (a) follows immediately, because a real symmetric matrix is just a real Hermitian matrix, so the condition for multiple eigenvalues in this case is simply that all off-diagonal entries have the same absolute value.

Also solved by Robert Calcaterra, Joseph DiMuro, James Duemmel, Dmitry Fleischmann, Eugene A. Herman, Elias Lampakis (Greece), Missouri State University Problem Solving Group, Ioannis Sfikas (Greece), Nicholas Singer, Jeffrey Stuart, Nora Thornber, and the proposer.

Answers

Solutions to the Quickies from page 311.

A1083. For any integer $n \geq 2$, let

$$a = 2^n - 1, \qquad b = 2^n + 1, \qquad c = 1, \qquad d = 2^{2n} - 1.$$

Because $c = 1$, we have $\gcd(a, b, c, d) = 1$. Since $n \geq 2$, it is clear that $c < a < b < d$, $ab = 2^{2n} - 1 = cd$, and $a + b = 2^{n+1}$ divides $c + d = 2^{2n}$.

A1084. We prove that such u, v exist. If f is constant it suffices to take u rational and v irrational in $[a, b]$. Otherwise, if f is not constant, there is $c \in [a, b]$ with $f(c) \neq f(a)$. Let $p = f(a) = f(b)$ and $q = f(c)$. We consider the case $q > p$ (the case $q < p$ is analogous). It follows from the intermediate value theorem and the choice of q that, for any $y \in [p, q]$, the bounded set $A(y) = \{x \in [a, b] : f(x) \geq y\}$ is nonempty. Define real-valued functions g, h, k on $[p, q]$ by

$$g(y) = \inf A(y), \qquad h(y) = \sup A(y), \qquad k(y) = \frac{g(y)}{h(y)}.$$

Note that $A(y)$ is closed and bounded, so $\inf A(y) = \min A(y)$ and $\sup A(y) = \max A(y)$. Again, by the intermediate value theorem, it is easy to see that $f(g(y)) = y = f(h(y))$ holds for $y \in [p, q]$, that g is strictly increasing while h is strictly decreasing, and hence that k is also strictly increasing and thus injective. Since $[p, q]$ is uncountable, k must take an irrational value $k(z)$ for some $z \in [p, q]$. Thus, we have $k(z) = u/v$ is irrational, where $u = g(z)$ and $v = h(z)$ satisfy $f(u) = z = f(v)$.

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.

Ciesielska, Danuta, and Krzysztof Ciesielski, Equidecomposability of polyhedra: A solution of Hilbert's Third Problem in Kraków before ICM 1900, *Mathematical Intelligencer* 40 (2) (2018) 55–63.

Thanks to this article, we now know that Ludwik Antoni Birkenmajer (1855–1929) in 1883 gave conditions for two tetrahedra with equal volumes so that either can be cut by planes into finitely many polyhedral pieces that can be reassembled to form the other. This solution preceded Hilbert's formulation of the problem in his famous lecture at the International Congress of Mathematicians in 1900. Birkenmajer won a prize of the Academy of Arts and Sciences in Kraków for his solution; his article, written in Polish, was never published. Max Dehn has been credited with solving the problem, and the authors of this article—despite their findings about Birkenmajer's solution—surprisingly reassert that attribution because "[Dehn] was the first one to publish the correct proof." Hmm. . . haven't we seen this sort of conflict before, e.g., with Newton and Leibniz?

Pieronkiewicz, Barbara, Mathematicians who never were, *Mathematical Intelligencer* 40 (2) (2018) 45–49.

Some mathematicians have written under pseudonyms (I was once tempted myself); a few have even created entire personas for their pseudonymous avatars. Nicolas Bourbaki (1934–1968), a product of the collective imagination of French mathematicians, is perhaps the most famous mathematician who never existed; he was rejected twice for individual membership in the American Mathematical Society (but perhaps could have secured an institutional membership). Other such phantasms are noted here, including H. Pétard (famous for the mathematical theory of big game hunting) and others created by Ralph Boas. The names of dogs, cats, and computers belonging to mathematicians have also appeared as coauthors of papers (justification for the computer: "credit where credit is due"). At the end of this article, the author asserts her own existence and that she is not an animal of any sort (but does not claim not to be a computer. . .).

Posamentier, Alfred S., and Christian Spreitzer, *The Mathematics of Everyday Life*, Prometheus Books, 2018; 398 pp, $25(P), ISBN 978-1-63388-387-1; ebook $11.99, ISBN 978-1-63388-388-8.

This book offers a compendium of "topics and ideas most of which will not have been encountered during the school instruction." The topics include the golden ratio and Fibonacci numbers in art, nature, and investment strategies; the optimal point on the sideline to shoot for a soccer goal; algorithms for the date of Easter; Reuleaux triangles and fire hydrant valves; wrapping presents; the rencontres problem (without mention of its application to the Midwestern holiday "guy gift" custom); ballistic trajectories (only without air resistance); billiards; bicycle gearing; Spirograph images; the rainbow; the Regiomontanus problem of optimal viewing angle of a statue; how GPS works; and lots more—altogether, plenty to whet the reader's curiosity to see mathematics in the world around, and to seek out more.

Math. Mag. **91** (2018) 317–318. doi:10.1080/0025570X.2018.1504590.

Neale, Vicky, *Closing the Gap*, Oxford University Press, 2018; vi + 164 pp, $25.95. ISBN 978-0-19-878828-7.

This book about the twin primes conjecture tells the tale of Yitang Zhang's discovery in 2013 that there are infinitely many pairs of prime numbers that differ by at most 70 million. The author sets out that and subsequent discoveries in chronological order, salting in material about Goldbach's conjecture, primes as sums of squares, and more—remarkably, with scarcely any formulas or notation. She notes the importance of a collaborative Polymath project in improving the 70 million bound in short order to 246 (and possibly 6, if a certain conjecture is true).

Davey, Samuel, et al., *Bayesian Methods in the Search for MH370*, Springer, 2016; ix + 114 pp, $24.99(P). ISBN 978-981-10-0378-3.

I would like to be able to update this book with news that the Malaysian Airlines plane that went missing in 2014 has been found in the very area prescribed for searching by the Bayesian analysis in this book. But it wasn't found there, nor in a subsequent search 1,000 miles north of there. The Malaysian government gave up the search in May 2018. The book's analysis describes the prior distribution used (based on radar data), a likelihood function concerning the measurements, and assumptions about dynamics of the aircraft. The posterior distribution was relatively unaffected by an initial finding of debris at Réunion Island (more debris has been found since, there and elsewhere).

Nahin, Paul J., *How to Fall Slower Than Gravity: And Other Everyday (and Not So Everyday) Uses of Mathematics and Physical Reasoning*, Princeton University Press, 2018; xxxiii + 271 pp, $27.95. ISBN 978-0-691-17691-8.

This book is the author's response to claims that "algebra is the new Latin" and that students' time would be better spent "learning life skills and . . . how to reason." So author Nahin assembles here a collection of problems that illustrate how mathematics (up through calculus) can "provide insight into *real* problems." He is a mathematical physicist, so the problems (with solutions both in mathematics and in computer code) come mostly from physics. Particularly meaningful for me was the last problem, how to find a fault in an underwater electrical cable; the 4-mile-long electrical cable to Washington Island, Wisconsin, failed in June.

Kerins, Bowen, Darryl Yong, Al Cuoco, Glenn Stevens, and Mary Pilgrim, *Fractions, Tilings, and Geometry* and *Probability and Games*, American Mathematical Society, 2017; xi + 158 pp and xi + 137 pp, $29(P) each. ISBN 978-1-4704-4064-0, 978-1-4704-4063-3.

These two books are part of a remarkable new series based on courses offered in the Summer School Teacher Program at the Park City Mathematics Institute. Designed for precollege teachers, as "serious mathematics connected to secondary teaching," each book consists of a sequenced collection of 14 problem sets offering mathematical discovery (nonperiodic tilings in 2D and space-filling polyhedra; mathematical election analysis through probability and statistics). There are also facilitator notes on each problem set and solutions. A touchstone: "The problems should lead to the appropriate mathematics rather than requiring it." Other volumes in the series treat applications of algebra and geometry (complex numbers, including Gaussian and Eisenstein integers), structural similarities (from card-shuffling to expansions of rational numbers), functions in number theory, and geometric thinking. This is *exactly* the sort of experience to invigorate in-service mathematics teachers in summer seminars—if only it were a national priority to fund such experiences, as it was in the 25 years following Sputnik in 1957.

Holt, Jim, *When Einstein Walked with Gödel: Excursions to the Edge of Thought*, Farrar, Straus and Giroux, 2018; xi + 369 pp, $28. ISBN 978-0-374-14670-2.

Beware: The title of this thoroughly enjoyable book is taken from the title of just one of its dozens of essays, though Einstein and Gödel do appear here and there in other essays. The essays appeared over the past 20 years in the *New York Review of Books*, *The New Yorker*, *The New York Times Book Review*, and elsewhere; they are gently instructive in an entertaining way about ideas in mathematics (Gödel incompleteness, the Riemann zeta function, Cantor's infinities, prime numbers, category theory, and fractals) and the foibles and fame of their originators.

Carl B. Allendoerfer Awards

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of articles of expository excellence published in MATHEMATICS MAGAZINE. The Award is named for Carl B. Allendoerfer, a distinguished mathematician at the University of Washington and president of the Mathematical Association of America, 1959–1960.

Fumiko Futamura and Robert Lehr "A New Perspective on Finding the Viewpoint." MATHEMATICS MAGAZINE 90:4, October 2017, 267–277.

Over the centuries, great painters have mastered perspective: the sleight of hand of conveying a three-dimensional visual experience in just two dimensions. But to fully enter into the illusion, a viewer needs to be standing where the artist stood; otherwise, the spectator will see "every false relation and disagreement of proportion that can be imagined. . . " (in the words of Leonardo da Vinci quoted in this article). So how, using only the painted image, do we find the exact spot where a viewer should stand?

The authors begin by taking us on a tour of the rich history of solutions to this problem. They do this through a cleverly chosen case study: Hendrick van Vliet's *Interior of the Oude Kerk, Delft,* 1660. This vivid rendering of the oldest church in Amsterdam conveys a sense of the spaciousness of a highly geometric interior, and does this through two-point perspective. The authors discuss a standard geometric method for finding the correct viewpoint that involves finding two pairs of vanishing points and taking the intersection of two appropriately drawn semi-circles passing through the vanishing points. When applied to the *Oude Kerk*, this yields the surprising conclusion that one should stand almost at the left edge of the painting! (This method—as do all the others discussed in the article—also gives the height at which one's eye should be, and the distance from the canvas.) Later in the article, the authors write that from this viewpoint, "... you see the interior from the height of an average person standing in the church ... the arches soar overhead and you can feel the spaciousness of the old church. The effect is magical."

However, the reader quickly sees that the standard geometric method, when applied to the *Oude Kerk*, involves drawing a very large semicircle, which would be impractical in all but the largest museum galleries. The authors then discuss refinements of the standard method—still geometric in nature—due to Johann Heinrich Lambert and Brook Taylor (the latter of Taylor Series fame). These methods are more practical, but introduce new drawbacks. The authors then skip ahead two centuries and describe an algebraic method due to R. Greene from 1983. They provide a new derivation of Greene's method using the cross ratio, which has the added benefit of leading to simplifications of Greene's original formulas. Then comes the coup de grace: a new method for solving this age-old problem, which the authors call the "perspective slope method." It combines algebraic and geometric methods and is superior to the other methods in certain circumstances (e.g., a two-point perspective painting with a tiled floor).

The applied nature of the question at the heart of this article is appealing. The authors close by saying, "Our great hope is that you will share this article with your local museums or at the very least feel empowered to use these techniques yourself on digital images to determine viewpoints prior to a museum visit." The article offers the reader a fascinating mix of geometry and algebra, a lesson in history, and a renewed knowledge that mathematics has for centuries played a critical role in problems of interest to artists and art-lovers alike.

Math. Mag. **91** (2018) 319–320. doi:10.1080/0025570X.2018.1489630

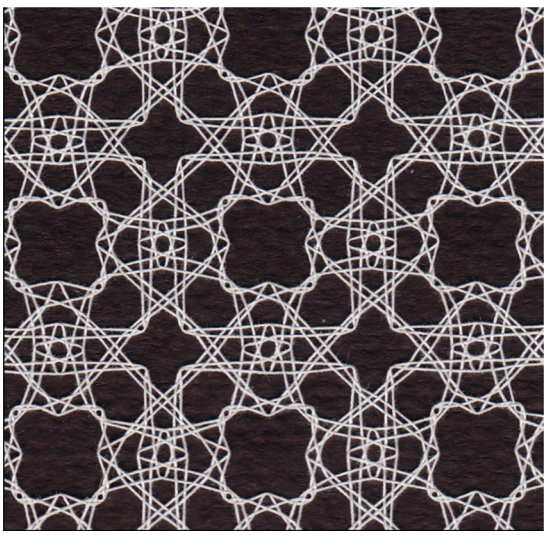
Response from the authors

This is such a pleasant surprise! We are thrilled and honored to receive the Allendoerfer Award this year. To give a little glimpse into the background of the paper, Robert's original intention was to create fractals out of harmonic sets, for his project in Futamura's Geometry course. In making a number of elaborate images, tweaking them and turning them upside down, he realized that he may have stumbled upon a new way of finding the viewpoint for a two-point perspective image. Together, we solidified and refined the idea, delved into the fascinating history of finding the viewpoint for perspective images and went through over a dozen drafts trying to find the best way to convey these ideas in a fun, informative, and accessible way. We truly hope that readers will take this with them on their next museum trip to gain a more immersive experience!

DR. FUMIKO FUTAMURA is a Professor of Mathematics at Southwestern University, a liberal arts college in Georgetown, TX. As an artist who enjoys working in a number of different media, she did not consider majoring in mathematics until she realized that mathematics could be considered an art form. She decided to get her Ph.D. in mathematics to be able to work with this medium at a more advanced level. She now gets to teach art in many of her math classes as well as do research in math and art, in particular, on projective geometry and its applications to perspective drawing. She is currently working with coauthors on an inquiry-based textbook on the subject.

ROBERT LEHR received his BA from Southwestern University and worked for the University of Texas in Austin as a research assistant analyzing transportation behavior before continuing his education at the University of Texas School of Architecture where he is pursuing an MSSD. He wishes to incorporate his mathematical knowledge and familiarity with geometry into novel, practical architectural constructs.

Artist Spotlight: Veronika Irvine



Easter(n), Veronika Irvine; White cotton thread, 2015. The *442 orbifold symmetry in this pattern was created using four- and two-fold mirror generators. The underlying directed graph, embedded on a torus, has 52 nodes, that is, 52 braids per repeat.

See interview on page 307–309.