AUTHORS

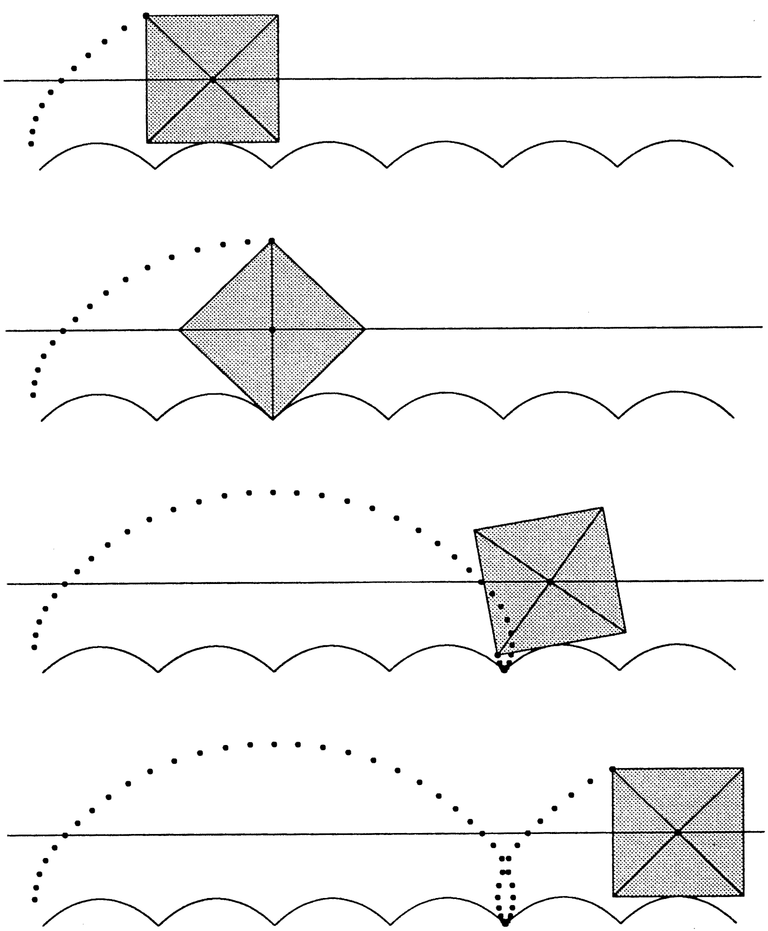
TITLES

PROBLEMS

P, S, Q, *and* C *refers to Proposals, Solutions, Quickies, and Comments, respectively. Page numbers are given in parantheses. Thus,* P1338(56) *refers to proposal number* 1338, *which appears on page 56.*

Vol. 65 No. 5, December 1992

MATHEMATICS MAGAZINE



- Roads and Wheels
- Euler's Constant, Taylor's Formula, Slowly Converging Series
- On Computing Euler's Constant

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

The aim of *Mathematics Magazine* is to provide lively and appealing mathematical exposition. This is not a research journal and, in general, the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for an article for the *Magazine*. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Articles on pedagogy alone, unaccompanied by interesting mathematics, are not suitable. Neither are articles consisting mainly of computer programs unless these are essential to the presentation of some good mathematics. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

The full statement of editorial policy appears in this *Magazine*, Vol. 64, pp. 71–72, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, nor published by another journal or publisher.

Send new manuscripts to: Martha Siegel, Editor, Mathematics Magazine, Towson State University, Towson, MD 21204. Manuscripts should be typewritten and double spaced and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit the original and two copies and keep one copy. In addition, authors should supply the full five-symbol Mathematics Subject Classification number, as described in *Mathematical Reviews*, 1980 and later. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added. Do not use staples.

AUTHORS

Leon Hall received his Ph.D. from the University of Missouri-Rolla in 1974, specializing in differential equations. He spent 11 years at the University of Nebraska before returning to UMR in 1985. As a result of his involvement in a departmental project using *Mathematica* in calculus, he attended the *Mathematica* Conference in San Francisco, where he met Stan Wagon and where they got rolling on this paper. Besides mathematics, he enjoys playing racquetball, hiking, and bicycling.

Stan Wagon's graduate training was in logic (Dartmouth Ph.D.), but he loves interesting mathematics wherever he can find it. His recent work has centered on how *Mathematica* can be used to enhance teaching and research. Stan is currently working on Animating Calculus (with Ed Packel; W. H. Freeman), a book that contains many applications of *Mathematica* to the teaching of calculus. When he can break himself away from the computer screen, Stan loves outdoor activities, such as 100-mile trail runs and two-week wilderness ski traverses.

John Baxley received his Ph.D. in analysis at the University of Wisconsin-Madison in 1966. Hooked on infinite series, as a source of mathematical problems and as a challenge in teaching elementary calculus, he began experimenting in the 1970s with computing approximations of sums of infinite series as a way of making the subject less abstract for students. John's research for the last 20 years has been devoted almost completely to the existence of solutions of boundary value problems for differential equations. He has been at Wake Forest University since 1968.

Jeffrey Nunemacher received his B.A. from Oberlin College and a Ph.D. in the area of several complex variables from Yale University in 1975. Since then, he has taught at the University of Texas at Austin, Kenyon College, Oberlin College, and now Ohio Wesleyan University, where he is professor and chair of the Department of Mathematical Sciences. The present article arose from a class project in a numerical analysis course to investigate how many different analytic techniques could be used to approximate Euler's constant.

Vol. 65 No. 5, December 1992

MATHEMATICS MAGAZINE

The *MATHEMATICS MAGAZINE* (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, bimonthly except July/August.
The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $16 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $64. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 40% dues discount for the first two years of membership.) The nonmember/library subscription price is $68 per year.
Subscription correspondence and notice of change of address should be sent to the Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Advertising correspondence should be addressed to Ms. Elaine Pedreira, Advertising Manager, The Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

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

Postmaster: Send address changes to Mathematics Magazine Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036-1385.

PRINTED IN THE UNITED STATES OF AMERICA

ARTICLES

Roads and Wheels

LEON HALL
University of Missouri-Rolla
Rolla, MO 65401

STAN WAGON
Macalester College
Saint Paul, MN 55105

Introduction

San Francisco's Exploratorium contains an intriguing exhibit of a square wheel that rolls smoothly on a road made up of linked, inverted catenaries (see FIGURE 4). That exhibit inspired us to generate a computer animation of a rolling square and further explore the relationship between the shapes of wheels and roads on which they roll. In a sense, we are bringing up to date the paper by G. Robison [4], showing how much more can be done, both numerically and graphically, with modern computer hardware and software. The problem of the square wheel has been rediscovered and solved several times; see [**5**, **7**].

All the diagrams and animations were prepared in *Mathematica*. Our package that generated the diagrams and the associated animations (see Section 5) can be obtained by sending a Macintosh disk to one of the authors. It is noteworthy that some of the results of this paper, in particular the discovery of a cycloidal locus generated by a noncircular wheel, were discovered only after viewing certain graphics. *Mathematica* was also used to do all the symbolic integrations that occur. For further applications of symbolic and graphic computation to wheel/road problems, in particular, a complete discussion of the cycloid, see [**6**, Chapter 2].

The paper is organized as follows. Section 1 discusses the theory and the fundamental differential equation. Section 2 contains many closed-form examples. Section 3 shows how numerically approximating the solution to the differential equation is an excellent approach to diverse examples, even those solvable in closed form. Section 4 squares the circle by considering Fourier approximations to the catenary. And Section 5 discusses the *Mathematica* package that we built.

1. Building a wheel

Suppose we are given a road in the form of a rectifiable curve in the lower half-plane parametrized by $f(t) = (x(t), y(t))$, where $x(t)$ is increasing, $x(0) = 0$, $y(t) \leq 0$. By the *wheel* corresponding to the road we mean a curve that will roll smoothly on the road. More precisely, a wheel will be a curve given by a polar function $r = r(\theta)$ such that the axle of the wheel, which initially is at $(0, 0)$, stays on the x-axis directly above the wheel–road contact point as the curve rolls along the road. The wheel's axle may or may not coincide with the wheel's geometric center. The road is assumed to provide enough friction so that there is never any slipping of the wheel. The rolling motion can be described by a function $\theta = \theta(t)$ that describes the amount of angular rotation for the wheel to roll from $f(0)$ to $f(t)$. These functions must satisfy the following conditions (see FIGURE 1):

1. *Initial condition.* The initial contact point is at $f(0)$, directly under the origin, whence $\theta(0) = -\pi/2$.
2. *Rolling condition.* The amount unravelled on the wheel matches the distance travelled on the road: For any t, the arc length of f between $f(0)$ and $f(t)$ equals the arc length of the polar curve between $\theta(0)$ and $\theta(t)$.
3. *Radius condition.* The radius of the wheel matches the depth of the road at the corresponding point: For any $t, r(\theta(t)) = -y(t)$.

FIGURE 1 illustrates the formation of the wheel in the case when the road is given as $y = f(x)$, with $f(x)$ nonpositive, in which case the conditions simplify accordingly (that is, x can be used as the parameter, and so θ becomes a function of x).

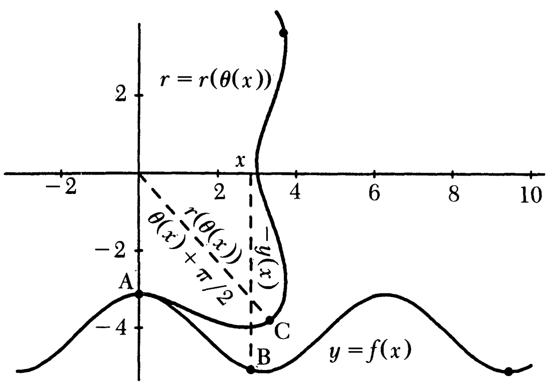


FIGURE 1

If a road is given by $y = f(x)$, then the relationship between θ and x is obtained from the equality of the arc lengths AB and AC and of the radius vector OC and the depth of the road (dashed lines). The road illustrated is given by $y = -\sqrt{17} + \cos x$, where $\sqrt{17}$ has been chosen so that the wheel closes up on itself (see Remark 5).

The key to getting a wheel is finding the function $\theta(t)$, since the radius condition will then yield $r(\theta)$. The first two conditions become a simple differential equation, which can lead to either a closed-form description of θ or a numerical approximation. The rolling condition is:

$$\int_0^t \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt = \int_{-\pi/2}^{\theta(t)} \sqrt{r(\theta)^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta.$$

Differentiating both sides with respect to t and squaring yields:

$$\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 = \left(\frac{d\theta}{dt}\right)^2 \left(r(\theta)^2 + \left(\frac{dr}{d\theta}\right)^2\right).$$

Now substitute $\dfrac{dr}{d\theta}\dfrac{d\theta}{dt} = -\dfrac{dy}{dt}$ (obtained by differentiating the radius condition) to get:

$$\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 = r(\theta)^2 \left(\frac{d\theta}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2,$$

which simplifies to:

$$\frac{d\theta}{dt} = \pm \frac{dx}{dt} \frac{1}{-y(t)}.$$

Because $d\theta/dt$ is to be positive, the differential equation we seek is:

$$\frac{d\theta}{dt} = -\frac{dx}{dt} \frac{1}{y(t)},$$

with initial condition $\theta(0) = -\pi/2$.

Remarks.

1. If the road is given by $y = f(x)$, the differential equation for θ becomes simply $d\theta/dx = -1/f(x)$. In this case

$$\theta(x) = \int_0^x \frac{-1}{f(x)}\, dx - \frac{\pi}{2}.$$

2. If the function $\theta(t)$ can be inverted to $t(\theta)$ then the wheel is given by the polar equation $r = -y(t(\theta))$.
3. An alternative approach to characterizing $\theta(t)$ proceeds by matching slopes instead of arc lengths. The rolling condition then becomes: For any t, the slope of the road at $f(t)$ equals the slope of the tangent to the polar curve at $\theta(t)$, rotated clockwise through $\theta(t) + \pi/2$ radians. This leads to the same differential equation.
4. The inverse problem starts with $r(\theta)$, a polar representation of a wheel, and seeks the appropriate road. The preceding discussion implies that the road is given by $y(x) = -r(\theta(x))$, where θ satisfies $\theta(0) = -\pi/2$ and $d\theta/dx = 1/r(\theta)$. One can also deal with the case that the wheel is given parametrically by $(x(t), y(t))$; see Case 4 in Section 3.
5. Suppose the road $y = f(x)$ is periodic with period a. Then the corresponding wheel does not necessarily close up on itself to form a topological disk (see FIGURE 2). The condition for such closure—the *closed-wheel condition*—is that there exists a rational number r so that $2\pi r = \theta(a) - \theta(0) = \int_0^a -1/f(x)\, dx$. If $r = 1/n$, n a positive integer, then the wheel rolls over n periods of the road during each complete revolution. As an example, consider the road given by $y = d + \cos x$,

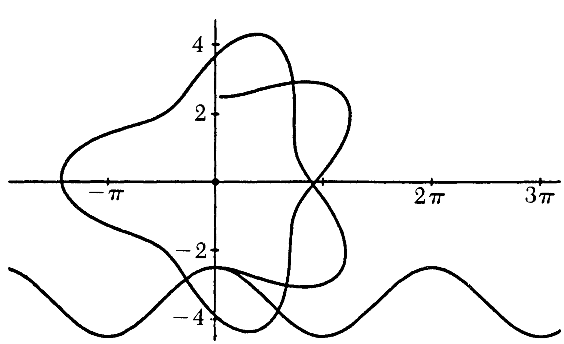


FIGURE 2
The wheel for a cosine road given by $-3.5 + \cos x$ winds around endlessly without closing up, because 3.5 is not one of the special values $\sqrt{1+n^2}$. This wheel was generated by the numerical technique discussed in Section 3.

where $d \leq -1$. For each positive integer n, there is a unique value of d, which turns out to be $-\sqrt{1+n^2}$, for which the wheel closes up into one that covers n periods per revolution. (See FIGURE 1 for the $n=4$ case and FIGURE 13 for the $n=1$ case.) To see that the values of d in the cosine case are as claimed, observe that $\int_0^{2\pi} -1/(d+\cos x)\,dx = 2\pi/\sqrt{d^2-1}$. It is easiest to integrate from 0 to π and then double. More generally, the road $y = d + b\cos(cx)$ yields a closed wheel that covers n periods when $d = -\sqrt{b^2+n^2/c^2}$. Cycloid roads and inverted cycloid roads provide two more examples in which the closed-wheel condition integral can be evaluated (see FIGURE 9(c) for the $n=0$ case, and FIGURE 19(d) for a fractional example, viz., $n = 1/2$). The closed-wheel condition has a completely analogous form in the case of a parametrically defined road as well; see the cycloid example in Case 3 of Section 3.

2. Closed-form solutions

This section discusses several examples for which the differential equation can be solved in closed form. It is remarkable that, although the arc length of familiar curves is generally not solvable in closed form, the wheel–road problem is solvable for a wide variety of functions.

Polygonal wheels The roads corresponding to polygonal wheels are derived from the case of a wheel that is nothing more than a straight line. Consider the polar equation $r = -\csc\theta$, $-\pi < \theta < 0$, whose graph is a horizontal line one unit below the x-axis. The results of Section 1 show that the road on which this polar line will roll is given by $y = f(x) = -r(\theta(x))$ where $\theta(0) = -\pi/2$ and

$$\frac{d\theta}{dx} = \frac{1}{-y(x)} = \frac{1}{r(\theta(x))} = -\sin\theta.$$

The solution to this initial-value problem is $x = -\log(-\tan(\theta/2))$, or $\theta = -2\arctan e^{-x}$. Hence the road is given by:

$$\begin{aligned} y = f(x) &= \csc(-2\arctan e^{-x}) \\ &= \frac{-1}{\sin(2\arctan e^{-x})} = \frac{-1}{2\sin(\arctan e^{-x})\cos(\arctan e^{-x})} \\ &= -\frac{1+e^{-2x}}{2e^{-x}} = -\frac{e^x+e^{-x}}{2} = -\cosh x, \end{aligned}$$

whose graph is an inverted catenary. This means that the polar line will roll on the catenary so that the polar origin, which we imagine as attached to the line, stays on the x-axis.

Modifying the straight-line example yields roads for wheels that are regular polygons. Consider the square wheel. By simply truncating the catenary where its slope is ± 1—this occurs at $x = \pm\operatorname{arcsinh} 1$—and forming a periodic road by translating copies of the truncated catenary, the angle at the junctions will be 90° Hence a square will smoothly pass over the junction. FIGURE 4 shows several images from an animation of a rolling square on such a road, along with the locus of a vertex, which is related to the involute of the catenary.

The road appropriate for a regular n-gon may also be obtained from the catenary $y = -\cosh x$. If the catenary is truncated at $x = \pm\operatorname{arcsinh}[\tan(\pi/n)]$ then the angle at the road's cusp matches the interior angle of the n-gon, and the amount of rotation to get the wheel into the cusp is $\theta(x) + \pi/2$. This works out to be exactly $2\pi/n$ (the details, which involve a horrendous-looking identity involving tan, arcsinh, and arctan, are left to the reader). So the wheel corresponding to the road made of pieces of inverted catenaries closes up exactly into a regular n-gon. The case of a triangle is noteworthy in that the rolling cannot happen physically: Because the cusp angle is less than 90°, the triangle will crash into the road before the vertex gets into the cusp (FIGURE 5).

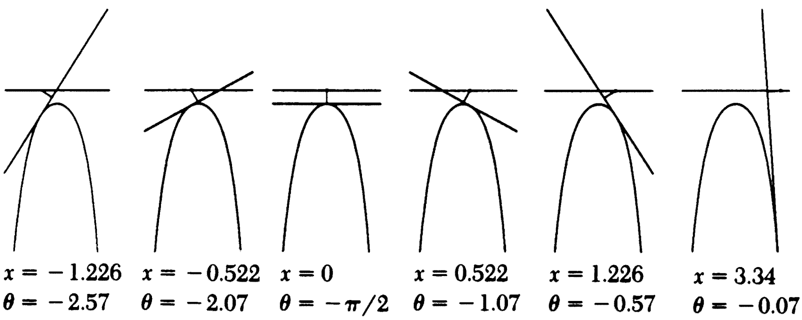


FIGURE 3

Some stills from an animation showing the polar line $r = -\csc\theta$, $-\pi < \theta < 0$, rolling over the catenary $y = -\cosh x$. The x-values and θ-values correspond to the x-coordinate of the point of tangency and the θ-value of the point of tangency viewed as a point on the polar line. Adding $\pi/2$ to the θ-value yields the amount of rotation of the horizontal line.

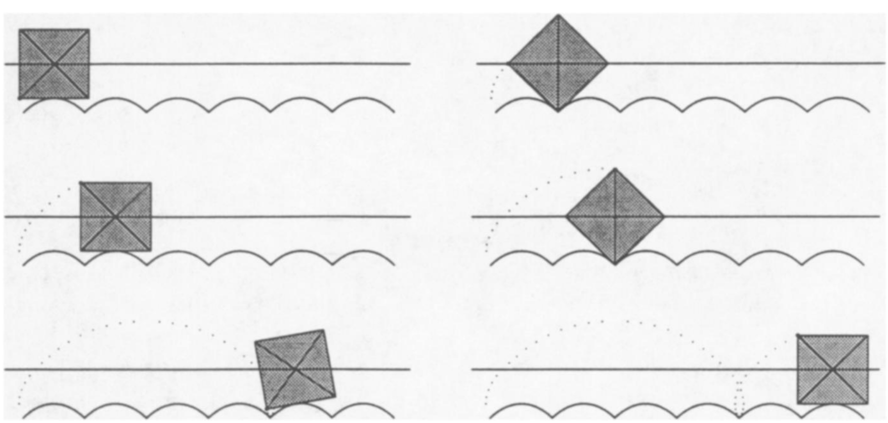


FIGURE 4

A road made up of pieces of an inverted catenary allows a square to roll smoothly. The dots are the locus of a vertex during the rolling. Note that the slope of the tangent to the locus has a discontinuity at the cusp.

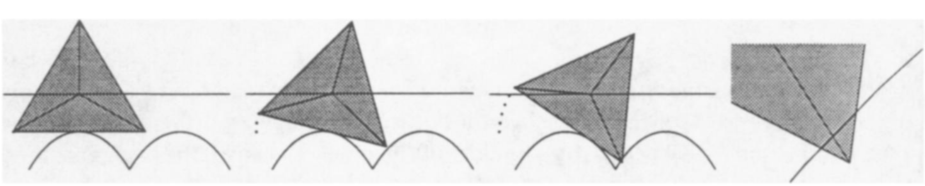
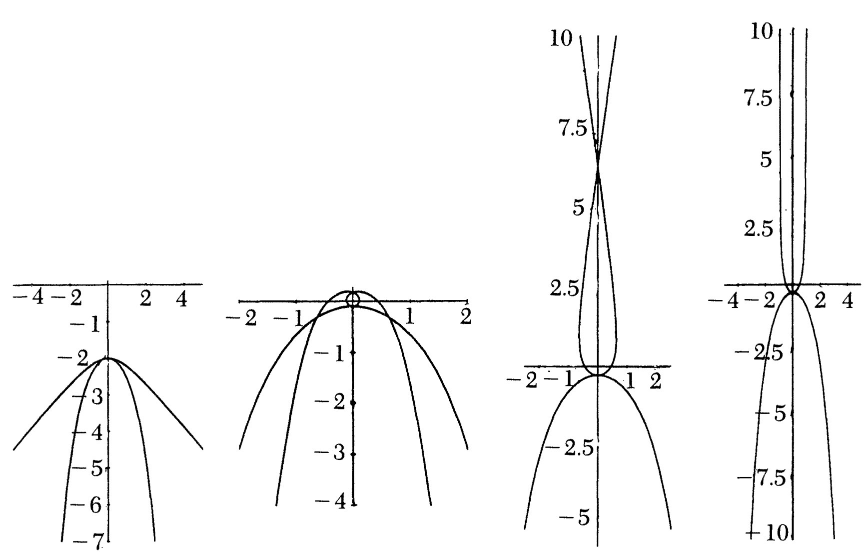


FIGURE 5

A vertex of a rolling triangle crashes into the road just before the vertex arrives at the cusp. The rightmost diagram is a close-up of the collision.

As we shall see several times in this paper, the depth of a road plays a crucial role in determining the shape of the wheel. If the catenary road $y = -\cosh x$ is raised or lowered, the shape of the wheel changes. Consider the family of roads $k - \cosh x$ where $k < 1$. If $k = 0$, then the wheel is a straight line, but other values of k yield radically different wheels, as shown in FIGURE 6. The closed-form solution, obtained with the help of *Mathematica*'s integrator, is given by $\theta(x) = \phi(x) - \phi(0) + \pi/2$ where

$$\phi(x) = \frac{-2\arctan\left(\dfrac{e^{-x} - k}{\sqrt{1-k^2}}\right)}{\sqrt{1-k^2}}.$$



roads: $y = -1 - \cosh x$ $\quad$ $0.9 - \cosh x$ $\quad$ $0.7 - \cosh x$ $\quad$ $0.676454 - \cosh x$

FIGURE 6

Raising or lowering a catenary road via $k - \cosh x$ leads to a variety of wheel shapes. Only the choice $k = 0$ leads to a straight-line wheel. The wheel in the last case has vertical asymptotes.

Tilted roads Let $f(x) = -1 - x, x \geq 0$, define a downward-sloping road. The solution to the initial-value problem for θ is then $\theta(x) = \log(x+1) - \pi/2$. Therefore the wheel for this oblique line is the equiangular spiral $r(\theta) = 1 + \exp(\theta + \pi/2) - 1 = \exp(\theta + \pi/2)$, which is shown in FIGURE 7. In the general case that the road is given by $-1 - mx$, the wheel is the spiral given by $r = \exp[(\theta + \pi/2)/m]$.

As with the catenary, we can turn the 45° road into a periodic function, in this case a sawtooth. In order to get a smoothly rolling wheel that covers four teeth of the road in one revolution, we need to find x such that $\theta(x) = -\pi/4$: $\theta(x) = -\pi/4$ if and only if $\log(x+1) - \pi/2 = -\pi/4$ if and only if $x = e^{\pi/4} - 1 = 1.19\ldots$. Now we can cut off the straight line at this point, generate a sawtooth road, and paste appropriately truncated pieces of the spiral together at right angles to get the sawtooth wheel shown in FIGURE 8. The pasting yields 90° angles because the tangent to the equiangular spiral makes a constant 45° angle with the radius vector; these angles plus their mirror images yield the right angles. A wheel that covers more teeth during each revolution can be obtained by truncating the line at a value of x that satisfies $\theta(x) = -\pi/n$ and pasting together more and shorter pieces of the spiral.

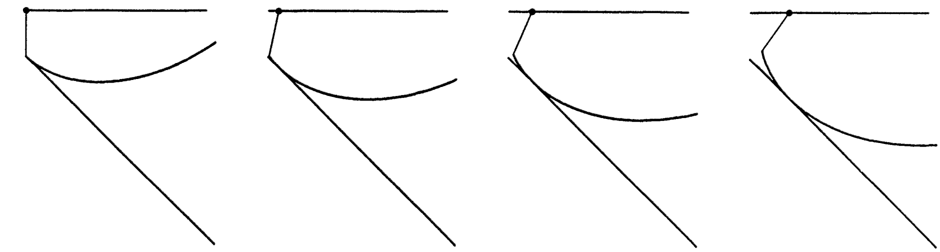


FIGURE 7
An exponential, or equiangular, spiral rolling along a tilted line.

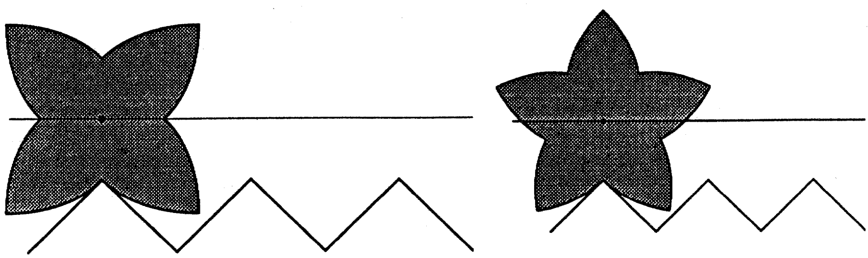


FIGURE 8
Pieces of an equiangular spiral can be pasted together to get a wheel that rolls on a sawtooth road. The examples shown cover four and five teeth per revolution, respectively.

Cycloidal roads The cycloid is the famous curve that is the path of a point on a traditional round wheel rolling on a straight road; inverting the cycloid leads to the parametric curve $f(t) = (t - \sin t, \cos t - 1)$. The wheel that rolls on an inverted cycloid can be found by solving the differential equation $d\theta/dt = (-1 + \cos t)/(\cos t - 1) = +1$, so $\theta(t)$ is simply $t - \pi/2$. Hence the wheel is given in polar form by $r = -y(t(\theta)) = 1 - \cos(\theta + \pi/2) = 1 + \sin\theta$, the polar form of a cardioid (FIGURE 9(a)). The cusp of the cardioid rolls over the cusp of the cycloid, at least in theory. In practice, there is a crash between the road and the cardioid, similar to the one that happens with a rolling triangle (FIGURE 5). As pointed out by Robison, a physical model can be built so as to avoid the cusp problem by introducing pieces of a catenary into the road and a straight line segment into the wheel so as to bypass the cusps. See [4, FIGURE 4] for details.

The locus of the top point of the cardioid seems to have the same general shape as a cycloid. As an exercise, the reader can verify that, indeed, this locus is a cycloid stretched vertically by a factor of 2. As further exercise, the reader can investigate the clover-like wheels that arise from lowering the cycloid so that the closed-wheel condition is met. *The case of* $(t - \sin t, -13/5 - \cos t)$ *is illustrated in* FIGURE *9(b).* One can also consider the uninverted cycloid that is tangent to the x-axis from below — $(t - \sin t, -1 - \cos t)$ — for which the wheel is derived from the function:

$$\theta(x) = -x + \frac{2\sin x}{1 + \cos x} - \frac{\pi}{2}.$$

This leads to a spiral wheel that requires infinitely many revolutions to pass over the cycloid's high point (FIGURE 9(c)). Cycloidal roads are discussed further in Case 3 of Section 3.

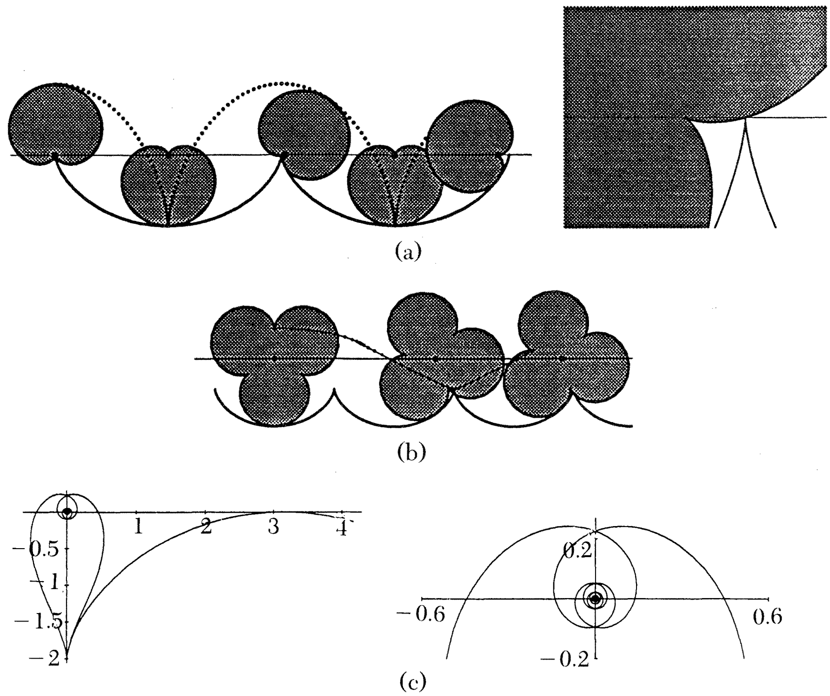


FIGURE 9
(a). A cardioid rolls on an inverted cycloid, and the locus of the point opposite the cardiod's cusp is a vertically scaled cycloid. The close-up view shows that, in actuality, a cardioid-shaped tire would be punctured by the cycloid's cusp. (b) Lowering the cycloid leads to the clover-shaped wheels. (c) A right-side-up cycloid touching the x-axis yields a wheel that takes infinitely many revolutions to pass over the cycloid's high point.

It turns out to be worth considering more general cycloidal roads, such as trochoidal roads represented by $(t - a \sin t, -1 + a \cos t)$. The usual computations show that the wheel is given by $r = 1 + a \sin \theta$, a limaçon that, when $a = 1$, is the just discussed cardioid. This limaçon intersects the positive y-axis at $(0, a \pm 1)$, points whose loci during the rolling are $(t \pm \sin t, (a \pm 1)\cos t)$. These loci are cycloids stretched in the y-direction. Note that there are two cases in which the stretching constant is 1—that is, the locus is an exact cycloid: The classical case $a = 0$, with

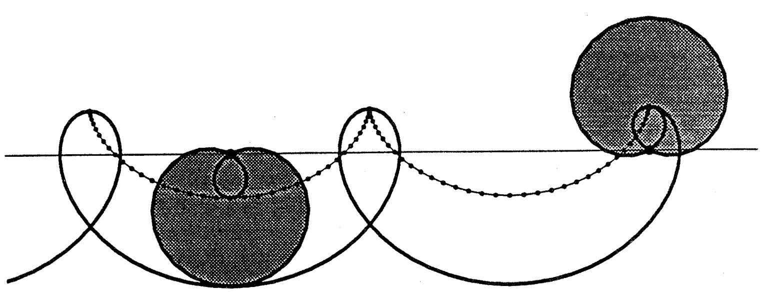
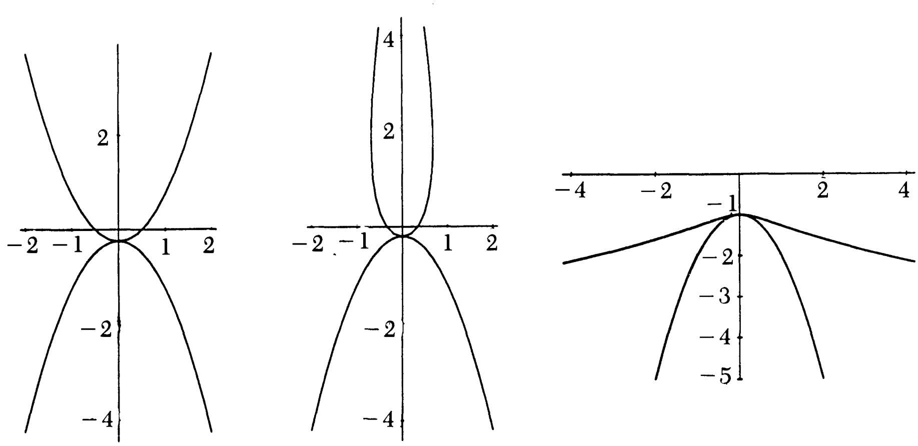


FIGURE 10
Rolling a limaçon along a trochoid yields an exact cycloidal locus.

straight road and round wheel, and the case when $a = 2$ (or -2), where the road is a trochoid, the wheel is a limaçon, and the locus of the top of the inner loop is an exact inverted cycloid (FIGURE 10). Are these the only two examples of road-wheel combinations for which a point on the wheel traces out an exact cycloid?

A road that is its own wheel Can there be a road for which the corresponding wheel is congruent to the road and for which the rolling motion matches points that correspond under the congruence? Consider the road that is given by the parabola $y = -x^2 - 1/4$. The differential equation yielding $\theta(x)$ is then $d\theta/dx = 1/(x^2 + 1/4)$, for which the solution is $\theta(x) = 2\arctan(2x) - \pi/2$, or $x = 1/2\tan(\pi/4 + \theta/2)$. The polar wheel then has the form $r(\theta) = -y(x(\theta))$, which simplifies to $r = 1/(2 - 2\sin\theta)$. This last is the graph of the parabola $y = x^2 - 1/4$, which is the reflection of the road in the line $y = -1/4$. FIGURE 11 shows this singular situation of a wheel rolling on itself. We leave the verification that corresponding points touch as an exercise. Robison [4] showed that this parabola is the only curve that has this property.

Note that raising or lowering the parabola changes the shape of the wheel dramatically (see FIGURE 11).



road: $y = -x^2 - \frac{1}{4}$; wheel: $y = x^2 - \frac{1}{4}$ road: $y = -x^2 - \frac{1}{5}$ road: $y = -x^2 - 1$

FIGURE 11

The wheel corresponding to the parabolic road given by $-x^2 - \frac{1}{4}$ is simply a reflection of the road itself. But the wheels for other parabolic roads are not parabolic.

Round wheels can roll on round roads A well-known puzzle can be interpreted as follows: What is the wheel corresponding to a road that is an upward-opening semicircle whose highest points are on the x-axis? Such a road is given by $f(t) = (\cos t, \sin t)$, $\pi \le t \le 2\pi$. The differential equation for $\theta(t)$ is simply $d\theta/dt = 1$, so $\theta(t) = t - \pi/2$ and $r(\theta) = -\sin(\theta + \pi/2) = -\cos\theta$, $\pi/2 \le \theta \le 3\pi/2$. Thus the wheel is a polar circle with geometrical center at $(-1/2, 0)$. The aforementioned puzzle is the one that asks for the locus of a point on the circumference of a circle that is rolling in a way tangent to the interior of a circle twice as large. The locus is a straight line, which shows itself in FIGURE 12 as the vertical lines in the arch-shaped locus.

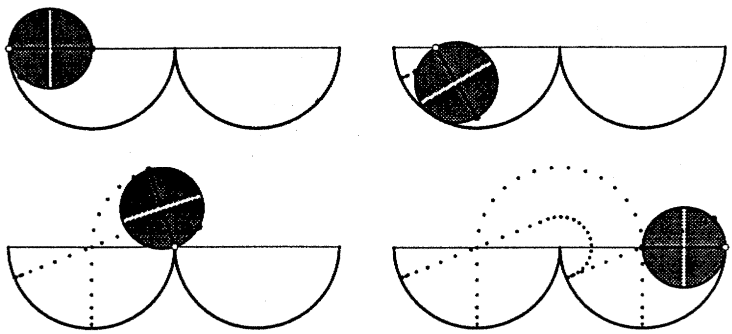


FIGURE 12
A circle rolls on inverted semicircle. The dotted paths are the loci of two points in the wheel's circumference.

Off-centered elliptical wheels Consider the ellipse given in polar form by $r = k\varepsilon/(1 - \varepsilon \sin\theta)$, where $0 < \varepsilon < 1$ and $k > 0$; the origin, which corresponds to the axle of the wheel as it rolls, is one focus of the ellipse, the other focus is on the positive y-axis, ε is the eccentricity, and k is the distance from the origin to the corresponding directrix. Such an elliptical wheel rolls on the road $y = -(k\varepsilon/a^2)(1 - \varepsilon\cos(cx))$, where $c = a/k\varepsilon$ and $a = \sqrt{1-\varepsilon^2}$. The derivation of this is slightly complex. Here's a sketch:

1. Solve the initial-value problem to get

$$\frac{ax}{2k\varepsilon} = \arctan\left(\frac{\tan(\theta/2) - \varepsilon}{a}\right) + \arctan\left(\frac{1+\varepsilon}{a}\right).$$

2. Take the tangent of the relationship in (1) and use some trig formulas to get

$$\frac{1+\varepsilon}{1-\varepsilon}\,\frac{1-\cos^2(cx)}{(1+\cos(cx))^2} = \frac{1+\sin\theta}{1-\sin\theta}.$$

3. Solve the preceding for $\sin\theta$, substitute into

$$y(x) = \frac{-k\varepsilon}{1-\varepsilon\sin\theta(x)},$$

and simplify to get the desired representation of the road as $(k\varepsilon/a^2)(1 - \varepsilon\cos(cx))$.

If we set $k = 1$ and $\varepsilon = 1/\sqrt{2}$ then the road is just $y = -\sqrt{2} + \cos x$ (see FIGURE 13); this is a special case of one of the proper depths to lower the cosine so as to get a closed wheel (see Remark 5 in §1).

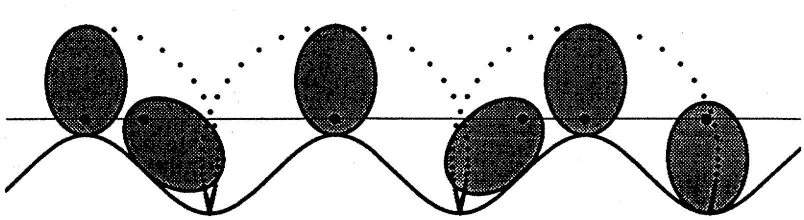


FIGURE 13
An ellipse rolls on a cosine curve. In the example shown the road is given by $y = -\sqrt{2} + \cos x$ and the ellipse has the polar form $r = 1/(\sqrt{2} - \sin\theta)$.

Centered elliptical wheels The preceding example involved an off-center elliptical wheel; that is, the axle is not at the ellipse's center. Let's now find the road on which an ellipse with centered axle will roll. The ellipse $(x/a)^2 + (y/b)^2 = 1$ has the polar representation $r = b/\sqrt{1 - m\cos^2\theta}$, where m abbreviates $1 - b^2/a^2$; in this representation the polar center—the axle of the wheel—coincides with the center of the ellipse. We can find the appropriate road by first finding the relationship between x and θ; the differential equation for θ separates to $\sqrt{1 - m\cos^2\theta}\, d\theta = (1/b)\,dx$ and the initial condition then leads to:

$$\int_{-\pi/2}^{\theta(x)} \frac{d\phi}{\sqrt{1 - m\cos^2\phi}} = \int_0^x \frac{dx}{b}.$$

Substituting $\psi = \phi + \pi/2$ yields $\sin\psi = \cos\theta$ and

$$\int_0^{\theta(x)+\pi/2} \frac{d\psi}{\sqrt{1 - m\sin^2\psi}} = \frac{x}{b}.$$

Now this is an incomplete elliptic integral of the first kind [**1**, 17.2.2, 17.2.17]. Therefore $x/b = F(\theta + \pi/2|m)$, which can be inverted by using what is known as the Jacobian elliptic function sn: $\sin(\theta + \pi/2) = \mathrm{sn}(x/b, m)$. The road is therefore given by $y = -r(\theta(x)) = -b/\sqrt{1 - m\,\mathrm{sn}^2(x/b, m)}$. The sn function is built into *Mathematica*, and so it is a simple matter to generate a diagram of the road and wheel. But some complications arise when one tries to use the closed form to generate an animation; they can be worked around by adding multiples of $\pi/2$ as x increases through the quarter-period value. But life is much simpler if we use the numerical approach of the next section to generate the animation. It is easiest to start with the ellipse and use the Case 2 discussion, as that avoids the computation of values of sn. The result of one such computation is shown in FIGURE 14 ($a = 1/2$, $b = 1$, eccentricity $= 0.87$). Ellipses with eccentricities greater than about 0.97 lead to a road-wheel crash similar to the one for the triangle (FIGURE 5).

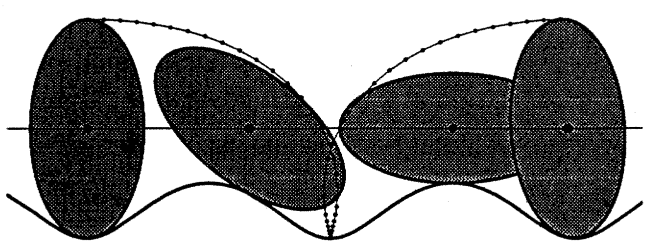


FIGURE 14
Four frames from an animation showing an elliptical wheel ($a = \frac{1}{2}, b = 1$) with its axle at its center rolling on a road defined using the elliptic sine function.

An elliptical road We can consider bounded roads as well as roads that protrude above the x-axis. As one example, consider the ellipse given parametrically by $(a\sin t, -b\cos t)$, where a and b are positive. The closed-wheel integral is $2\pi a/b$, and integer values of b/a lead to closed wheels that are familiar curves. The usual calculation shows that $t(\theta)$ is just $(b/a)(\theta + \pi/2)$, whence the wheel's polar form is $r(\theta) = -y(t) = b\cos[(b/a)(\theta + \pi/2)]$, a polar rose (FIGURE 15(a)). The locus here was a surprise to us as it turns out to be a piriform, which we had considered earlier in another context (§3, Case 4). If b/a is rational then the wheel is a rosette, as defined by Hall [**2**].

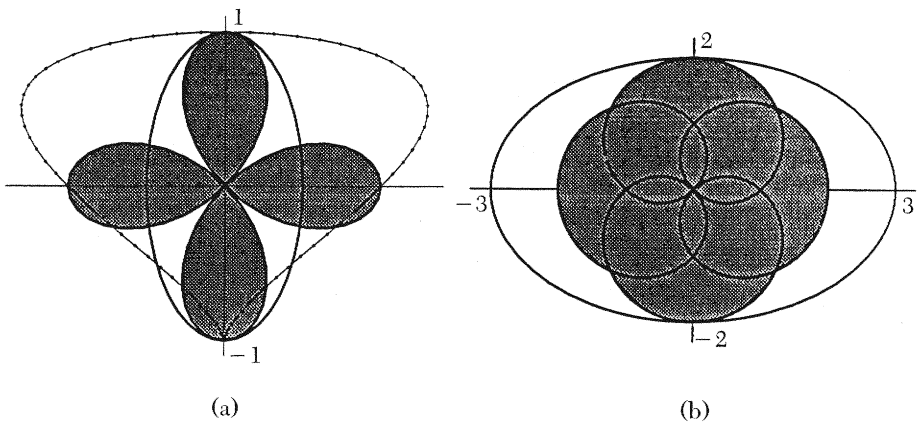


FIGURE 15
(a). A four-leafed rose rolls inside the ellipse $(\frac{1}{2}\sin t, -\cos t)$, and the locus of the tip of a petal is a piriform. (b) The wheels corresponding to more general ellipses [shown here is $(3\sin t, -2\cos t)$] are rosettes.

Vertical scaling We have seen in some of the preceding examples that raising or lowering the road usually changes the wheel significantly, and may destroy the closed-wheel condition. Another way to change the road is by scaling the y-coordinate: $y=f(x)$ becomes $y=kf(x)$, and $x=f_1(t)$, $y=f_2(t)$ becomes $x=f_1(t)$, $y=kf_2(t)$. The closed-wheel condition is affected as follows.

1. If the a-periodic road $y=f(x)$ has a closed wheel, then so does $y=kf(x)$ for any positive rational scaling factor k.
2. If the a-periodic road $y=f(x)$ does not have a closed wheel, the scaled road $y=kf(x)$ does have a closed wheel whenever k is a rational multiple of

$$\int_0^a \frac{-1}{2\pi f(x)}\,dx.$$

Similar results hold for roads defined parametrically.

As an example of scaling, consider the road $y=(\cos x)-b$, $b>1$. The scale factor $1/\sqrt{b^2-1}$ produces a closed wheel. Note that the scale factor goes to zero as b goes to infinity, and that the scaled road approaches $y=-1$. Since the wheels for these roads are ellipses, larger values of b correspond to wheels with smaller eccentricity.

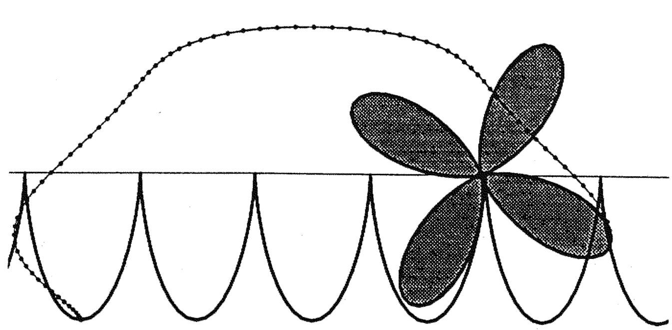


FIGURE 16
A scaled cycloidal road has a cuspitate rosette as its wheel.

Another case where scaling gives interesting road–wheel pairs is an inverted cycloid: $x = t + \sin t$, $y = k(-1 - \cos t)$. Assume k is a positive rational. Then the wheel is given by $r(\theta) = k(1 + \cos[k\theta + k\pi/2])$. For k a positive integer these wheels are cuspitate rosettes, curves similar in appearance to standard polar roses, but having cusps at their center. For more about cuspitate rosettes see **[2]** and the references therein.

Summary of some road–wheel relationships Here is a summary of road–wheel relationships including some not in closed form (12, 13) and some whose derivations are left as an exercise for the reader (14, 15, 16).

Wheel	Form of Road	Locus
1. Straight line	Catenary	
2. Regular polygons	Piecewise catenaries	
3. Circle	Horizontal line	Cycloid/Trochoid
4. Circle, axle on circumference	Circle, radius doubled	Circular arch
5. Equiangular spiral	Oblique line	
6. Piecewise equiangular spiral	Sawtooth	
7. Ellipse, axle at focus	Cosine	
8. Ellipse, axle at center	$1/\sqrt{1 - (\text{elliptic sine})^2}$	
9. Parabola $(x^2 - 1/4)$	Parabola $(-x^2 - 1/4)$	
10. Cardioid; see FIGURE 9(a)	Inverted cycloid	Scaled cycloid
11. Spirals, clovers; see FIGURE 9(b, c)	Lowered inverted cycloids	
12. Pointed wheels; see FIGURE 19	Lowered cycloids	
13. Piriform	See FIGURE 20	
14. Hippopede	See FIGURE 18	
15. Roses, rosettes	Ellipses	Piriform
16. Limaçon	Trochoid	Cycloid
17. Hyperbolic spiral $[r = 1/(\theta + \pi)]$	Exponential $[y = -(2/\pi)e^{-x}]$	
18. Cuspitate rosettes	Scaled cycloids	

3. Generating solutions numerically

As we demonstrated in the last section, surprisingly many road–wheel problems are solvable in closed form. When one deals with arc length, however, functions that are not integrable in terms of elementary functions eventually show up. In this section we describe how to generate plots of roads and wheels numerically. Our procedures will be described for *Mathematica*, but the methods, with the possible exception of the animations, could be adapted to other software with graphics capabilities. What is needed is the ability to apply a numerical differential equation algorithm and plot points and lines.

The key step, as we have seen, is to solve the initial-value problem relating the road parameter and the polar angle. The differential equation involved is separable, but the resulting integrations can be daunting, even when feasible. Furthermore, we have often found this closed-form approach too slow to generate plots of the roads and wheels. A more efficient method is to solve the initial-value problem numerically using a standard Runge–Kutta algorithm; then a set of points defining the wheel or road can be generated with the help of the radius condition. The points are joined by lines to get the final image. The relationship between the road parameter and the polar angle is almost always nonlinear (an exception: a circular wheel rolling on a straight road), which means that constant velocity and constant angular velocity cannot be achieved simultaneously. The use of a fixed step-size in the numerical method can guarantee that, in an animation, one of these velocities is constant.

When animating one of these wheels rolling along its road, the relation between the polar angle θ and the horizontal coordinate x must be known. The rolling can be broken down into two parts: a rotation, given by the change in θ, and a translation, given by the corresponding change in x.

There are four cases to consider. Some stripped-down code for Case 1, the simplest case, is given in Section 5.

Case 1. Suppose the road is given by $y = f(x)$ and we want to generate the wheel. The initial-value problem is $d\theta/dx = -1/f(x)$, $\theta(0) = -\pi/2$. The Runge–Kutta method generates a table of ordered pairs (x, θ). This table and the radius condition are used to produce the wheel.

Example. Let $y = -1.887365 - (2/3)\cos x + \sin x - (1/2)\sin 2x$, where the constant term has been chosen so that the closed-wheel condition holds for one revolution per period. The wheel (FIGURE 17) is generated by using a Runge–Kutta method to solve the initial-value problem, taking the independent variable to be x. Thus, when animated, this wheel's axle moves with constant linear velocity. This example shows that wheels need not have an axis of symmetry.

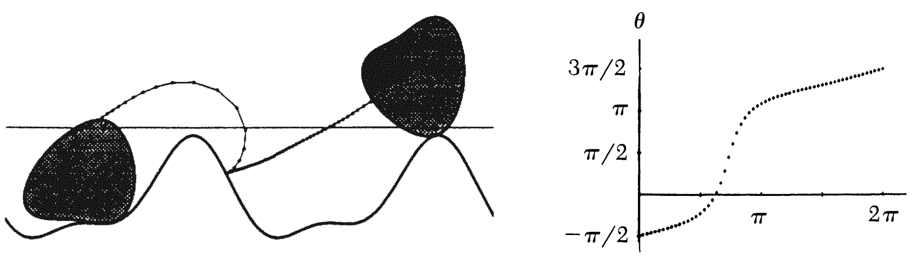


FIGURE 17

This example shows that the angular speed during rolling can vary a lot. The wheel rotates quickly when it is above the high bumps in the road, as illustrated by the steepness in the θ vs. x plot.

Case 2. Suppose the wheel is given by $r = g(\theta)$ and we want to generate the road. The initial-value problem is $dx/d\theta = g(\theta)$, $x(-\pi/2) = 0$. This time the numerical method gives pairs (θ, x), and the road can be generated using the radius condition.

Example. Let $r = 4\sqrt{5 - 4\sin^2\theta}$, a polar curve called a *hippopede* (see [3]). Again, Runge–Kutta is used to solve the initial-value problem, but this time it is convenient to take θ as the independent variable. Thus, when animated, the wheel (FIGURE 18) would exhibit constant angular velocity.

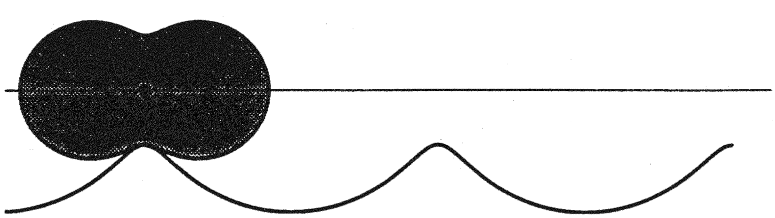


FIGURE 18

The example of a hippopedal wheel illustrates the case that the wheel is given in polar form and the road is found numerically.

Case 3. Suppose the road is given parametrically by $x = f_1(t)$, $y = f_2(t)$, where $f_1(0) = 0$. Now the initial-value problem relates the parameter t and the polar angle θ, and reduces to $d\theta/dt = -f_1'(t)/f_2(t)$, $\theta(0) = -\pi/2$. This time the numerical solution gives ordered pairs (t, θ), and the radius condition can be used to produce the wheel, very much as in Case 1.

Example. Let the road be a cycloid, lowered sufficiently so the closed-wheel condition (with an integer number, n, of periods per revolution) holds, and translated so the y-axis bisects one arch. The parametric equations are $x = t + \sin t$, $y = \cos(t) - d_n$. For this cycloid, the closed-wheel condition is

$$\int_0^{\pi} \frac{1 + \cos t}{d_n - \cos t}\, dt = \frac{\pi}{n}$$

for which the positive solutions are $d_n = 1 + 2n^2/(2n + 1)$. (One could also consider negative solutions: $d_1 = -1$ yields a road with a cardioid wheel similar to that in FIGURE 9.) Using the table of (t, θ) values obtained with the Runge-Kutta algorithm, and the radius condition, we get the wheels corresponding to different periods (see FIGURE 19). In this case both x and θ are given as functions of the parameter t, so in an animation neither will increase linearly. FIGURE 19 also shows the case of a period-1/2 roller; that is, $n = 1/2$, $d = 1\frac{1}{4}$ and the wheel rotates twice for each period of the cycloid. For the $n = 0$ case (infinitely many revolutions) see FIGURE 9(c).

Case 4. Suppose the wheel is given parametrically by $x = g_1(t)$, $y = g_2(t)$. In order to get a simple closed wheel, assume g_1 and g_2 are periodic with the same period. In terms of g_1 and g_2, the initial-value problem relating the x-coordinate of the road with the parameter t is:

$$\frac{dx}{dt} = \frac{g_1(t)g_2'(t) - g_1'(t)g_2(t)}{\sqrt{g_1(t)^2 + g_2(t)^2}}, \qquad x(0) = 0,$$

The numerical method produces pairs (t, x), and if care is taken to use the same t-values, the corresponding coordinates of the road are found from the radius condition to be given by: $y(t) = -\sqrt{g_1(t)^2 + g_2(t)^2}$.

To generate an animation of this case, we must also have the polar angle θ in terms of t. Unfortunately, there are complications involving branches of the arctangent function that prevent the direct use of $\theta = \arctan[g_2(t)/g_1(t)]$, so we generate another table of values, this time (t, θ), by applying Runge–Kutta to:

$$\frac{dx}{dt} = \frac{g_1(t)g_2'(t) - g_1'(t)g_2(t)}{\sqrt{g_1(t)^2 + g_2(t)^2}}, \qquad \theta(0) = -\frac{\pi}{2},$$

again being careful to use the same t-values as were used to generate the (t, x) pairs. Together, the (t, x) pairs and the (t, θ) pairs define a set of (x, θ) pairs that can be used to animate the wheel.

Example. Let the wheel be defined by $x = -\sin t + (1/2)\sin 2t$, $y = -\cos t$. This shape (FIGURE 20) is known as a *piriform* (again, see [3]). As in the example for Case 3, neither linear velocity nor angular velocity will be constant in an animation.

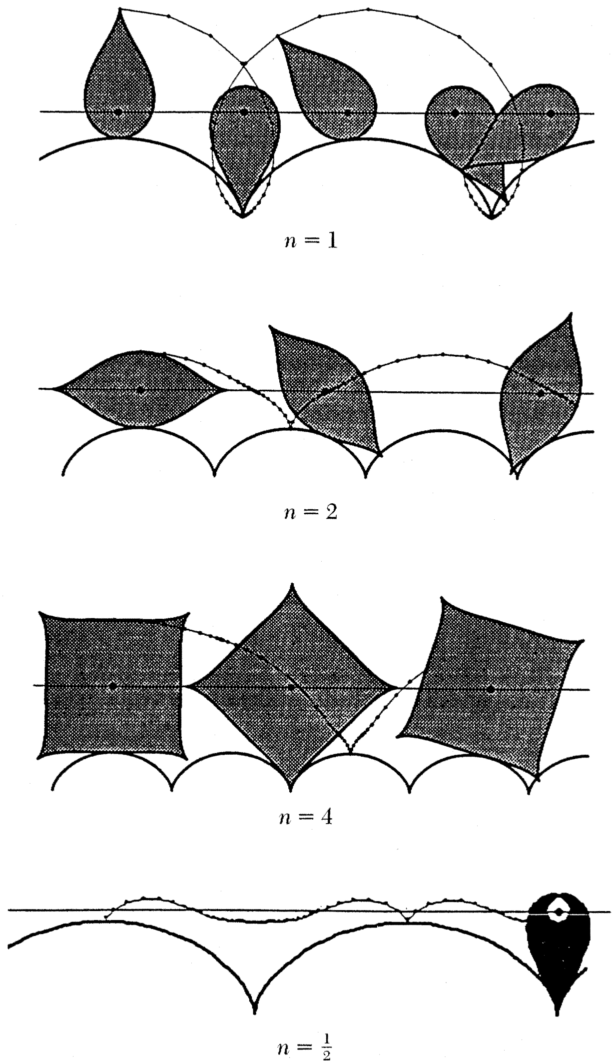


FIGURE 19

Some wheels corresponding to a cycloidal road. These wheels can be given in closed form, but when generating images or animations it is much more convenient to ignore the closed form and just use the numerical approach. The top wheel covers one cycloid period during each revolution; the next covers two; the next covers four. The bottom wheel covers a half-period of the road during each revolution and corresponds to $n = \frac{1}{2}$ in the closed-wheel condition.

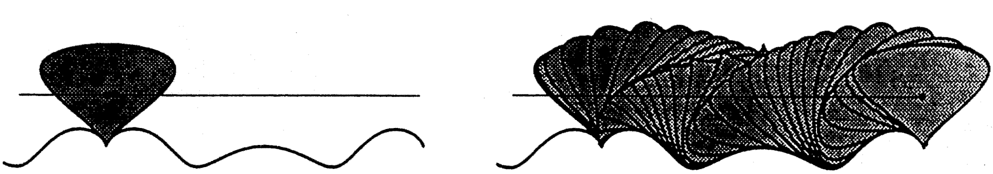


FIGURE 20

The piriform is an example of a parametrically given wheel. Its road is found, as in the other cases, by numerically solving the appropriate initial-value problem.

Remark. In all four cases, the choice of independent variable in the Runge–Kutta step is arbitrary. If an animation is planned, then this choice can be made to cause a particular quantity, usually x or θ, to change linearly in the animation.

4. Squaring the circle with Fourier series

A natural thought when dealing with periodic functions, such as the roads for closed wheels, is to look at the Fourier series. For wheels with an axis of symmetry through the axle, we can make the periodic function even, thus yielding a cosine series. The question then arises as to what shape the wheels for the various Fourier approximations will have. Clearly, for the approximation using only the constant term, the wheel will be a circle and, as the Fourier series more closely approximates the road, the wheels for the Fourier approximations will more closely approximate the original wheel. Thus if we begin with the road for a rolling square, the wheel for the Fourier approximations will "square the circle." Obviously, we are not restricting ourselves to Euclidean tools! There is one problem with this process, however: The Fourier approximations fail to satisfy the closed-wheel condition.

For example, suppose $p(x) = \cos x - \sqrt{5}$, which is a finite Fourier series already with period 2π, satisfying the closed-wheel condition. The constant term or 0th Fourier approximation, is $p_0(x) = -\sqrt{5}$, and so $\sqrt{5}$ is the radius of the corresponding wheel (a circle). It was shown earlier that the wheel for $p(x)$ traverses two periods of the road in one revolution, so the circumference of the circular wheel must be 4π (taking the period of the constant function the same as that of $p(x)$). But this makes the radius of the wheel equal to 2, a contradiction.

The case when $p(x)$ is the road for a square wheel is similar. Here, the circular wheel that rolls on the 0th Fourier approximation has circumference 7.12866 instead of $8 \operatorname{arcsinh} 1 \approx 7.05099$. And the closed-wheel condition fails here by 0.00138 for the two-term Fourier approximation road. Of course, as the trigonometric polynomials more closely approximate the original road, they will also come closer to satisfying the closed-wheel condition, but the condition fails nevertheless for each Fourier approximation.

What is needed, then, is a sequence of approximations, each satisfying the same closed-wheel condition as the road, which converges to the road. Such a sequence can be constructed by exploiting the orthogonality of the cosine functions in a Fourier series, along with the fact that the closed-wheel condition involves the reciprocal of the road function.

Road approximations having closed wheels Let $p(x)$ be a continuous, even, negative, periodic function with period T. Then $q(x) = 1/p(x)$ is also continuous, even, negative, and T-periodic, and so can be expanded in a Fourier cosine series. Denote the partial sums of the Fourier series for $q(x)$ by $q_n(x)$, $n = 0, 1, 2, \ldots$. The q_n's converge to q, are continuous, even, T-periodic, and, for large enough n, negative. If, in addition, we assume that the derivative of q is piecewise continuous, the convergence is uniform. Finally, we shall assume that q is "nice enough" so that all the q_n's are negative, which is the case in all our examples. We now form the sequence $\{p_n(x)\}$, where $p_n(x) = 1/q_n(x)$, which converges uniformly to $p(x)$, and we shall use this sequence to approximate the road. The p_n's have the same period as p and satisfy the same closed-wheel condition because

$$\int_{-T/2}^{T/2} -\frac{dx}{p_n(x)} = \int_{-T/2}^{T/2} -q_n(x)\,dx,$$

and, because the definite integrals of the cosines in the Fourier series vanish,

$$\int_{-T/2}^{T/2} -q_n(x)\,dx = \int_{-T/2}^{T/2} -q(x)\,dx = \int_{-T/2}^{T/2} -\frac{dx}{p(x)}.$$

Approximations to the square wheel Recall that the road for a square wheel with side 2 is the periodic extension of $y = -\cosh x$, for $-\operatorname{arcsinh} 1 \le x \le \operatorname{arcsinh} 1$. The first seven Fourier coefficients for the reciprocal are: -0.891107, -0.12537, 0.0230828, -0.0100959, 0.0056363, -0.00359458, 0.00249146. These were found using the standard formulas and integrating numerically. The first three approximations to the road are:

$$p_0(x) = -1.1222$$

$$p_1(x) = 1/(-0.891107 - 0.12537\cos(3.56443x))$$

$$p_2(x) = 1/(-0.891107 - 0.12537\cos(3.56443x) + 0.0230828\cos(7.12886x))$$

For the wheel corresponding to the approximation $p_k(x)$, the polar angle as a function of x is described by the initial-value problem $d\theta/dx = -1/p_n(x)$, $\theta(0) = -\pi/2$. This can be integrated in closed form since $-1/p_k(x)$ is a trigonometric polynomial, but the numerical method described in Section 3 is more efficient, especially for producing animations.

FIGURE 21 shows two corresponding positions of the road-wheel pairs for $p_0(x)$, $p_1(x)$, $p_2(x)$, and $p_6(x)$. Note how the circle becomes square-like very quickly.

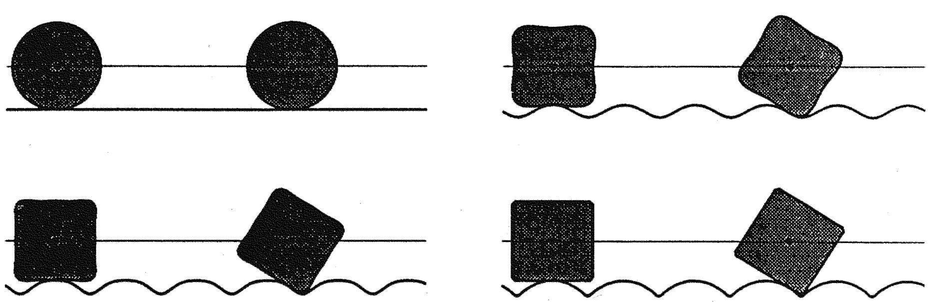


FIGURE 21
Approximating a catenary road with partial sums of its Fourier series yields wheels that transform a circle to a square.

5. A *Mathematica* wheel-building package

We have written a complete *Mathematica* package (version 1.2) that generates roads from wheels and vice versa. The package has options by which the user can generate stills or animations, with or without a locus, spokes, shading, and so on. The notebook can be obtained by sending a Macintosh disk to one of the authors. We include here a bare-bones version of the routine to give some idea of how such a program is written. This routine takes a function defining a road and displays the road and the wheel with its axle at the origin.

Even when a closed-form solution is available, it is often simpler to generate a diagram or animation by taking a numerical approach. Nevertheless, sometimes the closed form must be used. For example, rolling polygons are best generated directly, without numerical approximations. Some code for doing so can be found in the Appendix to [**6**].

```
RoadMovie[func_, {x_, xmax_}] := Module[{theta1, theta},
theta1[x_] = theta[x] /.
First @NDSolve[
          {theta' [x] == -1/func, theta[0] == -Pi/2},
          theta[x], {x, 0, xmax}, PrecisionGoal->2];
ParametricPlot[{-func {Cos[theta1[x]], Sin[theta1[x]]}, {x, func}},
  {x, 0, xmax}, PlotRange->All, AspectRatio->Automatic] ]
```

For example, `RoadMovie[Cos[x]-Sqrt[10], {x, 6 Pi}]` will generate the period-3 closed wheel for a cosine road (similar to FIGURE 2, but with $n = 3$).

REFERENCES

1. M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, 1965.
2. L. Hall, Trochoids, roses and thorns—beyond the Spirograph, *College Math. J.* 23(1992), 20–35.
3. J. D. Lawrence, *A Catalog of Special Plane Curves*, Dover Publications, Inc., New York, 1972.
4. G. B. Robison, Rockers and rollers, this MAGAZINE 33 (1960), 139–144.
5. P. Said, Problem E1809, *Amer. Math. Monthly* 74 (1967), 87.
6. S. Wagon, *Mathematica in Action*, W. H. Freeman & Co., New York, 1991.
7. D. G. Wilson, Problem E1668, *Amer. Math. Monthly* 72 (1965), 82–83.

If you ask mathematicians what they do, you always get the same answer. They think. They think about difficult and unusual problems. They do not think about ordinary problems: they just write down the answers.

M. Egrafov (translated from Russian), contributed by the late R.P. Boas, Jr.

Euler's Constant, Taylor's Formula, and Slowly Converging Series

JOHN V. BAXLEY
Wake Forest University
Winston-Salem, NC 27109

To R. P. Boas, Jr., with appreciation

1. Introduction

Because infinite series have fascinated me from the moment I first encountered them in calculus, I found the *Monthly* article [1] by Boas on partial sums of infinite series particularly interesting. Ever since I first read it, I have been experimenting with ways to communicate the essential ideas to my students in elementary calculus. For some years before Boas' article appeared, I had been using simple integral comparisons to illustrate some of the basic features, such as estimating the number n of terms required for the partial sum of the harmonic series to reach 100 ($n \approx 10^{43}$) or estimating the number of terms of

$$s = \sum_{k=2}^{\infty} \frac{1}{k(\log k)^2}$$

to achieve, say, two-decimal place accuracy ($|s - s_n| < .005$ requires about 10^{87} terms). What seemed harder to me for a while was computing the sum of this last and similar slowly converging series correct to several decimal places, since clearly one would not compute such a large partial sum. The approach taken by Boas depended on the Euler-Maclaurin summation formula, which he derives using Stieltjes integrals, not the kind of things I had time to discuss with my elementary calculus students.

The discussion below replaces the use of the Euler-Maclaurin formula by elementary manipulations with Taylor's formula. The knowledgeable reader will see, perhaps with some excitement, an interesting reversal of a familiar theme that is played when Taylor's formula is used to estimate the error in the midpoint rule for integration, where one replaces an integral by a sum and the replacement is good because the step size is small and the derivative is bounded. Here we replace a sum by an integral and the estimate is good because the derivative is small and the step size is bounded! These ideas are certainly not new; our purpose is only to encourage their use in elementary courses.

2. The Harmonic Series and Euler's Constant

We begin our story with a very elementary introduction to Euler's constant, the third member of the holy trinity (π, e, γ) of mathematical constants.

Let $s_n = \sum_{k=1}^{n}(1/k)$ be the nth partial sum of the harmonic series. That this series diverges is one of the first facts that a student learns about infinite series. Thus $\lim_{n\to\infty} s_n = +\infty$. By inscribing and circumscribing rectangles below and above the graph of $y = 1/x$ as in the proof of the integral test, it is easy to see that

$$\ln(n+1) < s_n < 1 + \ln(n).$$

This inequality quickly gives rough estimates of the number n of terms needed for s_n to exceed a given positive number A. For example, if $A = 7$, the left side of this inequality exceeds 7 for the first time when $n = 1{,}096$ and the right side exceeds 7 for the first time when $n = 404$. Thus we can be certain that s_n will exceed 7 for the first time for some n between 404 and 1,096. Repeating the exercise for $A = 13$, the corresponding range for n is 162,755 and 442,413, while for $A = 100$ the range is from about 9.89×10^{42} to about 2.69×10^{43}. Note that this crude approach gives a "ball park" estimate; the range is actually larger than the lower bound.

Reflecting on the simple derivation of the above inequality, we see that we have essentially used upper and lower sums in reverse; instead of approximating an integral by upper and lower sums, we have approximated upper and lower sums by an integral. Why not use the midpoint rule? Suppose we construct rectangles of width 1, centered at each integer k and with heights given by $1/k$. Although we lose the inequality, we obtain the approximation

$$s_n \approx \int_{\frac{1}{2}}^{n+\frac{1}{2}} \frac{1}{x}\,dx = \ln\left(n + \frac{1}{2}\right) + \ln 2.$$

Using this approximation, we would be led to the conclusion that s_n would reach the value A when n is approximately $(\exp(A) - 1)/2$. For $A = 7$, 13, 100, this estimate gives respectively, $n = 548$, 221,207, 1.344×10^{43}. We shall see that these estimates, while more accurate than the extremes obtained earlier, are all too small.

Our crude inequality implies that

$$0 < s_n - \ln n < 1,$$

giving reason to hope that

$$\gamma = \lim_{n\to\infty} (s_n - \ln n)$$

might exist. The "midpoint" approximation offers the naive guess that $\gamma \approx \ln 2 \approx .693$. Indeed, the limit does exist, it is called Euler's constant (Euler first discovered it in exactly this context), and is approximately .577. Here is a proof of the existence of γ, which is in the spirit of the above discussion and uses Taylor's formula in the same way as it is used to estimate the error in the midpoint rule. The proof produces at the same time an error bound for $s_n - \ln(n + \frac{1}{2}) - \gamma$, which simultaneously estimates the error in approximating γ by $s_n - \ln(n + \frac{1}{2})$ and the error in approximating s_n by $\gamma + \ln(n + \frac{1}{2})$. Let $f(x) = 1/x$ and use Taylor's formula to get

$$f(x) = f(k) + f'(k)(x-k) + \frac{f''(c_k)}{2}(x-k)^2,$$

where k is a positive integer and c_k is between k and x. Thus

$$\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} f(x)\,dx = f(k) + 0 + \frac{1}{2}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} f''(c_k)(x-k)^2\,dx. \tag{2.1}$$

This last integral cannot be evaluated because c_k depends on x and $f''(c_k)$ is not a constant. We let

$$I_k(f) \equiv \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} f(x)\,dx$$

and

$$r_k(f) = \frac{1}{2}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} f''(c_k)(x-k)^2\,dx$$

and rewrite (2.1) compactly as

$$f(k) = I_k(f) - r_k(f). \tag{2.2}$$

Since $f''(x) = 2/x^3$ is a decreasing positive function,

$$0 < r_k(f) < \frac{1}{2}f''\left(k-\frac{1}{2}\right)\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}(x-k)^2\,dx = \frac{1}{12\left(k-\frac{1}{2}\right)^3}. \tag{2.3}$$

Recalling that the p-series $\sum_{k=1}^{\infty}(1/k^p)$ converges for $p > 1$ and using the comparison test, we conclude that $\sum_{k=1}^{\infty} r_k(f)$ converges. By definition then, if we let $t_n = \sum_{k=1}^{n} r_k(f)$, we know that $t = \lim_{n\to\infty} t_n$ exists. But (2.2) gives

$$t_n = \sum_{k=1}^{n} I_k(f) - \sum_{k=1}^{n} f(k) = \int_{\frac{1}{2}}^{n+\frac{1}{2}} 1/x\,dx - s_n = \ln\left(n+\frac{1}{2}\right) + \ln 2 - s_n$$

and thus

$$s_n - \ln\left(n+\frac{1}{2}\right) = \ln 2 - t_n. \tag{2.4}$$

Moreover,

$$s_n - \ln n = -t_n + \ln 2 + \ln\left(1+\frac{1}{2n}\right)$$

and so

$$\gamma = \lim_{n\to\infty}(s_n - \ln n) = -t + \ln 2 \tag{2.5}$$

certainly exists. Subtracting (2.5) from (2.4) gives

$$s_n - \ln\left(n+\frac{1}{2}\right) - \gamma = t - t_n = \sum_{k=n+1}^{\infty} r_k(f). \tag{2.6}$$

Using (2.3) and inscribing rectangles under the graph of $1/(x-\frac{1}{2})^3$, we easily see that

$$0 < s_n - \ln\left(n+\frac{1}{2}\right) - \gamma < \int_n^{\infty}\frac{1}{12\left(x-\frac{1}{2}\right)^3}\,dx = \frac{1}{24}\frac{1}{\left(n-\frac{1}{2}\right)^2} \equiv E_n. \tag{2.7}$$

Thus $A_n s_n - \ln(n+\frac{1}{2})$ furnishes an over-approximation for γ with error given by (2.7). Rearranging (2.7), we get

$$\gamma + \ln\left(n+\tfrac{1}{2}\right) < s_n < \gamma + \ln\left(n+\tfrac{1}{2}\right) + \frac{1}{24\left(n-\frac{1}{2}\right)^2}. \tag{2.8}$$

TABLE A shows E_n and the corresponding approximation of γ for various values of n.

TABLE A. First approximation of γ

n	E_n	A_n
4	4×10^{-3}	.57926
10	5×10^{-4}	.577593
30	5×10^{-5}	.57726045
92	5×10^{-6}	.57722053
290	5×10^{-7}	.57721616
914	5×10^{-8}	.57721571
2888	5×10^{-9}	.57721567

Thus for example, calculating only for $n = 30$ gives the conclusion that $.5772104 < \gamma < .5772605$. Of course, close approximations of γ have been known for a long time; to 10 decimal places, γ is .5772156649.

3. Being lazy while finding γ

Using the results of the previous section, you can find γ to three decimal places using a simple calculator ($n = 10$) or to six or seven decimal places with a small computer ($n = 290$ or 914). But how could our grandparents have gotten γ to seven decimal places or how can we get γ to, say, 12 places with minimal computational effort? We shall see that with a little intelligent preparation, even a lazy person like me can have γ to high accuracy, using a simple calculator, and working very little. Our increasing accuracy will come by repeated mating of (2.2) and its descendants with Taylor's formula. Here is how we pass to the second generation. Begin with Taylor's formula of order three:

$$f(x) = f(k) + f'(k)(x-k) + \frac{f''(k)}{2}(x-k)^2 + \frac{f'''(k)}{3!}(x-k)^3 + \frac{f^{(4)}(c_k)}{4!}(x-k)^4,$$

where c_k is between x and k. Integrating from $k - \frac{1}{2}$ to $k + \frac{1}{2}$, we get

$$I_k(f) = f(k) + \frac{1}{24}f''(k) + r_k^{(4)}(f), \tag{3.1}$$

where

$$r_k^{(n)}(f) \equiv \frac{1}{n!}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} f^{(n)}(c_k)(x-k)^n\,dx. \tag{3.2}$$

Of course, the integral in (3.2) cannot be evaluated because c_k varies with x in an unknown fashion. We rearrange (3.1) and recycle (2.2) to substitute for $f''(k)$:

$$f(k) = I_k(f) - \frac{1}{24}I_k(f'') + \frac{1}{24}r_k^{(2)}(f'') - r_k^{(4)}(f), \tag{3.3}$$

where one should note that our $r_k^{(2)}$ is the same as r_k of the previous section. Note that both of the "r terms" involve integrals of $f^{(4)}$ so it is reasonable to assume that they will be very small. For convenience, let us for the time being abbreviate these combined terms by R_k. Recalling the definitions of t and t_n and that $f(x) = 1/x$, we sum (3.3) to get

$$t - t_n = \frac{1}{24}\int_{n+\frac{1}{2}}^{\infty} f''(x)\,dx + \sum_{k=n+1}^{\infty} R_k = \frac{1}{24\left(n+\frac{1}{2}\right)^2} + \sum_{k=n+1}^{\infty} R_k. \tag{3.4}$$

Let's temporarily "cheat," discard the last sum in hopes that it is small, and use the result in 2.6 to approximate γ with

$$\gamma \approx s_n - \ln\left(n+\tfrac{1}{2}\right) - \frac{1}{24\left(n+\tfrac{1}{2}\right)^2} = B_n. \tag{3.5}$$

Since we really do know γ to 10 decimal places, let's compute the right side (3.5) for various values of n and see how well we are doing. TABLE B shows the result for some selected values of n. (We will shortly describe the computation of E_n.)

TABLE B. Second approximation of γ

n	E_n	B_n
2	2.1×10^{-3}	.5770426
3	2.7×10^{-4}	.5771690
5	2.6×10^{-5}	.5772078
8	3.3×10^{-6}	.57721428
13	4.3×10^{-7}	.57721545
22	4.9×10^{-8}	.57721564
39	4.8×10^{-9}	.577215662

You will notice that we now have essentially 7 decimal places correct using only 22 terms, a significant improvement over the 914 required in TABLE A. Our ancestors, not as pampered as we are, did not blanch at that amount of arithmetic. However, they (and we) can do better yet. We just take more terms in Taylor's formula; you can easily verify that we now get, with the obvious notation:

$$f(k) = I_k(f) - \frac{1}{24}I_k(f'') + \frac{7}{8\cdot 24\cdot 30}I_k(f^{(4)}) + Q_k, \tag{3.6}$$

where Q_k consists of several error terms. Discarding the error terms, summing and integrating again, we now improve (3.5) to read

$$\gamma \approx s_n - \ln\left(n+\tfrac{1}{2}\right) - \frac{1}{24\left(n+\tfrac{1}{2}\right)^2} + \frac{7}{8\cdot 24\cdot 5\left(n+\tfrac{1}{2}\right)^4} = C_n. \tag{3.7}$$

Computing again with (3.7), we obtain TABLE C.

TABLE C. Third approximation of γ

n	E_n	C_n
2	8×10^{-4}	.5772293
3	4×10^{-5}	.5772176
4	5×10^{-6}	.57721611
6	4×10^{-7}	.577215715
9	3×10^{-8}	.57721567
12	4×10^{-9}	.577215666
17	5×10^{-10}	.577215665

We now have γ to 5 decimal places using $n=4$ and seven decimal places using $n=9$; not too much arithmetic for even us lazy people!

We return briefly to the error column in TABLE B. Examining (3.3) and (3.2), we see that each term $r_k^{(j)}$ is positive when j is even. Thus these error terms oppose each other and we may bound the error by finding a bound on each one and choosing the larger of the two bounds. An elementary calculation shows that the term involving $r_k^{(2)}(f'')$ leads to the larger bound and we get

$$|R_k| < \frac{f^{(4)}(k-\frac{1}{2})}{576}.$$

The term discarded in (3.4) satisfies

$$\left|\sum_{k=n+1}^{\infty} R_k\right| < \frac{1}{576}\sum_{k=n+1}^{\infty} f^{(4)}(k-\tfrac{1}{2}) < -\frac{1}{576}f^{(3)}(n-\tfrac{1}{2}) < \frac{1}{96(n-\frac{1}{2})^4}.$$

This is the bound we used for E_n in TABLE B.

A similar, albeit more tedious, calculation leads to the error bound

$$\frac{127}{12\cdot 24\cdot 49(n-\frac{1}{2})^6}$$

for the computation of E_n in TABLE C.

4. Partial sums of the harmonic series: examples

Our story continues with the following problem: given a positive number A, find the smallest positive integer N so that $s_N > A$. Our strategy is to replace s_n by a good approximation, solve the problem for the approximation, and then make sure that the approximation was good enough. For the harmonic series, we have several increasingly accurate choices based on our increasingly accurate approximations for γ:

$$s_n \approx \gamma + \ln(n+\tfrac{1}{2}) \tag{4.1}$$

$$s_n \approx \gamma + \ln(n+\tfrac{1}{2}) + \frac{1}{24(n+\frac{1}{2})^2} \tag{4.2}$$

$$s_n \approx \gamma + \ln(n+\tfrac{1}{2}) + \frac{1}{24(n+\frac{1}{2})^2} - \frac{7}{960(n+\frac{1}{2})^4} \tag{4.3}$$

We must face the fact that the more accurate approximation we use, the more difficult will be the algebra. The expedient route is to use the least accurate approximation to solve for N and then check afterwards. So we solve

$$\gamma + \ln(x+\tfrac{1}{2}) = A$$

for x to get

$$x = \exp(A-\gamma) - \tfrac{1}{2} \tag{4.4}$$

and then because N must be an integer, we take

$$N = 1 + \mathrm{int}(x), \tag{4.5}$$

where $\mathrm{int}(x)$ is the integer part of x. Is this value of N the correct answer to our question? We will deal with this problem more formally in section 7, but for the time being there are two ways to check. On the one hand we can compute s_n for $n = N$ and $n = N-1$ directly, or we can compute more quickly one of the approximations in (4.1), (4.2), or (4.3) and use the corresponding error estimate E_n to make sure that s_n for $n = N$ and $n = N-1$ are on opposite sides of A.

To illustrate, let $A = 7$. From (4.4) and (4.5), we get $N = 616$. This value is not too large to check directly on a small computer. We get $s_{615} \approx 6.99965073$ and $s_{616} \approx$

7.00127411. (Incidentally, when adding up such a partial sum on your computer, be sure to begin with the smallest number to reduce roundoff error.) Alternatively, for these values of n, we see from TABLE A that the error is about 10^{-7}; using (4.1), we get $s_{615} \approx 6.99965062$ and $s_{616} \approx 7.00127399$, allowing the certain conclusion that $N = 616$ is correct.

Now consider the case $A = 13$; (4.4) and (4.5) give $N = 248{,}397$. The time for a direct check is not overwhelming for a fast computer, as it is for small microcomputers, but in either case there is the nagging worry about the effect of roundoff on such a large partial sum. The error bound for TABLE A gives an error of about 10^{-13} for this N when using (4.1) to obtain $s_{248396} \approx 12.9999972$ and $s_{248397} \approx 13.0000012$, so once again we can be certain that 248,397 is the answer to our question.

To consider larger values of A (for example my computer gives $N = 36{,}865{,}412$ for $A = 18$), one must be able to compute using double precision for a careful check. Using single precision and the approximation (4.1), my microcomputer now does not distinguish between s_{N-1} and s_N and so cannot guarantee that this value of N is accurate. Using a larger computer with double precision, the approximation (4.1) gives $s_{36865411} \approx 17.999999977$ and $s_{36865412} \approx 18.0000000037$, thus certifying this value of N. For $A = 100$, we find $N \approx 1.51 \times 10^{43}$. TABLE D gives the value of N corresponding to integer values of A between 5 and 24.

TABLE D

A	N	A	N
5	83	15	1,835,421
6	227	16	4,989,191
7	616	17	13,562,027
8	1,674	18	36,865,412
9	4,550	19	100,210,581
10	12,367	20	272,400,600
11	33,617	21	740,461,601
12	91,380	22	2,012,783,315
13	248,397	23	5,471,312,310
14	675,214	24	14,872,568,831

You can have a lot of fun using these techniques on the examples

$$\sum_{k=2}^{\infty} \frac{1}{k \ln k} \tag{4.6}$$

and

$$\sum_{k=3}^{\infty} \frac{1}{k \ln k \ln \ln k} \tag{4.7}$$

and asking similar questions. For (4.5), I get

$$\gamma = \lim_{n \to \infty} \left(s_n - \ln \ln\left(n + \tfrac{1}{2}\right)\right) \approx .79467865.$$

(Note that this value differs, as it should, from the one in [1] by $\ln \ln 2$.)

5. Estimating other finite sums

Some popular elementary sums are

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

and

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}. \tag{5.1}$$

Such sums can be obtained quickly with the techniques of the previous sections. For example, to derive (5.1), you need only notice that in (3.3) the error terms for $f(x) = x^2$ vanish and so we sum to get

$$\sum_{k=1}^{n} k^2 = \int_{\frac{1}{2}}^{n+\frac{1}{2}} x^2\,dx - \frac{1}{24}\int_{\frac{1}{2}}^{n+\frac{1}{2}} 2\,dx,$$

which reduces after integration and some algebra to (5.1). But we can also consider more complicated sums like

$$s_n = \sum_{k=1}^{n} k^{3/2}. \tag{5.2}$$

Using (3.3) with $f(x) = x^{3/2}$, we get

$$f(k) - I_k(f) - \frac{1}{24} I_k(f'') = R_k$$

where

$$|R_k| < \frac{f^{(4)}(k-\frac{1}{2})}{576} = \frac{1}{1024(k-\frac{1}{2})^{5/2}}.$$

Thus

$$t_n = \sum_{k=1}^{n} R_k = s_n - \frac{2}{5}\left[\left(n+\frac{1}{2}\right)^{5/2} - \left(\frac{1}{2}\right)^{5/2}\right] + \frac{1}{16}\left[\left(n+\frac{1}{2}\right)^{1/2} - \left(\frac{1}{2}\right)^{1/2}\right], \tag{5.3}$$

and so

$$s_n = t_n + \frac{2}{5}\left[\left(n+\frac{1}{2}\right)^{5/2} - \left(\frac{1}{2}\right)^{5/2}\right] - \frac{1}{16}\left[\left(n+\frac{1}{2}\right)^{1/2} - \left(\frac{1}{2}\right)^{1/2}\right]. \tag{5.4}$$

Since t_n converges rapidly (the error can be estimated using the series $\sum_{k=n+1}^{\infty} R_k$), one gets a rather accurate estimate for t_n from (5.3) using quite small values of n ($t_n \approx 1.027$) and then (5.4) gives accurate values of the sum in (5.2) using this value for t_n. For example, direct computation of s_{100} gives 40,501.2245 while the right side of (5.4) gives 40,501.2248. Similar sums for larger values of the exponent can be closely approximated in similar fashion using the extensions of (3.3) in section 8.

6. Slowly converging series

For the discussion thus far, we have been thinking of the particular function $f(x) = 1/x$. However, all of our fundamental equations using Taylor's formula hold for any sufficiently differentiable function. When we estimated bounds on errors, we used the additional property that the appropriate even order derivative (f'', $f^{(4)}$, $f^{(6)}$) is decreasing.

To illustrate how our work may be applied to finding the sums of various slowly converging series, consider the example

$$\sum_{k=2}^{\infty} \frac{1}{k(\ln k)^2} \tag{6.1}$$

for which we let

$$f(x) = \frac{1}{x(\ln x)^2}. \tag{6.2}$$

Summing (2.2) gives

$$s - s_n \approx \int_{n+\frac{1}{2}}^{\infty} \frac{1}{x(\ln x)^2}\,dx$$

and

$$s \approx s_n + \frac{1}{\ln(n+\frac{1}{2})} = A_n. \tag{6.3}$$

The error in (6.3) can be estimated using the error term in (2.2). Some care must be taken since the derivatives become relatively complicated and one must take n sufficiently large that $f''(x)$ is decreasing for $x \geq n$. Using (6.3), we can find the smallest value of N for which the Nth partial sum s_N of this series approximates the sum s of the series with error less than a desired amount E. TABLE E shows the value of N corresponding to various values of E. The slowness of convergence is apparent; in excess of 17 million terms fails to give even one decimal place of accuracy!

TABLE E

E	N	E	N
.5	7	.11	8,874
.3	28	.10	22,026
.2	148	.09	66,910
.18	259	.08	268,337
.15	786	.07	1,600,320
.12	4,160	.06	17,307,780

Using (3.3) in place of (2.2) gives the more accurate approximation

$$s \approx s_n + \frac{1}{\ln(n+\frac{1}{2})} + \frac{1}{24}f'\left(n+\frac{1}{2}\right) = B_n. \tag{6.4}$$

The value of the sum s of the series (6.1) to 10 decimal places is 2.1097428012; TABLE F shows the accuracy of the approximation in (6.3) and (6.4) for various values of n.

TABLE F

n	A_n	B_n
3	2.115098574	2.109471296
5	2.110756887	2.109726878
10	2.109742376	2.109742735
20	2.109760852	2.109742788
40	2.109745657	2.1097428007
80	2.109743287	2.10974280122

7. Partial sums of the harmonic series: theory

We now use our estimates from section 2 and 3 to give a very elementary proof of the theorems of Boas in [2] and [3]. Let $A > 2$ and let γ be Euler's constant for the harmonic series. Let N be the largest integer less than or equal to $\exp(A-\gamma)$ and put $\delta = \exp(A-\gamma) - N$; note that $N \geq 4$. Then $x = N + \delta - \frac{1}{2}$ is the solution of $\gamma + \ln(x + \frac{1}{2}) = A$. We shall prove the following theorem, which is essentially the one Boas proved in [2].

THEOREM. *Let s_n be the nth partial sum of the harmonic series. Given $A > 2$, let N_A be the smallest positive integer for which $s_n \geq A$. Then, with N and δ as described above,*

(1) $$N_A \text{ is either } N \text{ or } N+1;$$

(2) $$\text{if } \delta < \frac{1}{2} + \frac{1}{24(N+\frac{1}{2})} - \frac{N+\frac{1}{2}}{96(N-\frac{1}{2})^4}, \quad \text{then } N_A = N;$$

(3) $$\text{if } \delta > \frac{1}{2} + \frac{1}{24(N+\frac{1}{2})} + \frac{1}{48(N-\frac{1}{2})^2} + \frac{N+1}{96(N-\frac{1}{2})^4}, \quad \text{then } N_A = N+1.$$

Proof. To prove (1), it suffices to show that $s_{N+1} > A$ and $s_{N-1} < A$. The first inequality follows immediately from the lower bound in (2.8), with $n = N+1$, since $N + \frac{3}{2} > x + \frac{1}{2}$. For the second inequality, we use the upper bound in (2.8), with $n = N-1$, to get

$$s_{N-1} < \gamma + \ln\left(N - \frac{1}{2}\right) + \frac{1}{24(N-\frac{3}{2})^2}. \tag{7.1}$$

Since $N \leqslant x + \frac{1}{2}$, it suffices to show that the right side of (7.1) is bounded by $\gamma + \ln(N)$, which is equivalent to

$$\ln(N) - \ln\left(N - \frac{1}{2}\right) > \frac{1}{24(N-\frac{3}{2})^2}. \tag{7.2}$$

Applying the mean value theorem to the left side of (7.2), it now suffices to show that

$$\frac{1}{2N} > \frac{1}{24(N-\frac{3}{2})^2},$$

which is easily seen to be true by elementary algebra.

To prove (2), we show that the stated conditions imply that $s_N > A$. Using the lower bound provided by (4.2) and its associated error bound (3.8), we need only show that

$$\ln\left(x+\frac{1}{2}\right)-\ln\left(N+\frac{1}{2}\right)<\frac{1}{24(N+\frac{1}{2})^2}-\frac{1}{96(N-\frac{1}{2})^4}. \tag{7.3}$$

Since the right side of (7.3) is clearly positive for $N \geq 4$, then (7.3) is obviously true if $x \leq N$, i.e. $\delta \leq \frac{1}{2}$. If $\delta > \frac{1}{2}$, then using the mean value theorem, (7.3) will be true if

$$\frac{x-N}{N+\frac{1}{2}}<\frac{1}{24(N+\frac{1}{2})^2}-\frac{1}{96(N-\frac{1}{2})^4}.$$

Using $x+\frac{1}{2}=N+\delta$, this last inequality is equivalent to the stated hypotheses on δ.

Passing to (3), we show that now $s_N < A$. Using the upper bound provided by (4.2) and (3.8), we only need to show that

$$\ln\left(x+\frac{1}{2}\right)-\ln\left(N+\frac{1}{2}\right)>\frac{1}{24(N+\frac{1}{2})^2}+\frac{1}{96(N-\frac{1}{2})^4}. \tag{7.4}$$

From the mean value theorem, (7.4) will be true if

$$\frac{x-N}{N+\delta}>\frac{1}{24(N+\frac{1}{2})^2}+\frac{1}{96(N-\frac{1}{2})^4},$$

which follows from stated hypotheses in (3) since $N+\delta<N+1$.

Our inequalities for δ look a little different from those of Boas [2] because of our preference for $N+\frac{1}{2}$; the form in [2] came from Boas' preference for the first form of the Euler-Maclaurin formula. It is not hard to see that our result is equivalent to the one in [2].

8. Concluding comments

Our purpose has been to make available to beginning students the fascinating ideas described by Boas [1], by using Taylor's formula instead of the Euler-Maclaurin formula. The alert reader will be asking for deeper insight into the relationship between these two formulas. It should come as no surprise that the Euler-Maclaurin formula is essentially a corollary of Taylor's formula. We shall briefly indicate how the Euler-Maclaurin formula may be derived from Taylor's formula.

The appearance of the error term in the Euler-Maclaurin formula will depend on which of the various forms we use for the remainder in Taylor's formula. The integral form of the remainder leads essentially to the usual form of the Euler-Maclaurin formula. It seems more elementary here to use the Cauchy form of the remainder. Our key is to first find the general form of (2.1), (3.3), and (3.6). Using the ideas of section 2, it is straightforward to prove by induction that

$$f(k)=\sum_{i=0}^{n}a_{2i}I_k(f^{(2i)})+R_k^{(2n)}(f), \qquad \text{for } n\geq 1, \tag{8.1}$$

where

$$|R_k^{(2n)}|\leq C_{2n}M_k^{(2n+2)},$$

$$M_k^{(j)}\equiv\max\left\{|f^{(j)}(x)|: k-\frac{1}{2}\leq x\leq k+\frac{1}{2}\right\},$$

and C_{2n} is independent of k while a_{2i} is given recursively by $a_0 = 1$ and

$$a_{2m} = -\sum_{i=1}^{m} \frac{a_{2(m-i)}}{2^{2i}(2i+1)!}.$$

Summing (8.1) on k, we obtain

$$\sum_{k=1}^{n} f(k) - \int_{\frac{1}{2}}^{n+\frac{1}{2}} f(x)\,dx = \sum_{i=1}^{N} a_{2i}\left(f^{(2i-1)}\left(n+\frac{1}{2}\right) - f^{(2i-1)}\left(\frac{1}{2}\right)\right) + \sum_{k=1}^{n} R_k^{(2N)}(f),$$

a form of the Euler-Maclaurin formula (cf. [**1**, p. 249, eqn. (2.4)]).

Our numbers a_{2i} bear, of course, a close relationship to the Bernoulli numbers, which the interested reader will enjoy working out.

I thank my colleague, Fred Howard, for several stimulating conversations about the Bernoulli numbers and the Euler-Maclaurin formula.

REFERENCES

1. Boas, R. P., Partial sums of infinite series and how they grow, *Amer. Math. Monthly* 84 (1977), 237–258.
2. Boas, R. P., Growth of partial sums of divergent series, *Math. of Comp.* 31 (1971), 257–264.
3. Boas, R. P. and J. W. Wrench, Jr., Partial sums of the harmonic series, *Amer. Math. Monthly* 78 (1971), 864–870.

On Computing Euler's Constant

JEFFREY NUNEMACHER
Ohio Wesleyan University
Delaware, OH 43015

Introduction

The basic constants of the mathematical universe, such as π, e, Euler's constant, and Khintchine's constant have an intrinsic fascination for many of us. They are the mathematical analogs of basic physical constants like Avogadro's number and Planck's constant that underlie the structure of the universe. That the same few constants appear in various guises in diverse branches of mathematics is evidence for the basic unity of the subject. The values of these constants are not easily expressible in finite form, so the problem of computing them to a prescribed accuracy is one of obvious interest and importance.

Euler's constant $\gamma = 0.5772156649\ldots$ is defined by the formula

$$\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \log n\right). \tag{1}$$

Many students first encounter it in a calculus course, where they are asked to prove that the limit exists. Euler's constant is a natural object: It measures the amount by which the partial sums of the simplest divergent series ultimately differ from the and C_{2n} is independent of k while a_{2i} is given recursively by $a_0 = 1$ and

$$a_{2m} = -\sum_{i=1}^{m} \frac{a_{2(m-i)}}{2^{2i}(2i+1)!}.$$

Summing (8.1) on k, we obtain

$$\sum_{k=1}^{n} f(k) - \int_{\frac{1}{2}}^{n+\frac{1}{2}} f(x)\,dx = \sum_{i=1}^{N} a_{2i}\left(f^{(2i-1)}\left(n+\frac{1}{2}\right) - f^{(2i-1)}\left(\frac{1}{2}\right)\right) + \sum_{k=1}^{n} R_k^{(2N)}(f),$$

a form of the Euler-Maclaurin formula (cf. [**1**, p. 249, eqn. (2.4)]).

Our numbers a_{2i} bear, of course, a close relationship to the Bernoulli numbers, which the interested reader will enjoy working out.

I thank my colleague, Fred Howard, for several stimulating conversations about the Bernoulli numbers and the Euler-Maclaurin formula.

REFERENCES

1. Boas, R. P., Partial sums of infinite series and how they grow, *Amer. Math. Monthly* 84 (1977), 237–258.
2. Boas, R. P., Growth of partial sums of divergent series, *Math. of Comp.* 31 (1971), 257–264.
3. Boas, R. P. and J. W. Wrench, Jr., Partial sums of the harmonic series, *Amer. Math. Monthly* 78 (1971), 864–870.

On Computing Euler's Constant

JEFFREY NUNEMACHER
Ohio Wesleyan University
Delaware, OH 43015

Introduction

The basic constants of the mathematical universe, such as π, e, Euler's constant, and Khintchine's constant have an intrinsic fascination for many of us. They are the mathematical analogs of basic physical constants like Avogadro's number and Planck's constant that underlie the structure of the universe. That the same few constants appear in various guises in diverse branches of mathematics is evidence for the basic unity of the subject. The values of these constants are not easily expressible in finite form, so the problem of computing them to a prescribed accuracy is one of obvious interest and importance.

Euler's constant $\gamma = 0.5772156649\ldots$ is defined by the formula

$$\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n} \frac{1}{k} - \log n\right). \tag{1}$$

Many students first encounter it in a calculus course, where they are asked to prove that the limit exists. Euler's constant is a natural object: It measures the amount by which the partial sums of the simplest divergent series ultimately differ from the approximating integral. These partial sums $H_n = \sum_{k=1}^{n} 1/k$, which are known as the harmonic numbers, occur frequently in the analysis of algorithms. As compared with π and e, its more famous siblings, not much is known about γ—even whether it is irrational is an open problem. The purpose of this note is to look briefly at some of the ways in which γ occurs and then to consider a variety of methods for computing its numerical value.

I have found this topic to be an attractive unifying theme for a numerical analysis course. It provides a good blend of theory and practical work (to implement the methods and to evaluate their accuracy) and clearly demonstrates that theory has a concrete payoff. It also gives a good opportunity for students to read some accessible short papers.

Background

A fundamental connection, which presumably gave rise to the usual symbol for Euler's constant—the Greek letter γ (Euler used C), is with the gamma function defined for $x > 0$ by

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\, dt. \tag{2}$$

This function provides a continuous extension of the discrete factorial function ($\Gamma(n) = (n-1)!$) and is a basic building block for many special functions of mathematical physics and standard distributions of statistics. Although the improper integral in (2) converges only for $x > 0$, Weierstrass proved that the gamma function could be expressed also as the infinite product

$$\Gamma(x) = e^{-\gamma x} \frac{1}{x} \prod_{k=1}^{\infty} \frac{e^{x/k}}{1 + x/k}, \tag{3}$$

which converges for all complex numbers except for 0 and the negative integers. This formula thus extends the gamma function into the complex plane, where it can be studied as an analytic function. Euler's constant γ is the only nonelementary object in this expansion. For a simple derivation of (3) as well as a beautiful treatment of most aspects of the gamma function see [**3**, p. 15].

An immediate consequence of (3) is that the derivative $\Gamma'(1)$ has the value $-\gamma$. If we differentiate under the integral sign in (2), we see that Euler's constant is the value of one of the simplest improper integrals:

$$\gamma = -\int_0^\infty e^{-t} \log t\, dt. \tag{4}$$

For what follows it will be useful to transform this integral. We write (4) in the form

$$\gamma = \int_0^1 \log t\, d(e^{-t} - 1) + \int_1^\infty \log t\, d(e^{-t})$$

and integrate by parts to obtain

$$\gamma = \int_0^1 \frac{1 - e^{-t}}{t}\, dt - \int_1^\infty \frac{e^{-t}}{t}\, dt. \tag{5}$$

The substitution $u = 1/t$ transforms the second integral into $\int_0^1 (e^{-1/u}/u)\, du$. Thus we obtain an integral representation for γ over a finite interval:

$$\gamma = \int_0^1 \frac{1 - e^{-t} - e^{-1/t}}{t}\, dt. \tag{6}$$

Notice that if the integrand is set equal to 1 at $t = 0$, formula (6) then gives a representation for γ as a proper Riemann integral.

There is another integral representation for γ that is quite elegant:

$$\gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2}\, dt,$$

where $\{t\} = t - [t]$ is the fractional part of t. This formula is a direct consequence of the Euler-Maclaurin summation formula discussed below (see [**2**, pp. 55–56]), but it is instructive to check it directly.

Euler's constant is also related to various other special functions. The exponential integral $E(x) = \int_x^\infty (e^{-t}/t)\, dt$ is a frequently occurring function in applied mathematics, since any integral of the form $\int R(x) e^x\, dx$, where $R(x)$ is a rational function, can be reduced to a linear combination of elementary functions and $E(x)$ [**15**, p. 42]. $E(x)$ has the useful expansion

$$E(x) = -\gamma - \log x - \sum_{k=1}^{\infty} \frac{(-x)^k}{k \cdot k!}. \tag{7}$$

To see this, write $E(x)$ in the form

$$E(x) = \int_1^\infty \frac{e^{-t}}{t}\, dt + \int_0^1 \frac{e^{-t} - 1}{t}\, dt - \int_0^x \frac{e^{-t} - 1}{t}\, dt - \int_1^x \frac{1}{t}\, dt.$$

By (5) the first two terms sum to $-\gamma$, and term by term integration of the Taylor expansion in the third term yields the result.

Another important analytic connection is with the Riemann zeta function that is defined for $\operatorname{Re} z > 1$ by the series $\zeta(z) = \sum_{n=1}^\infty 1/n^z$. The zeta function, aside from being the natural extension of the p-series of calculus into the complex plane, is a fundamental object of investigation in number theory. To see a simple reason for this, let p_k denote the kth prime. Then for each p_k the geometric series $\sum_{n=1}^\infty p_k^{-nz}$ sums to $1/(1 - p_k^{-z})$. Multiplying these expansions together and using the Fundamental Theorem of Arithmetic to express each n uniquely as a product of primes, we obtain the Euler product expansion

$$\zeta(z) = \prod_{k=1}^{\infty} \frac{1}{1 - p_k^{-z}}. \tag{8}$$

This argument can be made precise by proper attention to issues of convergence and demonstrates a direct connection between the zeta function and the sequence of primes.

The zeta function can be continued analytically into the complex plane as a meromorphic function whose only singularity is a pole of order one at $z = 1$. (One wants to do this, since to study any analytic function it is useful to have it defined on the largest natural domain.) The Laurent expansion for $\zeta(z)$ about the singularity, which is known as Kronecker's Limit Formula, is valid for all z different from 1 and shows a further connection with Euler's constant [**19**, p. 378]:

$$\zeta(z) = \frac{1}{z-1} + \gamma + \sum_{n=1}^{\infty} C_n (z-1)^n.$$

The Riemann hypothesis is a conjecture about the location of the zeros of the zeta function. It asserts that apart from the "trivial" zeros at the negative even integers (the existence of these zeros is far from obvious) all other zeros of $\zeta(z)$ are located on the line $\operatorname{Re} z = 1/2$. Many feel that this is the most significant open problem in mathematics. Strong theorems in number theory would follow if the truth of the conjecture could be established. A formula of Riemann shows a connection between this problem and Euler's constant. Let (z_i) denote the sequence of all nontrivial zeros of the zeta function, i.e., the zeros not located at the negative integers. Then Riemann showed that

$$\sum_i \frac{1}{z_i} = \tfrac{1}{2}\gamma + 1 - \log 2 - \tfrac{1}{2}\log \pi$$

and used this formula in his investigation of the hypothesis [**9**, p. 67].

We mention two other number theoretic formulas that involve γ. Mertens discovered a formula whose relation to the Euler product formula (8) is reminiscent of that of the defining formula (1) to the definition of the zeta function:

$$e^{-\gamma} = \lim_{n\to\infty} \log n \prod_{k=1}^{n}\left(1 - \frac{1}{p_K}\right)$$

[**17**, p. 162]. Let $d(k)$ denote the number of positive integer divisors of an integer k. Then the average number of divisors for all integers between 1 and n is approximately equal to $\log n$. More precisely, we have the asymptotic expansion [**2**, p. 57]

$$\frac{1}{n}\sum_{k=1}^{n} d(k) = \log n + (2\gamma - 1) + O\left(1/\sqrt{n}\right),$$

where $O(1/\sqrt{n})$ denotes a function $e(n)$ that satisfies $|e(n)| < c/\sqrt{n}$ for some constant c.

Finally, we give a few miscellaneous series expansions for γ. The definition of γ expresses the constant as the limit of partial sums of a divergent series minus a correction term. Simple algebraic manipulation allows (1) to be re-expressed as the sum of the convergent series

$$\gamma = \sum_{n=1}^{\infty}\left[\frac{1}{n} - \log\left(1 + \frac{1}{n}\right)\right]. \tag{9}$$

A series for γ that involves only fractions was discovered by Addison [**1**]:

$$\gamma = \frac{1}{2} + \sum_{n=1}^{\infty}\sum_{m=2^{n-1}}^{2^n-1} \frac{n}{(2m)(2m+1)(2m+2)}.$$

Euler found an expression for γ in terms of p-series. Let $S_n = \sum_{k=2}^{\infty} 1/k^n = \zeta(n) - 1$. Then

$$\gamma = 1 - \sum_{n=2}^{\infty}\frac{1}{n}S_n.$$

This formula can be verified using simple but clever manipulation with series (start with the Taylor expansion for $\log(1 - x)$).

Elementary approximation methods

Let us now consider the problem of determining the numerical value for γ correct, say, to 10 decimal places. We take this to mean that the computed value should differ from the exact one by less than 0.5×10^{-10}. A reasonable way to begin is by investigating how rapidly the expression following the limit in the defining formula (1) approaches γ. The rate of convergence turns out to be very slow, a fact that can be seen in a variety of ways.

Rao gives a pleasantly direct approach in [16]. Let us denote the expression following the limit in (1) by a_n. Then each term can be expressed as an integral, that yields

$$a_n = \int_0^{\infty} (e^{-x} + e^{-2x} + \cdots + e^{-nx})\, dx - \int_0^{\infty} \frac{e^{-x} - e^{-nx}}{x}\, dx.$$

Summing the series and regrouping show that

$$a_n = \int_0^{\infty} \left(\frac{e^{-x}}{1 - e^{-x}} - \frac{e^{-x}}{x} \right) dx + \int_0^{\infty} e^{-nx} \left(\frac{1}{x} - \frac{1}{e^x - 1} \right) dx.$$

The second term goes to zero as n goes to infinity and gives an explicit formula for the difference between γ and the first term, which must therefore be yet another integral representation for γ. The elementary inequality

$$\frac{1}{2} - \frac{x}{8} < \frac{1}{x} - \frac{1}{e^x - 1} < \frac{1}{2},$$

a consequence of the fact that $\tanh x < x$ for $x > 0$, then implies that

$$\frac{1}{2n} - \frac{1}{8n^2} < a_n - \gamma < \frac{1}{2n}. \tag{10}$$

Thus a_n tends to γ with speed approximately $1/(2n)$. But if we now replace a_n by $a_n - 1/(2n)$, the result is a sequence that converges to γ much more rapidly: The error is now less than $1/(8n^2)$. Thus a program to calculate $a_n - 1/(2n)$ with $n = 50$ will give 4 place accuracy (error $< .00005$), and $n = 50{,}000$ suffices for 10 place accuracy.

Another elementary method due to Horton [12] is based on formula (9) and a simple fact about series: If $\sum_{k=1}^{\infty} b_k$ is a convergent series and f is a decreasing function satisfying $f(k) = b_k$, then $\sum_{k=1}^{\infty} b_k$ differs from $\sum_{k=1}^{n} b_k + \int_n^{\infty} f(x)\, dx$ by less than b_n. The proof is a simple generalization of the usual argument leading to the integral test. In our case

$$f(x) = \frac{1}{x} - \log\left(1 + \frac{1}{x}\right),$$

and we find that γ differs from

$$\sum_{k=1}^{n} \left[\frac{1}{k} - \log\left(1 + \frac{1}{k}\right) \right] + \left[-1 - (n+1) \log \frac{n}{n+1} \right]$$

by less than

$$\frac{1}{n} - \log\left(1 + \frac{1}{n}\right).$$

To obtain an error less than $.5 \times 10^{-10}$ it suffices to set $n = 100{,}000$. This is twice the number of terms necessary using Rao's method.

A third elementary approach is to use a numerical integration technique such as Simpson's Rule to approximate an integral representation for γ. The simplest integral representation that we have encountered, namely (4), would not be a good choice for this purpose, since the interval of integration is infinite and the integrand becomes infinite at the left endpoint. Formula (6) suffers from neither of these problems, so we expect it to be numerically more tractable. Let us denote the integrand in (6) by h.

Care is required in entering h into a computer program to implement Simpson's Rule. Not only does the formula for h have no value at the left endpoint but for x near 0, evaluation of $\exp(-1/x)$ will result in a runtime error on any system where $\exp(-x)$ is evaluated as $1/\exp(x)$. The natural way to avoid this problem is to omit this term when it becomes negligible. Thus instead of h we integrate numerically the function g defined in Pascal as follows:

```
function g(x: real): double;
begin
  if x = 0 then g := 1
  else if x < .025 then g := (1 - exp(-x))/x
  else g := (1 - exp(-x) - exp(-1/x))/x
end;
```

Here the constant .025 is chosen so that g differs from h by less than 10^{-15}, a negligible value if we are trying for 10-place accuracy.

To assess the accuracy of Simpson's approximation requires information concerning the size of the fourth derivative of the integrand. The error in the Simpson approximation to $\int_a^b g(x)\,dx$ computed using a partition into n equal subintervals is less than

$$\frac{(b-a)^5}{180}\frac{1}{n^4}\max_{a \leqslant x \leqslant b}\left|g^{(iv)}(x)\right|$$

(see any calculus book). It is a good exercise in numerical calculus (either by hand assisted by a Newton's Method program or else by using an algebraic package such as Derive) to determine the maximum of $|g^{(iv)}(x)|$ in the interval $[0,1]$. The answer turns out to be a bit less than 4,000, which implies that 10-place accuracy can be achieved with n equal to 818. Thus of our three elementary methods it appears that Simpson's Rule is the most efficient. It does, however, require many evaluations of a complicated function, so depending upon the hardware being used, it may not be significantly faster than Rao's method.

If one uses the Trapezoidal Rule instead of Simpson's Rule, the error bound requires finding an upper bound for $|g''(x)|$, which is much easier (by hand) than the corresponding problem for $|g^{(iv)}(x)|$. But it turns out that $n = 113{,}000$ is required to guarantee that this error is less than $.5 \times 10^{-10}$.

There are, of course, much more efficient methods of numerical integration than Simpson's Rule. Some of these will achieve 10-place accuracy with very few function evaluations. For example, an adaptive Gaussian approach needs only seven points. But assessing the accuracy of these methods requires maximizing higher derivatives of the integrand, a process that rapidly becomes more and more unwieldy. To achieve accuracy much greater than 10 places we need other ideas.

More sophisticated methods

The three approaches discussed above all required nothing more advanced than elementary calculus. We now consider some more sophisticated methods with the aim of decreasing the number of operations required to attain 10-place accuracy and also of computing γ to arbitrarily high accuracy. In the last few years new ideas have enabled mathematicians to calculate π and other quantities to many millions of decimal places (see, for example, [5] and [8]). These new methods are based on deep properties of special functions that were better known to the general mathematical community in the nineteenth century than in the twentieth. One of the Borwein algorithms that is based on a formula due to Ramanujan converges to π quartically, i.e., each iteration quadruples the number of decimal places of accuracy. Analogous methods for Euler's constant are not known: The best current methods converge only linearly and are appropriate for calculation of γ to thousands, not millions, of decimal places. They are based, however, on much simpler mathematics than the very rapid algorithms for π.

A standard tool for evaluating a sum to a prescribed accuracy is the Euler-Maclaurin summation formula:

$$\sum_{a\leqslant k<b} f(k)=\int_a^b f(x)\,dx+\sum_{k=1}^{m}\frac{B_k}{k!}f^{(k-1)}(x)\bigg|_a^b+R_m, \tag{11}$$

where a and b are integers and

$$R_m=(-1)^{m+1}\int_a^b\frac{B_m(\{x\})}{m!}f^{(m)}(x)\,dx.$$

Here the constants B_k are the Bernoulli numbers, which are defined by the power series expansion

$$\frac{z}{e^z-1}=\sum_{m=0}^{\infty}\frac{B_m}{m!}z^m. \tag{12}$$

The function $B_m(\{x\})$ in the remainder term R_m is the mth Bernoulli polynomial evaluated at the fractional part of x, where the Bernoulli polynomial is defined by the formula $B_m(x)=\sum_{k=0}^{m}\binom{m}{k}B_kx^{m-k}$.

The Euler-Maclaurin formula is really just a consequence of clever integration by parts. Thus approximation based on it might more aptly fit under "elementary methods" except for the asymptotic analysis of the behavior of the Bernoulli numbers. For a nice discussion and proof of this formula see [**10**, Sect. 9.5 and 9.6] from which the following application and notation are drawn.

Since

$$\frac{z}{e^z-1}+\frac{z}{2}=\frac{z}{2}\coth\frac{z}{2},$$

an even function, it follows that B_k vanishes for k odd and greater than 1. The Bernoulli numbers can be calculated most readily from the implicit recurrence relation

$$B_k=\sum_{j=0}^{k}\binom{k}{j}B_j$$

together with the initial conditions $B_0 = 1$ and $B_1 = -1/2$. This relation follows directly from (12) and is valid for $k \geqslant 2$. It yields $B_2 = 1/6$, $B_4 = -1/30$, and $B_6 = 1/42$. Although initially small, the Bernoulli numbers grow rapidly: It can be shown that

$$\left|\frac{B_{2k}}{B_{2k-2}}\right| \approx \frac{2k(2k-1)}{4\pi^2}.$$

When (11) is applied to the case of $f(x) = 1/x$ with $a = 1$, $b = n$, and m replaced by $2m$, we obtain

$$\sum_{k=1}^{n-1} \frac{1}{k} = \log n + C_{2m} + \frac{B_1}{n} - \sum_{k=1}^{m} \frac{B_{2k}}{2kn^{2k}} + R_{2m}(n), \tag{13}$$

where C_{2m} is a constant independent of n. Let us write the remainder term in the form $R_{2m}(n) = R_{2m}(\infty) - R^*_{2m}(n)$, where the interval of integration is $[1, \infty]$ in $R_{2m}(\infty)$ and $[n, \infty]$ in $R^*_{2m}(n)$. Denoting by C the constant $C_{2m} + R_{2m}(\infty)$ and adding $1/n$ to both sides of (13), we now have

$$\sum_{k=1}^{n} \frac{1}{k} = \log n + C + \frac{1}{2n} - \sum_{k=1}^{m} \frac{B_{2k}}{2kn^{2k}} - R^*_{2m}(n). \tag{14}$$

By letting $n \to \infty$ it is easy to see that C is equal to γ. It is not difficult to show that when $f^{(2m+2)}(x) \geqslant 0$ and $f^{(2m+4)}(x) \geqslant 0$, as is the case for $f(x) = 1/x$, the remainder $R^*_{2m}(n)$ can be expressed in the form $\theta_{m,n} B_{2m+2}/[(2m+2)n^{2m+2}]$, where $\theta_{m,n}$ is a constant in $(0, 1)$. Thus we have obtained a very efficient formula for the calculation of γ (or of H_n if γ is known):

$$H_n = \sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + \frac{1}{2n} - \sum_{k=1}^{m} \frac{B_{2k}}{2kn^{2k}} + \theta_{m,n} \frac{B_{2m+2}}{(2m+2)n^{2m+2}}. \tag{15}$$

When $m = 2$ formula (15) becomes

$$H_n = \sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \theta_{2,n} \frac{1}{252n^6}.$$

The choice of $n = 21$ is enough to make the error term less than 0.5×10^{-10}, so we can obtain 10-place accuracy with very little calculation. Notice that Rao's inequality (10) is the case of $m = 0$ in (15) with a slightly coarser error estimate. Note also that for fixed m the error term can be made as small as required by choosing n sufficiently large, but if n is fixed then letting m grow results in a divergent series.

This method was used by Adams for the first high-precision calculation of γ in 1878. He obtained 263 decimal digits without a computer! In 1962 Knuth in [14] used formula (14) with $n = 10{,}000$ and $m = 250$ to obtain a value for γ correct to 1,271 decimal places. The most time-consuming part of Knuth's calculation was the determination of the first 250 Bernoulli numbers. The need for these constants is the major disadvantage of this very general summation method.

Sweeney in [18] suggested a different idea for a high-precision calculation of γ, which avoids the necessity of calculating the Bernoulli numbers. His approach is based on formula (7), which relates γ to an expansion of the exponential integral $E(x)$. Let us rewrite (7) in the form

$$\gamma = -\log x + S(x) - E(x),$$

where $S(x)=\sum_{k=1}^{\infty}(-1)^{k+1}x^k/(k\cdot k!)$ and consider each of the terms separately.

Recall that

$$E(x)=\int_x^{\infty}\frac{e^{-t}}{t}\,dt.$$

The elementary estimate $e^{-t}/t\leqslant e^{-t}/x$ for $t\geqslant x>0$ establishes that $0<E(x)\leqslant e^{-x}/x$. If x is a power of 2, then $\log x$ can be expressed in terms of $\log 2$, which can be calculated, for example, from the infinite series expansion

$$\log 2=\sum_{k=1}^{\infty}\frac{1}{k\cdot 2^k}.$$

If the mth partial sum is used as an approximation, the truncation error is bounded by $\int_m^{\infty}(1/x\cdot 2^x)\,dx$, which can be evaluated in terms of the exponential integral. The series for $S(x)$ is alternating, so the truncation error in approximating it by the mth partial sum is bounded by the first omitted term, which can be estimated using Stirling's formula. Thus for any $\varepsilon>0$ it is possible to determine x sufficiently large so that $E(x)<\varepsilon/3$, and the functions $\log x$ and $S(x)$ each can be computed to within $\varepsilon/3$. This gives a means of calculating γ to within ε.

This method was used by Sweeney to obtain 3,566 decimal digits of Euler's constant, then later in modified form by Brent in [6] to obtain 20,700 digits. As is the case in any high precision calculation, care is needed in determining how many digits to keep in the intermediate calculations to control the error due to round-off and cancellation. A similar method based on a formula relating γ to certain Bessel functions has been developed by Brent and McMillan in [7] and used to obtain 30,100 decimal digits of γ.

One reason for wanting to know a constant to such high accuracy is that it is then possible to compute accurately other quantities that involve the constant. Any real number x can be expressed in the form

$$x=a_0+\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+\cdots}}}$$

where the a_k's are integers with $a_k\geqslant 0$ for $k>1$ (see [11], Chap. 10). This representation is called the regular continued fraction expansion of x. If the continued fraction is truncated at the nth stage, the resulting fraction p_n/q_n is called the nth convergent of the expansion. It gives the best rational approximation to x of order q_n in the sense that no other rational number with denominator smaller than q_n can approximate x more closely. From knowing the decimal expansion of γ to many digits of accuracy it is possible to compute many convergents. It follows from Brent and McMillan's calculations of these convergents that if γ is rational, then when expressed in lowest terms the denominator of γ must exceed $10^{15,000}$.

There are two further bits of numerical evidence for the irrationality of Euler's constant. Recent work of Bailey [4] shows that if γ is an algebraic number that satisfies a degree eight polynomial equation with integer coefficients, then the Euclidean norm of the vector of coefficients must exceed 3.57×10^9. This is numerical evidence for the transcendence of Euler's constant, hence for its irrationality as well. There is a theorem of Khintchine [13] that states that for almost all real numbers x the geometric means $G_n(x)=(a_1\cdot a_2\cdot a_3\cdot\cdots\cdot a_n)^{1/n}$ of the partial quotients a_k

occurring in the regular continued fraction expansion of x converge to an absolute constant $K = 2.685452001\ldots$, which is known as Khintchine's constant. The excluded set of measure zero includes all rationals, since for them the partial quotients a_k eventually vanish. The calculations in [6] show that $G_{20,000}(\gamma) = 2.6908$, which again suggests that γ is irrational.

Note. The author wishes to thank the referees for their useful suggestions and corrections.

REFERENCES

1. A. W. Addison, A series representation for Euler's constant, *Amer. Math. Monthly* 74 (1967), 823–824.
2. Tom Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, New York, 1976.
3. Emil Artin, *The Gamma Function*, Holt, Rinehart and Winston, New York, 1964.
4. David Bailey, Numerical results on the transcendence of constants involving π, e, and Euler's constant, *Math. of Comp.* 50 (1988), 275–281.
5. J. M. Borwein and P. B. Borwein, Ramanujan, modular equations, and approximations to pi or how to compute one billion digits of pi, *Amer. Math. Monthly* 96 (1989), 201–219.
6. Richard Brent, Computation of the regular continued fraction for Euler's constant, *Math. of Comp.* 31 (1977), 771–777.
7. Richard Brent and Edwin McMillan, Some new algorithms for high precision computation of Euler's constant, *Math. of Comp.* 34 (1980), 306–312.
8. D. V. Chudnovsky and G. V. Chudnovsky The computation of classical constants, *Proc. Natl. Acad. Sci.* USA 86 (1989), 8178–8182.
9. Harold Edwards, *Riemann's Zeta Function*, Academic Press, New York, 1974.
10. Ronald Graham, Donald Knuth, and Oren Patashnik, *Concrete Mathematics*, Addison Wesley, Reading, MA., 1989.
11. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, 4th edition, 1960.
12. Goldie Horton, A note on the calculation of Euler's constant, *Amer. Math. Monthly* 23 (1916), 73–74.
13. A. Ya. Khintchine, *Continued Fractions*, 3rd edition, Noordhoff, Groningen, 1963.
14. Donald Knuth, Euler's constant to 1271 places, *Math. of Comp.* 16 (1962), 275–281.
15. N. N. Lebedev, *Special Functions and their Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1965.
16. S. K. Kakshmana Rao, On the sequence for Euler's constant, *Amer. Math. Monthly* 63 (1956), 572–573.
17. Paulo Ribenboim, *The Book of Prime Number Records*, Springer-Verlag, New York, 1988.
18. Dura Sweeney, On the computation of Euler's constant, *Math. of Comp.* 17 (1963), 170–178.
19. Giovanni Sansone and Johan Gerretsen, *Lectures on the Theory of Functions of a Complex Variable*, Noordhoff, Groningen, 1960.

Lagrange (1736-1813) was revolted by the cruelties of the Terror that followed the French Revolution. When the great chemist Lavoisier went to the guillotine, Lagrange expressed his indignation at the stupidity of the execution: "It took the mob only a moment to remove his head; a century will not suffice to reproduce it."

An Introduction to the History of Mathematics,
Howard Eves, Fifth Edition, Saunders Publishing.

NOTES

Optimal Steiner Points

REGINA B. COHEN
Villanova University
Villanova, PA 19085

Given three points P, Q, and R in the plane, the shortest path connecting the points using only straight lines between the points is equal to the sum of the shorter two sides of $\triangle PQR$. By introducing a fourth point interior to $\triangle PQR$ when all its angles are less than 120°, a shorter network is obtained. There is a point S in $\triangle PQR$ such that the sum of the Euclidean distances of S from the three vertices is minimized. This point is called the *Steiner point* of the triangle, after the mid-19th-century Swiss geometer Jacob Steiner, an early investigator of the question of minimal networks among a finite number of points [2].

For example, in an equilateral triangle of side 1, a minimal path that connects the vertices and stays on the perimeter has length 2, whereas a network of vertices and Steiner point S has length $\sqrt{3} \doteq 1.732$, still more than half the perimeter.

By introducing a well-known Minkowski metric and constructing the Steiner point with respect to that metric, the length of the minimal network can be improved for any three vertices to half the length of the perimeter. Such a Steiner point is optimal in the sense that the lengths in the network could not be shortened and still satisfy the triangle inequality required in a metric space. A point S in the metric space (d, X) is said to be an *optimal Steiner point* of $\triangle PQR$ if the d-length of the network connecting P, Q, R, and S is half the d-length of the perimeter of $\triangle PQR$. We shall see that, if the metric is Minkowski, all triangles have optimal Steiner points if and only if the defining unit ball is a parallelogram.

We start by noting some familiar facts. Given any three positive real numbers $[a, b, c]$, some or all of which may be the same, the triple can be the distances between three points in the Euclidean plane, denoted by E^2, as long as the sum of any two is greater than or equal to the third. $[1,1,1]$ can be realized in E^2 as the distances between the vertices P_1, P_2, P_3 of any equilateral triangle of side length 1. $[3,4,5]$ and $[1,1,2]$ will do, but $[1,1,2.001]$ will not. These are particular examples of the more general concept of a metric set.

A set F of n abstract points together with a distance function $d\colon F \times F \to [0, \infty)$ satisfying the axioms for a metric space is called a *metric set of n points*. Any metric set of three points can be embedded in the plane with the Euclidean distance function as a triangle or as collinear points. When a fourth point is introduced, there are six distances. Denoting $d(P_i, P_j)$ by d_{ij} for $i, j = 1,2,3,4$, we always keep the order $[d_{12}, d_{23}, d_{31}, d_{14}, d_{24}, d_{34}]$.

Going back to an equilateral triangle of side 1 and vertices P_1, P_2, and P_3 in E^2, let us look for a fourth point P_4 that has distance 1 from each vertex. The metric set $\{P_1, P_2, P_3, P_4\}$ will have associated with it the list of six distances $[1,1,1,1,1,1]$. There is no such point in the plane. If we are willing to go to Euclidean 3-space, the four points can be realized as vertices of a tetrahedron with all sides length 1.

Now let P_4 vary in E^2 over the interior of the above triangle. For each position of P_4, there will be a corresponding set of three last entries in the list of six distances for the metric set $\{P_1, P_2, P_3, P_4\}$. When P_4 is the Steiner point of the triangle, the sum of the Euclidean distances is minimized. The set of distances for the vertices of the triangle and the Steiner point is $[1,1,1,\sqrt{3}/3,\sqrt{3}/3,\sqrt{3}/3]$. There are many points P_4 not realizable in E^2 such that the distances associated with the metric set of the equilateral triangle in the plane together with P_4 undercut $\sqrt{3}$ with the sum of the last three entries. The limiting case for such points would be the metric set with distances $[1,1,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}]$. There exists no fourth point, even in Euclidean space of three dimensions, which has distance of $\frac{1}{2}$ from each vertex of the triangle. There are infinitely many metric sets of four points that cannot be embedded in E^2.

Let's look at another familiar metric in the plane, the *max metric*, where the unit "circle" is the square,

$$\{(x,y): |x|=1, |y| \leqslant 1\} \cup \{(x,y): |y|=1, |x| \leqslant 1\}.$$

This metric is similar, except for orientation, to the taxicab metric d_T where the unit "circle" is $\{(x,y): |x|+|y|=1\}$. These are particular examples of a Minkowski metric in the plane, where the defining unit ball is a closed, convex, centrally symmetric set. Denoting the max metric by d_M, we have

$$d_M(P,Q) = \max(|x-x'|, |y-y'|),$$

where (x,y) and (x',y') are the Cartesian coordinates of P and Q respectively. A surprising feature of the max metric is that any metric set of four points can be embedded in the plane endowed with this metric. The construction from Wolfe [4] is given in the appendix. Four points with six mutual distances of 1 are $(0,0)$, $(0,1)$, $(1,1)$, and $(1,0)$. Four points whose set of distances is $[1,1,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}]$ are $(0,0)$, $(\frac{1}{2},1)$, $(1,0)$, and $(\frac{1}{2},\frac{1}{2})$.

We will consider this last example more closely. In any metric space (X,d) a point R is said to be (metrically) *between* P and Q if $d(P,Q)=d(P,R)+d(R,Q)$. In the above example, the point $(\frac{1}{2},\frac{1}{2})$ is between each pair of the other three points. This cannot happen in Euclidean space, where a fourth point can be between at most two pairs of the other three points. The geometry of betweenness is rather simple in E^2: R is between P and Q if it lies on the line segment $\overline{PQ}$ from P to Q, because the circle with center P and radius $d_E(P,R)$ is tangent to the circle with center Q and radius $d_E(Q,R)$ at the one point R. The situation is different in the max metric. In FIGURES 1a and 1b, $d_M(P,Q)=3$ and all points on the heavy line are at a distance 1 from P and 2 from Q. The distance-measuring "circles" in the max metric touch along a vertical segment, unless P and Q both lie on a line of slope $+1$ or -1, as in FIGURE 1c. In the Euclidean metric, or in any Minkowski metric where the defining unit ball is strictly convex, the set of points between two given points consists of the line segment connecting them. In the max metric, however, the set of points between two points P and Q is a rectangle and its interior with sides of slope ± 1 and the line segment $\overline{PQ}$ as a diagonal. See FIGURE 2. Denote this set by $B_M(P,Q)$.

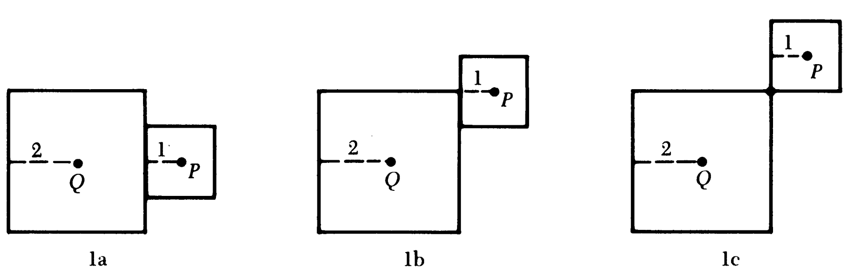


FIGURE 1

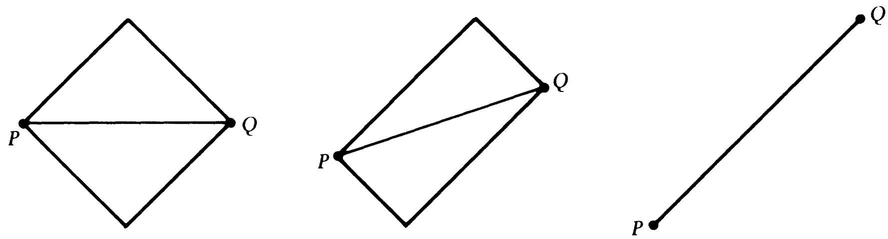


FIGURE 2
$B_M(P, Q)$

Using the concept of betweenness, the Steiner point S of a triangle in the max metric is easy to construct. Further, S is optimal in the sense that the sum of the distances from S to the vertices is equal to half the length of the perimeter, because S is between each pair of vertices. (Note: For a step-by-step construction of the Steiner point of $\triangle PQR$ in E^2, see Ehrmann [3]. If the angle measure at any vertex is 120° or more, the Steiner point in E^2 is at that vertex; otherwise, S is interior to the triangle, and all angles between segments $\overline{SP}$, $\overline{SQ}$, and $\overline{SR}$ are 120°.)

Given the vertices of a triangle P, Q, and R, we construct two lines, and the Steiner point will be their intersection. First, find the line l of slope $+1$ such that one vertex is on one side of l, the second is on l, and the third is on the other side. If it is the case that two of the vertices are on a line of slope $+1$, let l be that line. Second, find the analogous line l' of slope -1. $S = l \cap l'$ and also $S = B_M(P, Q) \cap B_M(Q, R) \cap B_M(P, R)$, the intersection of pairwise betweenness sets of the vertices. S may be at a vertex, on a side, or interior to the triangle. See FIGURE 3.

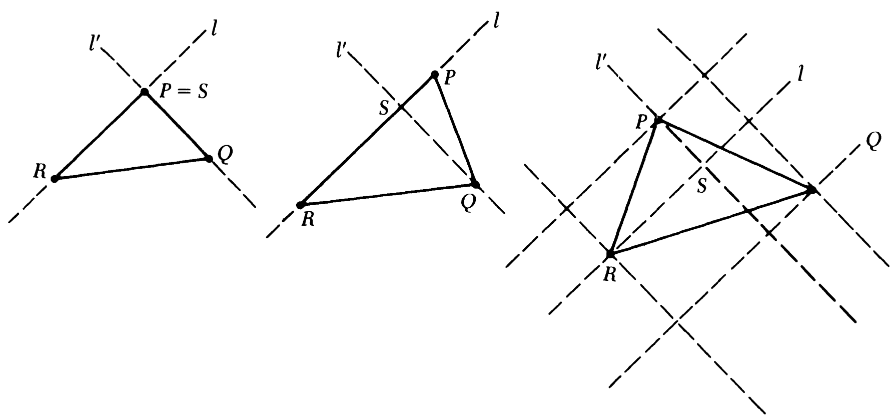


FIGURE 3

There is another Minkowski metric of interest in connection with Steiner's problem where the unit ball defining the metric is a regular hexagon of side 1 oriented so that two parallel sides are horizontal. The reader can verify that the (hexagon metric) betweenness sets, denoted $B_H(P, Q)$ for points P and Q, are parallelograms or line segments, of six basic types altogether, depending on the direction of the line segment $\overline{PQ}$, as shown in FIGURE 4.

The taxicab and the hexagon metric share an interesting feature that simplifies computation of distances. $d_T(P, Q)$ can be seen geometrically as the sum of the Euclidean lengths of the legs of the right triangle, legs parallel to the axes, for which $\overline{PQ}$ is the hypotenuse. Another way of looking at it would be to say that $d_T(P, Q)$ is equal to half of the Euclidean length of the perimeter of the betweenness set $B_T(P,Q)$, with the ordinary length of $\overline{PQ}$ in the limiting cases when P and Q are on the same vertical or horizontal line. Similarly, $d_H(P,Q)$ is equal to half of the Euclidean length of the perimeter of $B_H(P,Q)$, reducing to the usual length of the chord $\overline{PQ}$ when P and Q both lie on lines making angles of 0°, 60°, or 120° with the horizontal axis.

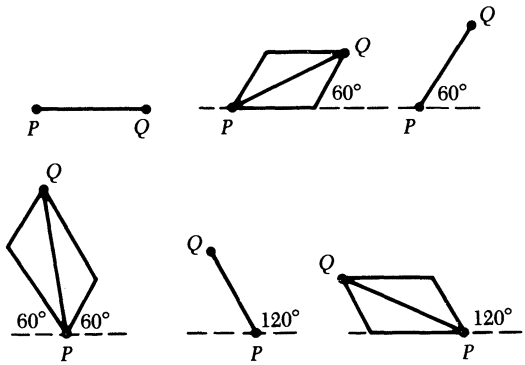


FIGURE 4
$B_H(P,Q)$

In the hexagon metric, it is not true that every triangle will have an optimal Steiner point metrically between each pair of vertices. For example, an equilateral triangle with horizontal base does not admit such a point, since metric betweenness for the vertices means lying on the line segment joining them. Suppose, however, we take any triangle with all angles less than 120°. The Steiner point S in the Euclidean metric is interior to $\triangle PQR$, and segments $\overline{PS}$, $\overline{QS}$, and $\overline{RS}$ make angles of 120° with one another. Now rotate and translate $\triangle PQR$ together with S so that the resulting $\triangle P'Q'R'$ has one vertex P' on the positive x-axis with S' at the origin. Since the hexagon distance from P' to Q' is the Euclidean length of $\overline{P'T}$, the "radius" of the oriented hexagon centered at P' such that Q' is on its perimeter, $d_H(P',Q')$ is equal to the sum of the Euclidean lengths of $\overline{P'S'}$ and $\overline{S'Q'}$, as in Figure 5. Then it is easily seen that the sum of the Euclidean lengths of $\overline{P'S'}$, $\overline{S'Q'}$, and $\overline{S'R'}$ is equal to half the length of the perimeter in the hexagon metric. The Steiner point S' of $\triangle P'Q'R'$ in the Euclidean metric is also the optimal Steiner point in the hexagon metric, because it is (hexagon metrically) between each pair of vertices, see Figure 6.

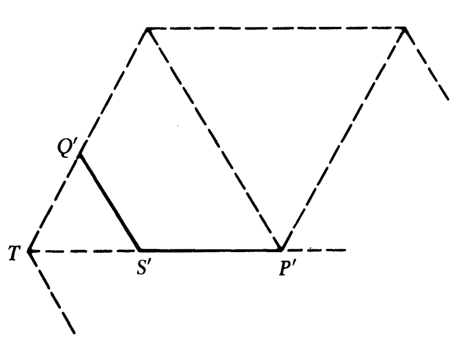


FIGURE 5

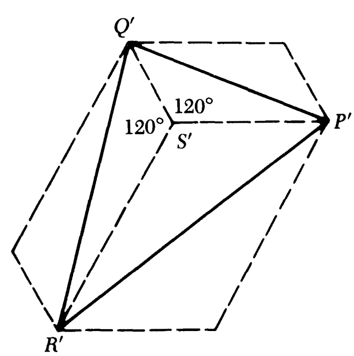


FIGURE 6

We will characterize betweenness sets for a general Minkowski metric in the plane. Then we can show that only when the defining unit ball is a parallelogram is it true that every triangle has an optimal Steiner point and every metric set of four points can be embedded. In what follows, d is a Minkowski metric in the plane and $\mathscr{B}$ is the defining centrally symmetric, closed, convex unit ball. We denote the space by $(d, \mathbb{R}^2)$.

The betweenness set $B(P,Q)$ in $(d, \mathbb{R}^2)$ with $\mathscr{B}$ centered at O depends on the direction of $\overline{PQ}$. Let $\overline{PQ}$ be parallel to $\overline{OX}$, a radius of $\mathscr{B}$ where $X \in \partial\mathscr{B}$, the boundary of $\mathscr{B}$. If X is a point of strict convexity of $\partial\mathscr{B}$, then $B(P,Q) = \overline{PQ}$, as $\mathscr{B}$ is centrally symmetric and a supporting line will not be changed in slope by a similarity transformation. For the same reason, if X is an extreme point of $\partial\mathscr{B}$, then $B(P,Q) = \overline{PQ}$ also, as the measuring balls will again touch at only one point. If X lies on a face, a straight line portion of the boundary, $B(P,Q)$ is a parallelogram with $\overline{PQ}$ as a diagonal and sides parallel to the "radii" of $\mathscr{B}$ from O to the endpoints of the face. The midpoint set, points in $B(P,Q)$ that are equidistant from P and Q, will be a chord parallel to that face and passing through the center of the parallelogram.

Theorem 1. *Every triangle in the plane has an optimal Steiner point in $(d, \mathbb{R}^2)$ if, and only if, $\mathscr{B}$ is a parallelogram.*

The technique for constructing a separating point in the max metric works just as well for the taxicab metric if horizontal and vertical lines are used instead of lines of slope $+1$ and -1. In fact, if the defining unit ball is any parallelogram $\mathscr{Q}$, proceeding in the same way using lines with the slopes of the diagonals of $\mathscr{Q}$, we find an optimal Steiner point as the intersection. The betweenness sets in such a metric are parallelograms with sides parallel to the diagonals of $\mathscr{Q}$.

If $\mathscr{B}$ is strictly convex, clearly no triangle has a point that is between each pair of vertices. As we have seen, in the hexagon metric some triangles have optimal Steiner points and others do not. When $\mathscr{B}$ is not a parallelogram, we will show how to construct a triangle that does not have an optimal Steiner point in that metric.

If $\mathscr{B}$ is not strictly convex, there are at least two parallel linear segments in $\partial\mathscr{B}$. Suppose there are just two, $\overline{P_1Q_1}$ and $\overline{P_2Q_2}$, as in Figure 7a. Choose any point P between P_1 and Q_2 on $\partial\mathscr{B}$ and draw the chord $\overline{OP}$. Any triangle with sides parallel to lines containing $\overline{OP_1}$, $\overline{OQ_2}$, and $\overline{OP}$ will not have an optimal Steiner point since the betweenness sets are all linear segments that connect vertices.

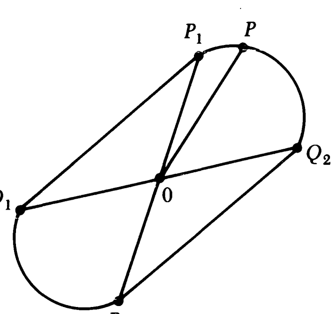


FIGURE 7a

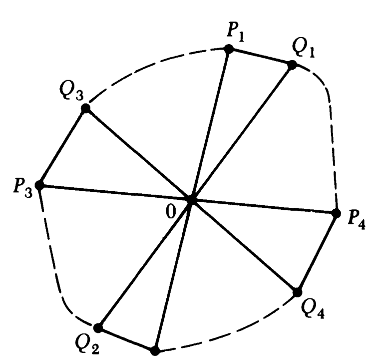


FIGURE 7b

Suppose now that there are at least two pairs of parallel faces in $\partial\mathscr{B}$, $\overline{P_1Q_1} \| \overline{P_2Q_2}$ and $\overline{P_3Q_3} \| \overline{P_4Q_4}$, but $\mathscr{B}$ is not a parallelogram. As before, the diagonals $\overline{P_1P_2}$, $\overline{Q_1Q_2}$, $\overline{P_3P_4}$, and $\overline{Q_3Q_4}$ all pass through O, since $\mathscr{B}$ is centrally symmetric (Figure 7b).

Among the endpoints of the faces are at least three pairs of distinct extreme points of $\partial\mathscr{B}$, because only one of the cases $P_1 = Q_3$ and $P_3 = Q_2$ is possible. Any triangle with sides oriented as the chords $\overline{OP_1}$, $\overline{OQ_1}$, and $\overline{OP_3}$ (or $\overline{OQ_3}$) will have the property that points metrically between vertices lie on the connecting segment. Such a triangle does not have an optimal Steiner point.

THEOREM 2. *An arbitrary metric set of four points can be embedded in* $(d, \mathbb{R}^2)$ *if, and only if,* $\mathscr{B}$ *is a parallelogram.*

We show that the metric set with distances $[1,1,2,1,2,1]$ can be embedded only when the Minkowski metric has a parallelogram as its defining unit ball. Suppose this set can be embedded in the plane with some Minkowski metric d where $\mathscr{B}$ is the unit ball. Taking P_1 as center, P_2 is on a ball of radius 1 and P_3 on a ball of radius 2. Now both P_2 and P_4 must be in $B(P_1, P_3)$, yet $d(P_2, P_4) = 2$ so that $\overline{P_1P_3}$ must not be in the direction of any extreme point or a point of strict convexity on $\partial\mathscr{B}$. If it is possible to embed this metric set, $\partial\mathscr{B}$ must not be strictly convex and P_3 must be in the relative interior of the linear segment of the boundary. P_2 and $P_4 \in B(P_1, P_3)$, a parallelogram with sides parallel to $\overline{P_1T}$ and $\overline{P_1T'}$.

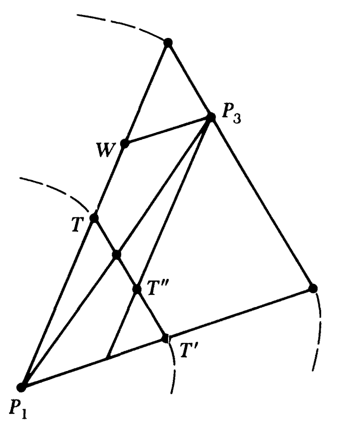


FIGURE 8a

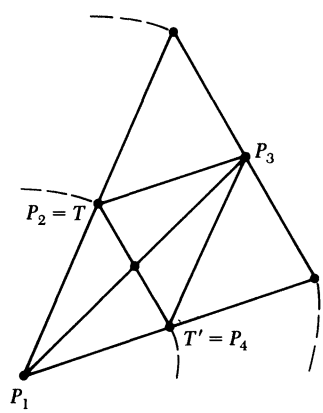


FIGURE 8b

Now suppose P_3 is not the midpoint of the face. P_2 and P_4 lie on $\overline{TT''}$, as in FIGURE 8a. P_1 and $P_3 \in B(P_2, P_4)$ so that $\overline{P_2P_4}$ cannot be in the direction of an extreme point or a point of strict convexity of $\partial\mathscr{B}$. $B(P_2, P_4)$ must also be a parallelogram. Whatever the placement of P_2 on $\overline{TT''}$, the angle measure of $\sphericalangle P_1P_2P_3$ must be greater than that of $\sphericalangle P_1WP_3$. Since $d(P_1, P_2) = d(P_2, P_3) = 1$, $\sphericalangle P_1P_2P_3$ is equal in measure to a central angle of $\mathscr{B}$. This cannot happen, as $\mathscr{B}$ is centrally symmetric and the sum of the angle measures of $\sphericalangle P_1P_2P_3$ and $\sphericalangle TP_1T'$ must not be more than 180°. Thus P_3 must be the midpoint of the face as in FIGURE 8b.

Reasoning as above about central angles of $\mathscr{B}$, P_2, and P_4 must be at the endpoints of this face of the unit ball. Then $\overline{P_1P_3}$ must be congruent to another face of the unit ball and $\sphericalangle P_1P_2P_3$ is equal in measure to a central angle. But the sum of the angle measures of $\sphericalangle P_4P_1P_2$ and $\sphericalangle P_1P_2P_3$ is 180°. Since $\mathscr{B}$ is centrally symmetric, $\partial\mathscr{B}$ must be a quadrilateral with sides parallel to $\overline{P_2P_4}$ and $\overline{P_1P_3}$. Briefly and intuitively, two points can be on the same face and also on opposite faces of a centrally symmetric convex set iff that convex set is a parallelogram.

To show the converse, let d_1 and d_2 be two Minkowski metrics in the plane, so that $(d_1, \mathbb{R}^2)$ and $(d_2, \mathbb{R}^2)$ are the respective metric spaces with $\mathscr{B}_1$ and $\mathscr{B}_2$ the defining unit balls. If there is a nonsingular transformation of the plane mapping $\mathscr{B}_1$

onto $\mathscr{B}_2$, then $(d_1, \mathbb{R}^2)$ is isometric to $(d_2, \mathbb{R}^2)$ [1]. Any metric set of four points can be embedded in the plane endowed with the max metric [4], and any parallelogram centered at the origin is the image of a square under some nonsingular linear transformation of the plane. Thus a given metric set of four points can be embedded in the plane with a Minkowski metric where the unit ball is a parallelogram.

Appendix The construction to embed a metric set of four points in the plane is the following: Let $F = \{P_1, P_2, P_3, P_4\}$ and $d_{ij} = d(P_i, P_j)$ for $i, j = 1, 2, 3, 4$. The set of six distances is $[d_{12}, d_{23}, d_{31}, d_{14}, d_{24}, d_{34}]$. We find the maximum sum of disjoint distances, $a = \max(d_{12} + d_{34}, d_{23} + d_{14}, d_{13} + d_{24})$. Relabel the points if necessary so that $a = d_{12} + d_{34}$. Then introduce P_0, called a separating point of F, so that $d_{12} = d_{01} + d_{02}$ and $d_{34} = d_{03} + d_{04}$ (extending the above notation in an obvious way so that i or j may be 0). Then an isometric embedding of F in the plane with the max metric consists of the following points:

$$\begin{aligned}
&P_0 = (0,0);\\
&P_1 = (d_{01}, d_{14} - d_{04});\\
&P_2 = (-d_{02}, d_{03} - d_{23});\\
&P_3 = (d_{01} - d_{13}, d_{03});\text{ and}\\
&P_4 = (d_{24} - d_{02}, -d_{04})
\end{aligned}$$

For example, suppose we have the set of six distances $[3,4,5,2,2,5]$ ordered as above. Check out the sums of disjoint distances: They are $3+5$, $4+2$, and $5+2$. Then $a = 8$ and let $d_{01} = 2$, $d_{02} = 1$, $d_{03} = 3$, and $d_{04} = 2$; the embedding will not be unique, as other choices as possible. Here the embedding of F and a separating point will be $P_0 = (0,0)$; $P_1 = (2,0)$, $P_2 = (-1,-1)$; $P_3 = (-3,3)$; $P_4 = (1,-2)$.

REFERENCES

1. Russell V. Benson, *Euclidean Geometry and Convexity*, McGraw-Hill Book Co., New York, 1966, p. 230.
2. Richard Courant and Herbert Robbins, *What is Mathematics?*, Oxford University Press, Inc., Fair Lawn, NJ, pp. 354–361, pp. 377–379.
3. Sister Rita M. Ehrmann, Minimal surfaces rediscovered, *Math Teacher*, Vol. 69, no. 2 (1976), 146–152.
4. Dorothy Wolfe, Imbedding a finite metric set in an n-dimensional Minkowski space, *Nederl. Acad. Wetensch. Proc. Ser.* A, Vol. 70 (1967), 136–140.

A Biased Random Walk with Symmetry

J. N. BOYD
P. N. RAYCHOWDHURY
Virginia Commonwealth University
Richmond, VA 23284-2014

Let us consider a random walk on an equilateral array of points and connecting segments such as that shown in FIGURE 1. In the triangular array, which we denote by its vertices as $\Delta V_1V_2V_3$, each interior point Q has six adjacent points A_i and each boundary point B (which is not a vertex) has two adjacent points on its side, C_1, C_2.

We wish to find the probability of reaching a particular vertex, which we may as well take to be V_1, from any starting point X. That is, we set for ourselves the problem of computing $p(X)$, the probability of reaching V_1 from starting point X subject to the conditions to follow.

First, suppose that $X = Q$, an interior point of the array. That situation is illustrated in FIGURE 2, where we also see that $\overline{A_4A_1} \| \overline{V_1V_2}$, $\overline{A_5A_2} \| \overline{V_1V_3}$, and $\overline{A_6A_3} \| \overline{V_2V_3}$. Then imagine that the walk drifts toward V_1 in a manner symmetric with respect to motions toward sides $\overline{V_1V_2}$ and $\overline{V_1V_3}$. Letting p_k be the probability of moving from Q to A_k, we take $p_4 = p_5 = s/3$ ($\frac{1}{2} < s \le 1$), $p_1 = p_2 = (1-s)/3$, and $p_3 = p_6$. (If $s = \frac{1}{2}$, we would be back to a simpler random walk that we have previously discussed **[1]**.) Our aim is to produce a solution in convex (or barycentric) coordinates. The original impetus for our work was provided by the development of random walks on rectangular networks given by Doyle and Snell **[2]**.

We can write

$$p(Q) = \frac{(1-s)}{3}[p(A_1) + p(A_2)] + \frac{s}{3}[p(A_4) + p(A_5)] + \frac{1}{6}[p(A_3) + p(A_6)] \tag{1}$$

where $p(A_i)$ is the probability of reaching V_1 from A_i. We take the boundary of $\Delta V_1V_2V_3$ to be a trap with $p(V_2) = p(V_3) = 0$ and $p(V_1) = 1$. In other words, the random walk can never reach V_1 from either V_2 or V_3. If the walk reaches a point of $\overline{V_2V_3} - \{V_2, V_3\}$, the probability of moving along the side toward V_2 is $\frac{1}{2}$ as is the probability of moving toward V_3. The probability of moving back into the interior is zero. Thus the probability of reaching V_1 from any point of the entire side $\overline{V_2V_3}$ is zero.

If $X = B \in \overline{V_1V_2} \cup \overline{V_1V_3} - \{V_1, V_2, V_3\}$ as in FIGURE 1, the probability of moving along a side toward V_1 is s and the probability of moving along a side toward the other vertex is $1-s$. The walk cannot return to an interior point and the probability of reaching V_1 from B is

$$p(B) = sp(C_2) + (1-s)p(C_1). \tag{2}$$

With the probability of reaching V_1 defined at V_1 by $p(V_1) = 1$ and on $\overline{V_2V_3}$ as zero and with the average value properties given by (1) and (2), we now have the problem of writing an expression for the probability of reaching V_1 from any starting point X of the array. We will write that expression in terms of the convex coordinates of X with respect to V_1, V_2, V_3.

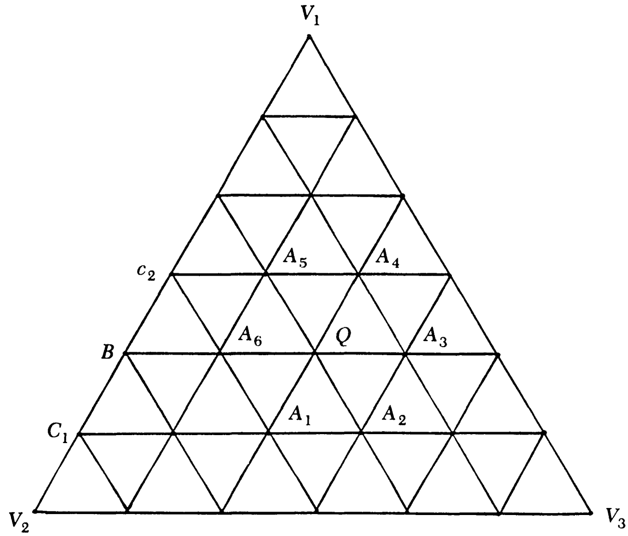


FIGURE 1
An equilateral grid for a random walk.

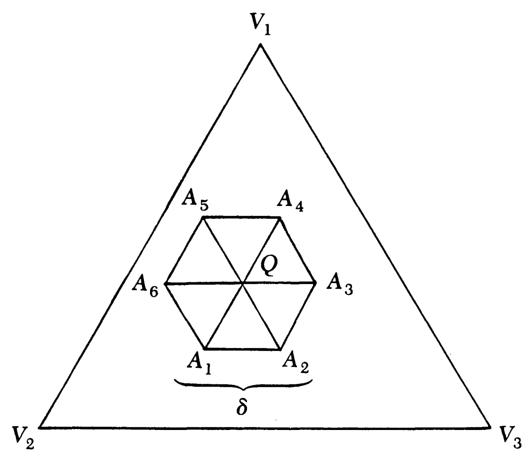


FIGURE 2
Interior point Q as starting point for a random walk with drift toward V_1.

Solution. Let us suppose that an equilateral array of points has $n+1$ points on each side. Then there are $\frac{1}{2}(n+1)(n+2)$ points in the array and the separation between nearest neighbors is $\delta = 1/n$. A typical point of the array has convex coordinates with respect to V_1, V_2, V_3 given by $(\alpha_1, \alpha_2, \alpha_3) = (a/n, b/n, c/n)$ where a, b, and c are nonnegative integers with $a+b+c=n$ (see [3]). If point Q (as shown in FIGURE 2) has convex coordinates $(a/n, b/n, c/n)$, then the nearest neighbors have convex coordinates as listed in TABLE 1.

Next we need to find a function of the convex coordinates of Q that satisfies the average value properties (1) and (2) on the array. Such a function ought to involve some ratio of the probabilities of moving toward and away from V_1. A likely candidate in our problem is $f(Q) = (1 - s/s)^{n\alpha_1} = (1 - s/s)^a$ where the convex coordinates of Q are as given above.

TABLE 1. The convex coordinates of Q and its nearest neighbors.

Point	$\alpha_1, \alpha_2, \alpha_3$	$n\alpha_1$
Q	$\frac{a}{n}, \frac{b}{n}, \frac{c}{n}$	a
A_1	$\frac{a}{n} - \frac{1}{n}, \frac{b}{n} + \frac{1}{n}, \frac{c}{n}$	$a-1$
A_2	$\frac{a}{n} - \frac{1}{n}, \frac{b}{n}, \frac{c}{n} + \frac{1}{n}$	$a-1$
A_3	$\frac{a}{n}, \frac{b}{n} - \frac{1}{n}, \frac{c}{n} + \frac{1}{n}$	a
A_4	$\frac{a}{n} + \frac{1}{n}, \frac{b}{n} - \frac{1}{n}, \frac{c}{n}$	$a+1$
A_5	$\frac{a}{n} + \frac{1}{n}, \frac{b}{n}, \frac{c}{n} - \frac{1}{n}$	$a+1$
A_6	$\frac{a}{n}, \frac{b}{n} + \frac{1}{n}, \frac{c}{n} - \frac{1}{n}$	a

From our table, we can evaluate f at the seven crucial points, Q and its nearest neighbors:

$$f(Q) = \left(\frac{1-s}{s}\right)^a$$

$$f(A_1) = f(A_2) = \left(\frac{1-s}{s}\right)^{a-1}$$

$$f(A_3) = f(A_6) = \left(\frac{1-s}{s}\right)^a$$

$$f(A_4) = f(A_5) = \left(\frac{1-s}{s}\right)^{a+1}.$$

Substituting these values into the right-hand side of (1), we obtain

$$\frac{1}{3}\left[\frac{2(1-s)^a}{s^{a-1}} + \frac{2(1-s)^{a+1}}{s^a} + \frac{(1-s)^a}{s^a}\right] = \left(\frac{1-s}{s}\right)^a = \left(\frac{1-s}{s}\right)^{n\alpha_1},$$

thereby substantiating our claim for our candidate. It follows from further, simple substitutions that

$$p(Q) = M\left(\frac{1-s}{s}\right)^{n\alpha_1} + N$$

(where M and N are constants) must also satisfy (1).

If point B of FIGURE 1 has convex coordinates $(a/n, b/n, 0)$, then the convex coordinates of C_1 and C_2 are

$$\left(\frac{a}{n} - \frac{1}{n}, \frac{b}{n} + \frac{1}{n}, 0\right) \quad \text{and} \quad \left(\frac{a}{n} + \frac{1}{n}, \frac{b}{n} - \frac{1}{n}, 0\right),$$

respectively, and it is an easy matter to show that $f(B) = (1 - s/s)^a$ also satisfies (2) as does $p(B) = M(1 - s/s)^{n\alpha_1} + N$. Had B been a point of $\overline{V_1V_3}$, the same result would hold.

We can now adjust M and N so that $p(V_1) = 1$ and $p(V_2) = p(V_3) = 0$. Since

$\alpha_1 = 1$ at V_1, we require that

$$M\left(\frac{1-s}{s}\right)^n + N = 1. \tag{3}$$

Since $\alpha_1 = 0$ at V_2 and V_3, we also require that

$$M + N = 0. \tag{4}$$

Solving (3) and (4) for M and N, we find that

$$M = \frac{-1}{1-\left(\frac{1-s}{s}\right)^n} < 0 \quad \text{and}$$

$$N = \frac{1}{1-\left(\frac{1-s}{s}\right)^n} > 0.$$

Then

$$p(X) = \frac{1-\left(\frac{1-s}{s}\right)^{n\alpha_1}}{1-\left(\frac{1-s}{s}\right)^n} \tag{5}$$

for X any point of the array.

The uniqueness of the solution Equations (1) and (2) give an average value property for the probability function $p(X)$ and would become defining equations for an harmonic function if s were permitted to take on the value $\frac{1}{2}$. By reasoning analogous to that for harmonic functions, the average value property implies that $p(X)$ acquires both its maximum and minimum values on the boundary of $\Delta V_1V_2V_3$.

Thus, if $p(X)$ and $p'(X)$ both satisfy (1) and (2) and agree on the boundary of $\Delta V_1V_2V_3$, then their difference $p(X) - p'(X)$ also satisfies the equations and has the value 0 at all boundary points. It follows that $p(X) - p'(X)$ has both maximum and minimum values of 0. Therefore $p(X) = p'(X)$ at all points of the array and our solution is unique.

REFERENCES

1. J. N. Boyd and P. N. Raychowdhury, Discrete Dirichlet problems, convex coordinates, and a random walk on a triangle, *College Math. J.* 20 (1989), 385–92.
2. P. G. Doyle and J. L. Snell, *Random Walks and Electrical Networks*, MAA, Washington, DC, 1984.
3. J. N. Boyd, J. S. Cooper, and P. N. Raychowdhury, Euclidean geometry from classical mechanics, *Internat. J. Math. Ed. Sci. Tech.*, 20 (1989), 383–90.

Euler's Theorem for Polynomials

WILLIAM P. WARDLAW*
U. S. Naval Academy
Annapolis, MD 21402

The similarity of the theory of divisibility for integers and that for polynomials is striking. For example, the Euclidean algorithm, the formula

$$(a,b) = ca + db \tag{1}$$

for the greatest common divisor (a, b) of a and b, the arithmetic of elements modulo a fixed element m, and the criterion

$$a \text{ is invertible modulo } m \quad \text{if and only if} \quad (a,m) = 1, \tag{2}$$

all apply equally well for integers, or for polynomials, a, b, c, d, and m. Both theories measure the distance of an element from a unit, by the absolute value $|a|$ of an integer a, or by the degree ∂a of a polynomial a. (Of course, the properties mentioned are shared by all Euclidean domains.)

This similarity suggests a polynomial analogue to Euler's pretty theorem on modular arithmetic:

EULER'S THEOREM. *If a and m are integers with $(a,m) = 1$ and $\phi(m) = |\{k \in Z: 0 < k < |m|, (k,m) = 1\}|$, then*

$$a^{\phi(m)} \equiv 1 (\operatorname{mod} m).$$

Such an analogue does indeed exist, and the analogy is almost exact!

THEOREM 1 (Euler's Theorem for Polynomials). *Let $m \in K[x]$, where K is a finite field, and let $\varphi(m) = |\{f \in K[x]: 0 \leqslant \partial f < \partial m$ and $(f,m) = 1\}|$. Then for any $f \in K[x]$ with $(f,m) = 1$,*

$$f^{\varphi(m)} \equiv 1 (\operatorname{mod} m). \tag{3}$$

Proof. Let $K_m = K[x]/(m)$. Then K_m is a ring. Let K_m^* be the group of invertible elements of K_m. Then $(f,m) = 1$ implies that $f_m = f + (m) \in K_m^*$. Since $|K_m^*| = \varphi(m)$, it follows by Lagrange's theorem that

$$f^{\varphi(m)} \equiv f_m^{\varphi(m)} \equiv 1 (\operatorname{mod} m).$$

It should come as no surprise that an immediate corollary is an analogue to Fermat's Little Theorem. There are q^d polynomials of degree less than d over the q element field K, one of which is 0. Thus it is clear that a polynomial g over K of degree d is irreducible if and only if $K_g^* = K_g \setminus \{0\}$, that is, if and only if $\varphi(g) = q^d - 1$. Replacing m by g in (3) and multiplying by f (to handle the $f \equiv 0 \pmod g$ case) gives

COROLLARY 1.1 (Fermat's Little Theorem for Polynomials). *Let K be a field of q elements and let g be an irreducible polynomial over K of degree d. Then $f \in K[x]$*

*Research supported in part by Naval Academy Research Council and by Naval Research Laboratory.

implies

$$f^{q^d} \equiv f(\bmod g). \tag{4}$$

Note that taking $f = x$ in (4) shows that if g is an irreducible polynomial of degree d over a field K of q elements, then

$$g | x^{q^d} - x. \tag{5}$$

This result can be useful. For example, the negation of (5) can be used to show that g is reducible in $K[x]$. If g is irreducible and $g(0) \neq 0$, (5) can be written as the congruence

$$x^{q^d - 1} \equiv 1(\bmod g). \tag{6}$$

The powers of x modulo a fixed polynomial f are fairly easy to calculate and can give considerable information about the factorization of f. We conclude with two examples showing how the above can be applied to investigate the irreducibility of polynomials over finite fields.

Example 1. Let $K = GF(2)$ and $f = x^7 + x + 1$. Taking congruences modulo f, $x^7 \equiv x + 1$, $x^8 \equiv x^2 + x$, $x^{16} \equiv x^4 + x^2$, $x^{32} \equiv x^8 + x^4 \equiv x^4 + x^2 + x$, $x^{64} \equiv x^8 + x^4 + x^2 \equiv x^4 + x$, $x^{128} \equiv x^8 + x^2 \equiv x$, and $x^{127} \equiv 1$. Since the order of x modulo f is $127 = 2^7 - 1$, it follows that $\varphi(f) = 127$, and f is irreducible by the remarks preceding Corollary 1.1.

Example 2. Let $K = GF(2)$ and $g = x^7 + x^6 + x^2 + x + 1$. Taking congruences modulo g for x^k with $k = 7$, 8, 10, 12, 16, 32, 64, and 128, one obtains (after some tedium) $x^{128} \equiv x^5 + x^4 \not\equiv x$. Hence, g is reducible, by (5). (Actually, $g = (x^3 + x^2 + 1)(x^4 + x + 1)$.)

These notions are developed further in [2].

After this note was written, the author was made aware of Lemma 3.69 on page 122 of [1] and the portion of Note 4 in the first paragraph of page 137 of [1], which indicate that some of the ideas presented here were known to Dedekind. Indeed, many of these notions appear in [1].

REFERENCES

1. R. Lidl and H. Niederreiter, *Finite Fields*, Addison-Wesley, Reading, MA, 1983.
2. William P. Wardlaw, *Euler's Theorem for Polynomials*, NRL Report, Naval Research Laboratory, Washington, DC.

The nth Power of a 2×2 Matrix

KENNETH S. WILLIAMS
Carleton University
Ottawa, Ontario, Canada K1S 5B6

Let α and β be the eigenvalues of the matrix

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We present a very simple derivation of the formula for A^n (n is a positive integer) in terms of α and β. As A satisfies its characteristic polynomial, we have

$$A^2-(\alpha+\beta)A+\alpha\beta I=0, \quad \text{where } I=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{1}$$

The matrices X, Y, Z defined by

$$\begin{cases} X=\dfrac{A-\beta I}{\alpha-\beta},\ Y=\dfrac{A-\alpha I}{\beta-\alpha}, & \text{if } \alpha\neq\beta, \\ Z=A-\alpha I, & \text{if } \alpha=\beta, \end{cases}$$

satisfy

$$\begin{cases} X^2=X,\ XY=YX=0,\ Y^2=Y, & \text{if } \alpha\neq\beta, \\ Z^2=0, & \text{if } \alpha=\beta, \end{cases}$$

so that for $k \geqslant 2$

$$\begin{cases} X^k=X,\ Y^k=Y, & \text{if } \alpha\neq\beta, \\ Z^k=0, & \text{if } \alpha=\beta. \end{cases}$$

Hence we have

$$A^n=\begin{cases} (\alpha X+\beta Y)^n=\alpha^n X^n+\beta^n Y^n=\alpha^n X+\beta^n Y, & \text{if } \alpha\neq\beta, \\ (\alpha I+Z)^n=\alpha^n I+n\alpha^{n-1}Z, & \text{if } \alpha=\beta, \end{cases}$$

giving

$$A^n=\begin{cases} \alpha^n\left(\dfrac{A-\beta I}{\alpha-\beta}\right)+\beta^n\left(\dfrac{A-\alpha I}{\beta-\alpha}\right), & \text{if } \alpha\neq\beta, \\ \alpha^{n-1}(nA-(n-1)\alpha I), & \text{if } \alpha=\beta. \end{cases} \tag{2}$$

If the matrix A is invertible ($\alpha\neq 0, \beta\neq 0$), it is easy to see that (2) holds for all integral values of n.

If A is real but its eigenvalues $\alpha=p+iq$ and $\beta=p-iq$ are nonreal ($q\neq 0$) with some power of them real, say $(p+iq)^m=(p-iq)^m=r$, then, by (2), we have $A^m=rI$.

In the case of distinct eigenvalues, the reader will recognize the matrices X and Y as supplementary projections ($X+Y=I$) and

eigenspace of α = range of X = null space of Y,
eigenspace of β = null space of X = range of Y.

Another Elementary Proof of Heron's Formula

SIDNEY H. KUNG
Jacksonville University
Jacksonville, FL 32211

Modern proofs of Heron's formula for the area of a triangle are more algebraic or trigonometric than geometric [1], [2], [3]. This note gives a simple geometric proof of the formula

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

where K denotes the area of a triangle ABC (FIGURE 1) with sides a, b, c, and $s=(a+b+c)/2$.

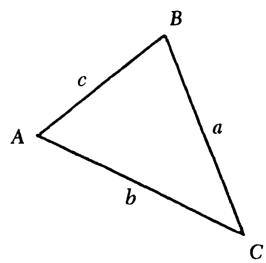


FIGURE 1

The incircle of the triangle (FIGURE 2) is centered at I with radius r. One of the excircles is shown with center I' and radius r'. By combining the areas of triangles IAB, IBC, and ICA, we get

$$K=rs. \tag{1}$$

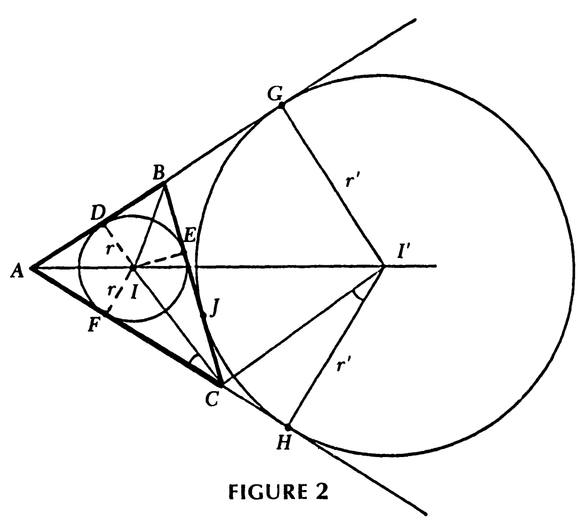


FIGURE 2

We see that tangents AG and AH from the exterior point A are equal; similarly, $BG=BJ$, and $CH=CJ$. Since

$$\begin{aligned} AG + AH &= (AD + DB + BG) + (AF + FC + CH) \\ &= (AD + DB + BJ) + (AF + FC + CJ) \\ &= (AD + DB) + (BJ + CJ) + (AF + FC) \\ &= c + a + b = 2s, \end{aligned}$$

$AG = AH = s$. Since $AD = AF = (a + b + c - 2(BE + CE))/2 = s - a$, we obtain, by logical symmetry, $FC = s - c$. Furthermore, $rt.\triangle\, ADI \sim rt.\triangle\, AGI'$, so

$$\frac{DI}{AD} = \frac{GI'}{AG}, \qquad \text{or} \qquad \frac{r}{s-a} = \frac{r'}{s}. \tag{2}$$

Note that the angle bisectors CI and CI' are perpendicular to each other; and $\angle I'CJ + \angle JCI = \angle I'CH + \angle ICF = \angle I'CH + \angle CI'H = \pi/2$. Therefore, $\angle ICF = \angle CI'H$, and $rt.\triangle\, ICF \sim rt.\triangle\, CI'H$. Thus, we have

$$\frac{FI}{FC} = \frac{CH}{HI'}, \qquad \text{or} \qquad \frac{r}{s-c} = \frac{s-b}{r'}. \tag{3}$$

Multiplying (2) and (3) we find

$$r = \sqrt{(s-a)(s-b)(s-c)/s}\,. \tag{4}$$

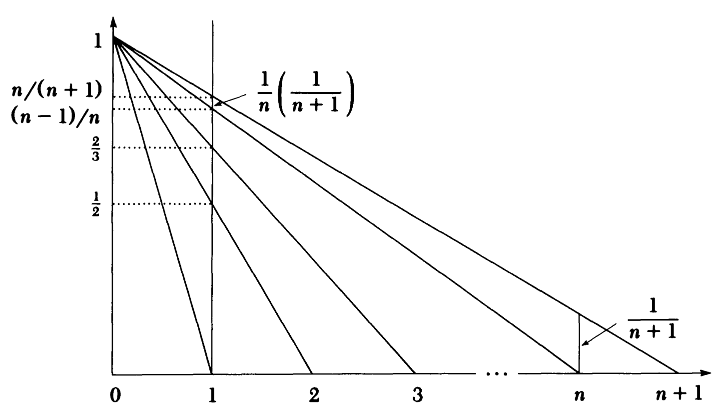
Substituting (4) in (1) gives

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

REFERENCES

1. C. Alperin, Heron's area formula, *College Math. J.* 18 (1987), 137–138.
2. David E. Dobbs, Proving Heron's formula tangentially, *College Math. J.* 15 (1984), 252–253.
3. William Dunham, An ancient/modern proof of Heron's formula, *Mathematics Teacher* 78 (1985), 258–259.

Proof Without Words: $\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{n(n+1)} = \frac{n}{n+1}$.



—Roman W. Wong
Washington and Jefferson College
Washington, PA 15301

$$\begin{aligned} AG + AH &= (AD + DB + BG) + (AF + FC + CH) \\ &= (AD + DB + BJ) + (AF + FC + CJ) \\ &= (AD + DB) + (BJ + CJ) + (AF + FC) \\ &= c + a + b = 2s, \end{aligned}$$

$AG = AH = s$. Since $AD = AF = (a + b + c - 2(BE + CE))/2 = s - a$, we obtain, by logical symmetry, $FC = s - c$. Furthermore, $rt.\triangle\, ADI \sim rt.\triangle\, AGI'$, so

$$\frac{DI}{AD} = \frac{GI'}{AG}, \qquad \text{or} \qquad \frac{r}{s-a} = \frac{r'}{s}. \tag{2}$$

Note that the angle bisectors CI and CI' are perpendicular to each other; and $\angle I'CJ + \angle JCI = \angle I'CH + \angle ICF = \angle I'CH + \angle CI'H = \pi/2$. Therefore, $\angle ICF = \angle CI'H$, and $rt.\triangle\, ICF \sim rt.\triangle\, CI'H$. Thus, we have

$$\frac{FI}{FC} = \frac{CH}{HI'}, \qquad \text{or} \qquad \frac{r}{s-c} = \frac{s-b}{r'}. \tag{3}$$

Multiplying (2) and (3) we find

$$r = \sqrt{(s-a)(s-b)(s-c)/s}\,. \tag{4}$$

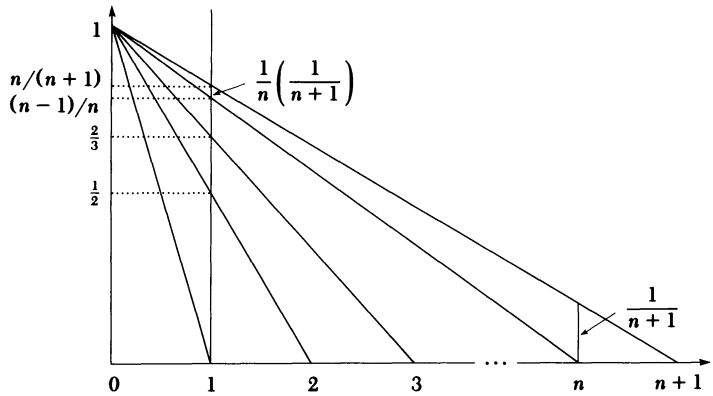
Substituting (4) in (1) gives

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

REFERENCES

1. C. Alperin, Heron's area formula, *College Math. J.* 18 (1987), 137–138.
2. David E. Dobbs, Proving Heron's formula tangentially, *College Math. J.* 15 (1984), 252–253.
3. William Dunham, An ancient/modern proof of Heron's formula, *Mathematics Teacher* 78 (1985), 258–259.

Proof Without Words: $\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{n(n+1)} = \frac{n}{n+1}$.



—Roman W. Wong
Washington and Jefferson College
Washington, PA 15301

Pairs of Points: Antigonal, Isogonal, and Inverse

JAN VAN YZEREN
Technological University
Eindhoven, NL 5600 MB

For many centuries Euclid's geometry maintained its classic style. But, suddenly, the slogans "Down with Euclid" and "Death to triangles" broke the spell; then the scientific bathwater nearly washed away one of mankind's finest babies.

As a happy coincidence "Death to triangles" misses the point here, because, in this article, surpassing the triangle is one of the first moves. However, admittedly, the old geometry needs some changes in style to conform with modern mathematics. One of these changes is the consistent use of a slightly modified concept of angles. The idea is by no means new; on the contrary, "directed angles between undirected lines" were strongly advocated in **[1]** and **[2]**. Nevertheless, the concept did not take root and some books even completely ignored it **[3]**. Signs of life can be found in **[4]**, **[5]**, and **[6**, p. 9**]**. Surely, the fundamental modern character of the concept and its very practical usefulness may modify parts of Euclid's geometry in a way that will please even the advocates of the above slogans.

Anyway, the various theorems that are treated in this article could not be formulated and proved adequately without the consistent use of the modern idea of angles. Among the theorems some special results seem to be new, although they are rooted in well-known ideas.

1. ABC of angles The best way to get acquainted with the modified concept of angles is to look at some simple pictures. Obviously, an undirected line, when rotated through any multiple of π, returns to its original direction. Let rotations be measured in the positive sense, then $\angle ab$ is defined as a rotation of a that makes it parallel to b and is measured modulo π. Therefore, $\angle ba = -\angle ab \bmod \pi$ and $\angle ab + \angle bc = \angle ac \bmod \pi$. Henceforth, "mod π" will be omitted; if points on the lines are used, now it should be clear that, for instance, $\angle AOB = -\angle BOC = \angle COD = -\angle DOA$.

The angles of a triangle ABC give $\angle ABC + \angle BCA + \angle CAB = 0$. For the opposite sides AB and CD of a quadrangle $ABCD$, it is found that

$$\begin{aligned}\angle(AB\ \ CD) &= \angle ABC + \angle BCD = \angle BAD + \angle ADC\\ &= \angle ABD + \angle BDC = \angle BAC + \angle ACD.\end{aligned}$$

Clearly, such equations can be found automatically by concatenation of angles. In principle, there is no need to make pictures. However, the case of cyclic quadrangles

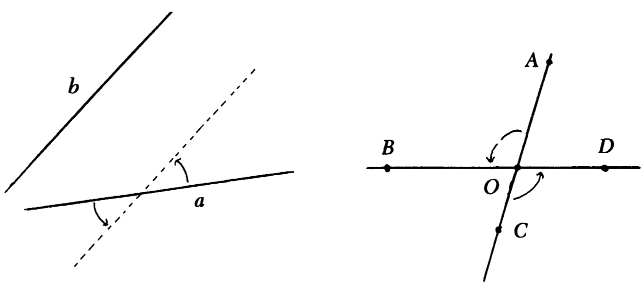


FIGURE 1

needs special attention, because they show the core property of the modified concept of angles: If P_1, P_2, P_3, P_4 are on a circle, then $\angle P_h P_i P_k = \angle P_h P_j P_k$ for every permutation (h,i,j,k) of $(1,2,3,4)$. Additionally, the tangents p_h at P_h give $\angle(p_h P_h P_k) = \angle P_h P_i P_k$ and $\angle(P_k P_h p_h) = \angle P_k P_i P_h$ (FIGURE 2).

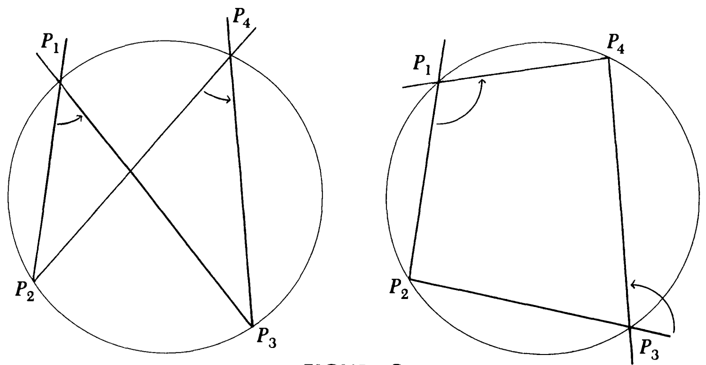


FIGURE 2

2. Antigonal pairs A given triangle $A_1A_2A_3$ determines pairs of points (X, X') by opposite angles:

$$\angle A_h X A_k + \angle A_h X' A_k = 0; \qquad h,k \in \{1,2,3\}.$$

Such a pair X and X' are called *antigonal* points [7]. A simple construction (FIGURE 3) leads from a given X to X': Let A_1X meet the circle XA_2A_3 at U, then $\angle A_1XA_2$, $\angle A_2XA_3$ and $\angle A_3XA_1$ are the angles of $\triangle UA_2A_3$; reflecting U in A_2A_3 gives the opposite triangle $U'A_2A_3$ whose angles, in the reverse way, lead to X'.

A special pair is well known: Equilateral $U_1A_2A_3$ and $U_1'A_2A_3$ (FIGURE 4) give the points T and T'—of Torricelli—that view A_1, A_2, and A_3 under angles of, respectively, 120° and $-120°$ ($= 60°$ mod 180°). The point T is also named after Fermat [8].

Many other antigonal pairs can be found. If A_0 is the orthocenter of $\triangle A_1A_2A_3$ and X is, for example, on A_1A_0, then X' is found by reflection in A_2A_3. Furthermore, if X is on the circumcircle of $\triangle A_1A_2A_3$, then $X' = A_0$ (in FIGURE 3: $U = A_1$). Obviously, A_0 is an intriguing point and the following theorem even puts A_0 on an equal footing with A_1, A_2, and A_3.

THEOREM. *Let $A_1A_2A_3A_0$ be an orthocentric set ($A_hA_i \perp A_jA_k$), then every point $X \neq A_i$, $i \in \{0,1,2,3\}$, has an antigonal point X' with $\angle A_hXA_k + \angle A_hX'A_k = 0$; $h,k \in \{0,1,2,3\}$.*

Proof. The triangle $B_1B_2B_3$ (FIGURE 5) is formed by the perpendicular bisectors of A_1A_0, A_2A_0, and A_3A_0. Then, conversely, the latter are the perpendicular bisectors of, respectively, B_2B_3, B_3B_1, and B_1B_2 (because $C_1B_2 = C_2C_3 = B_3C_1$, etc.). Now, for every point X not on A_1A_0 there is a circle XA_1A_0 with center M_1 on B_2B_3. Reflecting M_1 in A_1A_0 gives M_1' on B_2B_3 with $M_1B_2 = B_3M_1'$. Similarly M_2, M_2', M_3, M_3' are defined. Since M_1, M_2, and M_3 are points of the perpendicular bisector of XA_0, they are collinear and, therefore, make

$$\frac{M_1B_2}{M_1B_3}\,\frac{M_2B_3}{M_2B_1}\,\frac{M_3B_1}{M_3B_2} = 1$$

(Menelaus' theorem). Clearly, the analogous M'-product equals unity too. Therefore M_1', M_2', and M_3' are collinear, that is, the reflections of the circles XA_1A_0, XA_2A_0,

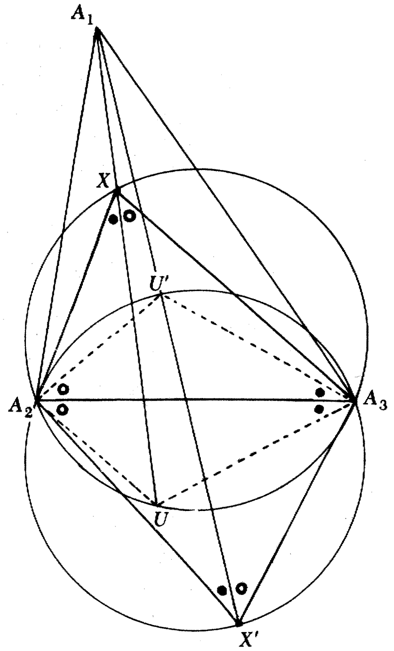


FIGURE 3

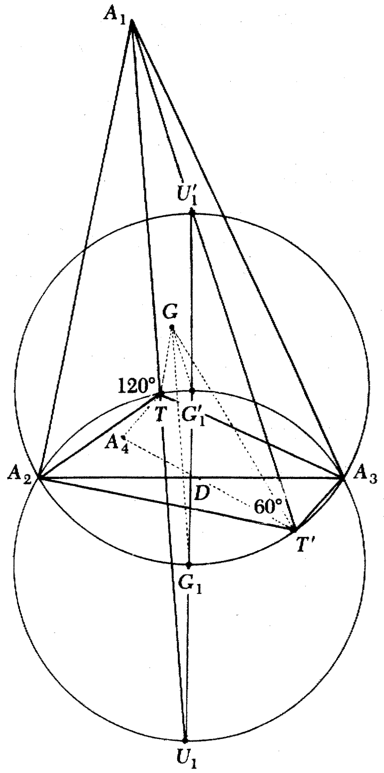


FIGURE 4

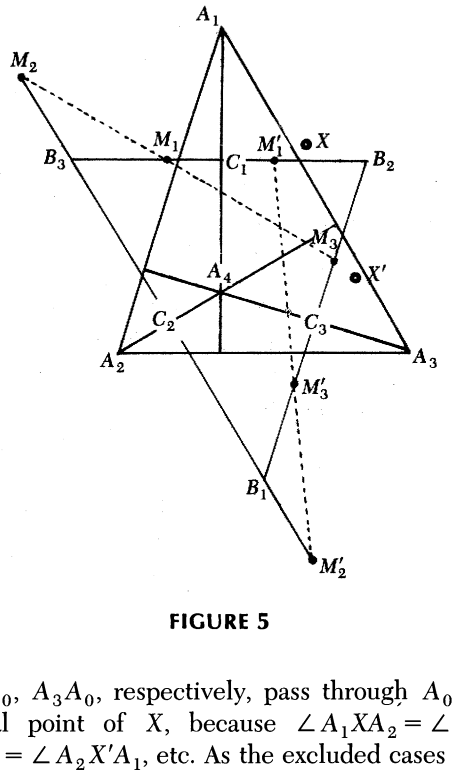


FIGURE 5

XA_3A_0 in A_1A_0, A_2A_0, A_3A_0, respectively, pass through A_0 and a point X'. The latter is the antigonal point of X, because $\angle A_1XA_2 = \angle A_1XA_0 + \angle A_0XA_2 = \angle A_0X'A_1 + \angle A_2X'A_0 = \angle A_2X'A_1$, etc. As the excluded cases (X on A_1A_0, etc.) are trivial (see above) the proof is complete.

This theorem agrees with the internal symmetry of $A_1A_2A_3A_0$ that gives in four ways a triangle with an orthocenter. The four circumcircles are mutual reflections in the six sides. The points of circumcircle $A_hA_jA_k$ are the antigonal images of A_i. The exclusion of A_i from the above theorem is motivated by its multiple images. Again, when scaled down in the ratio $\frac{1}{2}$, from its own orthocenter, each circumcircle becomes the nine-point circle, i.e., the circle through the midpoints of the six sides and the three pedal points $(A_hA_i \cap A_jA_k)$.

There is a special reason to look back to the Torricelli points (FIGURE 4), because the centroid G of $\triangle A_1A_2A_3$ appears to extend the angular T, T'-contrast to *five* points, namely, A_1, A_2, A_3, A_0, and G. Clearly, a proof of $\angle GTA_1 = \angle A_1T'G$ is sufficient. It is found by using the fact that A_1T is parallel to GG_1 (on the median A_1D: $GD/A_1D = 1/3 = G_1D/U_1D$) and $U_1T \perp TG_1'$. This, indeed, implies that G_1G is the perpendicular bisector of TG_1'. Hence, $\angle G_1GG_1' = \angle TGG_1 = \angle GTA_1$ and then, by symmetry, $\angle GTA_1 = \angle A_1T'G$.

Rather remarkably $\angle TGT'$ has GG_1 and GG_1' as a pair of *tri*sectors. Even more, if FIGURE 4 is supplemented with the obvious points U_2, U_3, G_2, G_3, G_2', G_3', then A_iU_i, $i \in \{1,2,3\}$ pass through T at angles of 60° and appear of equal length [8, p. 83]; again, GG_i $(= (1/3)A_iU_i)$ determines an equilateral triangle $G_1G_2G_3$ with center G and circumcircle through T'. This $\triangle G_1G_2G_3$ and its counterpart $\triangle G_1'G_2'G_3'$ are the Napoleon triangles of $\triangle A_1A_2A_3$ [8, p. 63]: GG_i, GG_i' form the three pairs of trisectors of $\angle TGT'$.

3. Locus of midpoints Other suggestive, special cases are provided by points that are far away from $A_1A_2A_3A_0$. In fact, if X "nears infinity", the lines XA_i become practically parallel and it is seen that X' must go far away in the opposite direction. Then, one wonders about the behaviour of the midpoint of XX'! After scrutinizing the various known special cases (e.g., X and X' on A_1A_0, or $X = A_2$, $X' =$ any point of circle $A_1A_0A_3$) a general theorem comes in sight:

THEOREM. *The midpoint of XX' lies on the nine-point circle.*

Proof. Let a_1 (FIGURE 6) be the tangent of circle A_1XX' at A_1: $\angle(a_1A_1X) = \angle A_1X'X$. The intersections T, V and V' of, respectively, a_1, A_1X and A_1X' with circle $A_1A_2A_3$ give reflections in A_2A_3, viz., T_0, V_0 and V_0' on circle $A_2A_0A_3$. Again, on this circle, let N be the point with $T_0N \| XX'$. Then, $NV_0 \| A_1X'$, because $\angle V_0NT_0 = \angle TA_1V = \angle(a_1A_1X) = \angle A_1X'X$. Furthermore, $RX \cdot RU = RA_2 \cdot RA_3 = RA_1 \cdot RV$ leads to $RX/RA_1 = RV_0/RU'$; hence, X is on NV_0. Similarly, $NV_0' \| A_1X$ and X' is on NV_0'. Now, A_1XNX' is a parallelogram and the dilatation with center A_1 and ratio $\frac{1}{2}$ transforms, respectively, T_0N into the line XX', N into the midpoint M of XX', and circle $A_2A_0A_3$ into the nine-point circle.

Remark 1. The tangents a_i of the circles A_iXX' are parallel. For instance,

$$\begin{aligned}\angle(a_1A_1A_0) &= \angle(a_1A_1X) + \angle XA_1A_0 = \angle A_1X'X + \angle A_1XA_0 + \angle XA_0A_1 \\ &= \angle A_1X'X + \angle A_0X'A_1 + \angle XA_0A_1 = \angle A_0X'X + \angle XA_0A_1 \\ &= \angle(a_0A_0X) + \angle XA_0A_1 = \angle(a_0A_0A_1).\end{aligned}$$

The parallels of a_i through the pedal points P_1, P_2, P_3 appear to be isogonally related (see section 4) to P_1S, P_2S, P_3S, with S the midpoint of A_1T_0 (e.g., $P_1S \| A_1'T_0$). At S, as at M, XX' meets the nine-point circle.

Remark 2. A_1 and N form an antigonal pair with respect to $A_2XX'A_3$ (not an orthocentric set). Such pairs are common in this part of geometry. From a higher viewpoint, orthogonal hyperbolas appear to be involved.

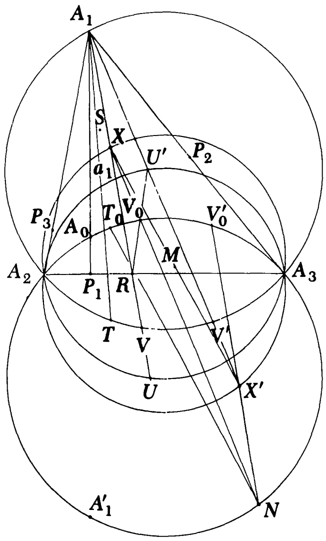


FIGURE 6

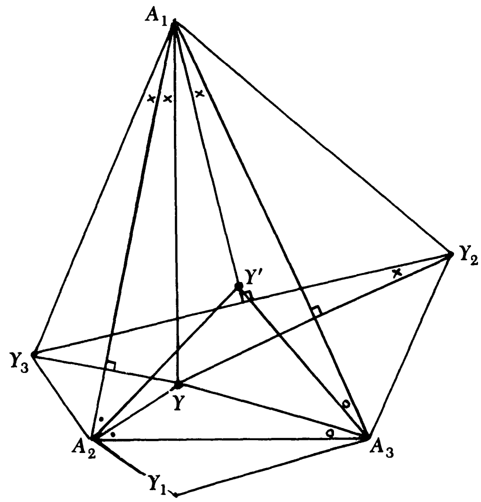


FIGURE 7

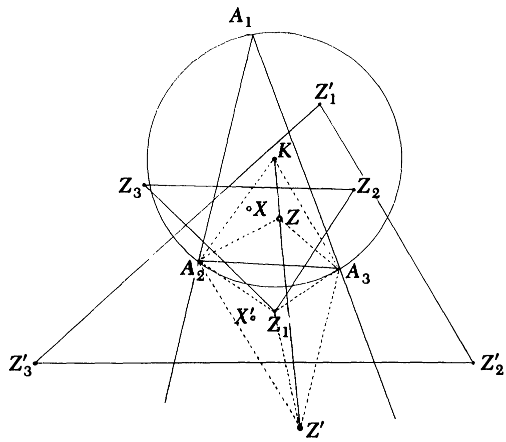


FIGURE 8

4. Characteristic equations When returning to the basic triangle $A_1A_2A_3$ there appear to be other pairs of points that are related to the antigonal ones.

Firstly, there are the *isogonal* pairs Y, Y' (FIGURE 7; [4] and [8]) with

$$\angle A_hA_iY = \angle Y'A_iA_k; \qquad h, i, k \in \{1,2,3\}.$$

Whereas antigonal X, X' give, for instance, $\angle A_2XA_3 + \angle A_2X'A_3 = 0$, it now follows that

$$\begin{aligned}\angle A_2YA_3 + \angle A_2Y'A_3 &= \angle A_2YA_3 + \angle Y'A_2A_3 + \angle A_2A_3Y' \\ &= \angle A_2YA_3 + \angle A_1A_2Y + \angle YA_3A_1 = \angle A_2A_1A_3.\end{aligned}$$

Secondly, there are the *inverse* pairs Z, Z' (FIGURE 8; K: circumcenter) with

$$KZ \cdot KZ' = (KA_i)^2, \qquad i \in \{1,2,3\}.$$

Obviously, similar triangles: $\triangle KA_2Z \sim \triangle KZ'A_2$, $\triangle KA_3Z \sim \triangle KZ'A_3$ give $\angle A_2ZA_3 = \angle A_2ZK + \angle KZA_3 = \angle KA_2Z' + \angle Z'A_3K$; hence $\angle A_2ZA_3 + \angle A_2Z'A_3 = \angle A_2KA_3 = 2\angle A_2A_1A_3$. This corresponds to the last of the following characteristic equations:

antigonal pairs: $\angle A_hXA_k + \angle A_hX'A_k = 0$; $\quad h, k \in \{1,2,3\}$;

isogonal pairs: $\angle A_hYA_k + \angle A_hY'A_k = \angle A_hA_iA_k$; $\quad h, i, k \in \{1,2,3\}$;

inverse pairs: $\angle A_hZA_k + \angle A_hZ'A_k = 2\angle A_hA_iA_k$; $\quad h, i, k \in \{1,2,3\}$.

From these equations it is seen that the three types of pairs are interconnected by the following:

THEOREM. *If X, X' is an antigonal pair and if both X, Z and X', Z' are isogonal pairs, then Z, Z' is an inverse pair.*

As an interesting special case, it is stated without proof that the Torricelli points T, T' are isogonally related to the inverse pair I, I' that are the intersections of the three Appollonian circles (circles $A_iB_iB_i'$ with B_i, B_i' the intersections of A_jA_k and the bisectors of $\angle A_i$; $i, j, k \in \{1,2,3\}$).

The theorem gives a kind of quadruplication in the context of the orthocentric quadrangle $A_1A_2A_3A_0$. The idea is to make a three-step transformation $X \to Z \to Z' \to X'$, built up of, in succession, isogonality, inversion, isogonality. However, this can be done with each of the four triangles $A_1A_2A_3$, $A_2A_3A_0$, $A_3A_0A_1$, and $A_0A_1A_2$, shown in duplicate in FIGURE 9, with their respective circumcenters heavily marked. Now, the left side of FIGURE 9 shows the point X and its four isogonal images Z_0, Z_1, Z_2, Z_3. Then, by comparing the left and the right, the four inversions $Z_i \to Z_i'$ can be seen. At last, on the right side, four isogonal mappings give the same image: $Z_i' \to X'$.

The figure shows a piece of "parallel processing": the four inversions are effected along parallel lines. For a proof, let Z_1^* be the reflection of Z_1 in A_2A_3. Then, $\angle A_3A_2Z_1^* = \angle Z_1A_2A_3 = \angle A_0A_2X = \angle A_0A_2A_3 + \angle A_3A_2X = \angle A_1A_2K + \angle Z_0A_2A_1 = \angle Z_0A_2K$. Similarly, $\angle Z_1^*A_3A_2 = \angle KA_3Z_0$. Hence, Z_1^*, Z_0 is an isogonal pair in $\triangle KA_2A_3$ and, therefore, $\angle Z_0KA_3 = \angle A_2KZ_1^* = \angle Z_1K^*A_2$. Clearly, $KA_3 \| K^*A_2$, hence $Z_0K \| Z_1K^*$ and, similarly, the other parallelisms.

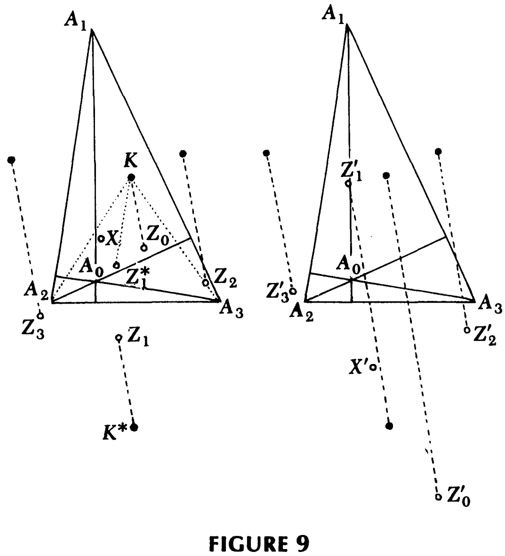


FIGURE 9

5. Reflections in the sides Antigonal, isogonal, and inverse pairs are interconnected in a direct way by reflections in the sides of the given triangle $A_1A_2A_3$. In FIGURE 7 the point Y gives the reflected images Y_1, Y_2 and Y_3. Clearly, Y_2, Y_3, and Y are on a circle around A_1. Hence, $\angle Y_3Y_2Y = \frac{1}{2}\angle Y_3A_1Y = \angle A_2A_1Y$. Now, let Y' be the circumcenter of $\triangle Y_1Y_2Y_3$. As A_1Y' and A_1A_3 are perpendicular bisectors of Y_2Y_3 and YY_2, it follows that $\angle Y'A_1A_3 = \angle Y_3Y_2Y = \angle A_2A_1Y$. That is, A_1Y and A_1Y' are isogonal lines at A_1 and similarly at A_2 and A_3. Hence, Y' is the isogonal image of Y with respect to $\triangle A_1A_2A_3$. Again, reflecting the inverse pair Z, Z' of FIGURE 8 gives the triangles $Z_1Z_2Z_3$ and $Z'_1Z'_2Z'_3$ that will be shown to be oppositely similar. Now, it is seen (note circles around A_3 and A_2) that $\angle Z_2Z_1Z = \angle A_1A_3Z$ and $\angle ZZ_1Z_3 = \angle ZA_2A_1$. Addition leads to $\angle Z_2Z_1Z_3 = \angle A_2ZA_3 + \angle A_3A_1A_2 = \angle A_2ZA_3 - \angle A_2A_1A_3$. Similarly with primes added, hence,

$$\angle Z_2Z_1Z_3 + \angle Z'_2Z'_1Z'_3 = \angle A_2ZA_3 + \angle A_2Z'A_3 - 2\angle A_2A_1A_3 = 0.$$

Thus, $\triangle Z_1Z_2Z_3$ and $\triangle Z'_1Z'_2Z'_3$ are oppositely similar indeed and this property appears to be equivalent to the characteristic equations of inverse pairs.

6. Inversive geometry Isogonal pairs and their characteristic equations lead to another scene: the geometry of circles, the "inversive geometry" [8, p. 103]. A pair Y, Y', isogonal with respect to $\triangle P_1P_2P_3$, satisfies the equation

$$\angle P_2YP_3 + \angle P_2Y'P_3 = \angle P_2P_1P_3.$$

As a consequence, the circles $YP_2P_3, Y'P_2P_3$ and circumcircle $P_1P_2P_3$ give

$$\angle(yP_2P_3) + \angle(y'P_2P_3) = \angle(pP_2P_3), \qquad y, y', p\text{: tangents at } P_2.$$

In words: The circles YP_2P_3 and $Y'P_2P_3$ are isogonal with respect to the circumcircle and P_2P_3. In this statement P_2P_3 is the only line.

Inversive geometry, however, closes the plane with one point $\bar{P}$ "at infinity" and conceives lines as circles through $\bar{P}$. Thus, $\triangle P_1P_2P_3$ and its circumcircle form a set of four circles: $P_1P_2\bar{P}$, $P_2P_3\bar{P}$, $P_3P_1\bar{P}$, and $P_1P_2P_3$, and Y determines six circles: YP_1P_2, YP_2P_3, YP_3P_1, $YP_1\bar{P}$, $YP_2\bar{P}$, and $YP_3\bar{P}$, that combine with six analogs through Y' to pairs that are isogonal with respect to the set of four. This is a reminder of earlier findings and, indeed, FIGURE 6 shows the orthocentric set $A_1A_2A_3A_0$ and the antigonal pair X, X' with circles XA_2A_3 and $X'A_2A_3$ that are *iso*gonal with the circumcircles $A_1A_2A_3$ and $A_0A_2A_3$. Thus, X, X' are an *iso*gonal pair with respect to the four circumcircles of $A_1A_2A_3A_0$. Even more (FIGURE 10), the incenter I and the excenters I_1, I_2, I_3 of $\triangle P_1P_2P_3$ form an orthocentric set and antigonality with respect to $I_1I_2I_3I$ will appear closely related to isogonality with respect to $\triangle P_1P_2P_3$. As a preliminary result it is useful to check that the four circles of $I_1I_2I_3I$ and those of $P_1P_2P_3\bar{P}$ have the same bisector circles. Six of them are lines, e.g., $I_1IP_1\bar{P}$, $I_2I_3P_1\bar{P}$, and six are circles, e.g., $I_1IP_2P_3$, $I_2I_3P_2P_3$.

Remark. The four circles through four non-concyclic points determine 12 bisector circles that again pass through four points, with four circles and the same bisectors. For a proof: Invert a basic point into $\bar{P}$. A simple example involves parallelograms. The circles XI_iI and $X'I_iI$ through antigonal X, X' are mutual reflections in $I_iIP_i\bar{P}$ and the isogonal "circles" $YP_i\bar{P}$ and $Y'P_i\bar{P}$ are too. It appears to be a matter of commuting I's and P's. However, note that antigonality appears at the points X and X' themselves whereas isogonality of Y, Y' appears at the points P_1, P_2, P_3. Additionally, both appear in "directions", i.e., at $\bar{P}$. Inversion, the great tool of inversive geometry, transforms any circle into a circle and any angle between circles into its exact opposite. As for the I's and P's, the inversion with center I and product: $II_1 \cdot IP_1 = II_2 \cdot IP_2 = II_3 \cdot IP_3$ commutes $I_1I_2I_3I \leftrightarrows P_1P_2P_3\bar{P}$. Let it be denoted by (I). The analogous inversion (I_1) commutes $I_1I_2I_3I \leftrightarrows \bar{P}P_3P_2P_1$; similarly $(I_2), (I_3)$. It is (I) that transforms a pair X, X', antigonal with respect to $I_1I_2I_3I$ (recall the six pairs of reflecting circles) into a pair Y, Y', isogonal with respect to $\triangle P_1P_2P_3$ (and reversely). Again, (I_i), $i \in \{1,2,3\}$, transforms X, X' into Y_i, Y_i' giving $YY' \| Y_iY_i'$. Indeed, YY' (i.e. $\bar{P}YY'$) is parallel to the tangent of circle IXX' at I and this tangent is parallel to those of I_iXX' at I_i (see remark 1 in 3.). A direct proof of $YY' \| Y_iY_i'$ will be given in due time, by using the two-fold transformations $(I)(I_i)$.

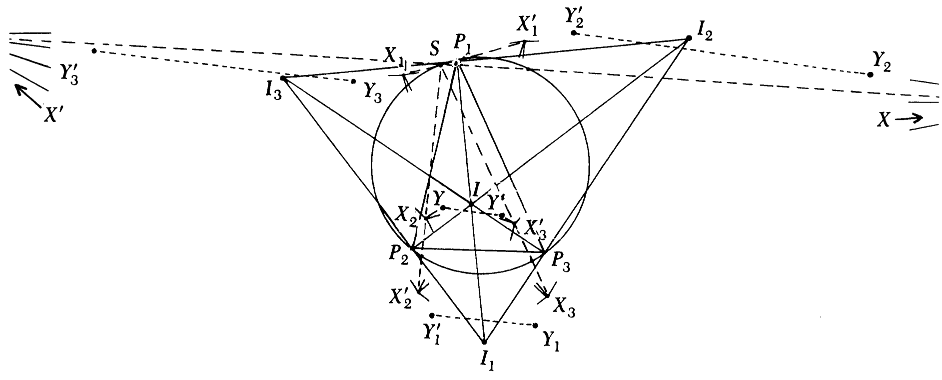


FIGURE 10

These three transformations $(I)(I_i) =: K_i$, $i \in \{1,2,3\}$, just run things. It is easily found that K_1 commutes $I \leftrightarrows I_1$, $I_2 \leftrightarrows I_3$. By symmetry it follows that $K_1 = (I)(I_i) = (I_1)(I) = (I_2)(I_3) = (I_3)(I_2)$ and similarly K_2 and K_3. Clearly, $K_i^{-1} = K_i$, $H_hK_i = (I_h)(I)(I)(I_i) = (I_h)(I_i) = K_j$, $h, i, j \in \{1,2,3\}$, that is, {Identity, K_1, K_2, K_3} is a group, the "four group" of Klein. The effects of the K_i are remarkable. Any point W forms a set of four with $K_iW =: W_i$. If V is one of the latter, say $V = K_1W$, then K_iV are again W, W_3, W_2. Two sets are at hand: $\bar{P}$ with $P_i = K_i\bar{P}$ and I with $I_i = K_iI$. Renumbering the I's, with I an *ex*center and I_i on P_iI, leaves the numbering of K_1, K_2, K_3 unaffected. They are associated to P_1, P_2, P_3.

The sets of four points can form pairs. If Y, Y' are *iso*gonally related with respect to $\triangle P_1P_2P_3$ then Y_i, Y_i' are too. Furthermore, in order to give a direct proof of, for example, $YY' \| Y_1Y_1'$, let (P_1) be the inversion with center P_1 and product $P_1I \cdot P_1I_1$, and let (P_1I) be the reflection in P_1I. Then, $(P_1I)(P_1)$ appears to commute $I \leftrightarrows I_1$, $I_2 \leftrightarrows I_3$, hence it is K_1. Now, $(P_1I)(P_1)$ refers Y to Y_1 on P_1Y' and Y' to Y_1' on P_1Y, also giving $P_1Y \cdot P_1Y_1 = P_1Y' \cdot P_1Y_1'$, i.e., $P_1Y/P_1Y_1' = P_1Y'/P_1Y_1$. Hence, $YY' \| Y_1'Y_1$.

Again, if X, X' are *anti*gonally related with respect to $I_1I_2I_3I$, then X_i, X_i' are too. Furthermore, the working of $(P_1I)(P_1)$ makes $\triangle P_1X_2X_2' \sim \triangle P_1X_3'X_3$. Consequently, the (not lettered) midpoints M_2 and M_3 of X_2X_2' and $X_3'X_3$ give $\angle P_1M_2X_2 = \angle P_1M_3X_3'$. Because P_1, M_2, and M_3 are on the nine-point circle, M_2X_2 and M_3X_3' meet on it. The conclusion is that the four lines XX', X_iX_i' meet on the nine-point circle. (For their meet S, see Remark 1 in 3.) Each of the sets $\{X, X_i\}$, $\{X', X_i'\}$, $\{Y, Y_i\}$, and $\{Y', Y_i'\}$ can be seen as an example of a set of K_i-connected points W, W_i. Any such set forms angles with $I_1I_2I_3I$ that are interrelated in a simple way. It is sufficient to look at relations that are caused by K_1:

$$\angle IW_1I_1 = \angle I_1WI,\ \angle I_1W_1I_2 = \angle IWI_3 \pm \tfrac{1}{2}\pi,\ \angle I_2W_1I_3, = \angle I_3WI_2.$$

The first and the last equation follow from the fact that K_1 reflects any circle through I and I_1 in II_1 and any circle through I_2 and I_3 in I_2I_3. The second equation expresses that K_1 transforms any circle through I and I_3 into a circle through I_1 and I_2, especially the "circle" $II_3\bar{P}$ into the bisector circle perpendicular to I_1I_2. This explains the term $\pm \frac{1}{2}n$.

Finally, K_1, K_2, and K_3 transform two antigonal sets of four lines: $XI\bar{P}$, $XI_i\bar{P}$ and $X'I\bar{P}$, $X'I_i\bar{P}$ into antigonal sets of circles through, respectively, P_1, P_2, and P_3. Thus, antigonality with respect to $I_1I_2I_3I$ does appear not only at $\bar{P}$ (viz. from "directions") but also at the pedal points P_1, P_2, P_3. Even more, by the inversion (I) antigonality enters the case of isogonality:

THEOREM. *If Y, Y' is an isogonal pair with respect to $\Delta P_1P_2P_3$ and I is the incenter, then the line $YI\bar{P}$ and the circles YIP_i are antigonally related to the line $Y'I\bar{P}$ and the circles $Y'IP_i$; and similarly for each excenter.*

A direct proof of this theorem (and the analogs for I_i) is left to the reader.

The author thanks the referee for helpful suggestions.

REFERENCES

1. R. A. Johnson, Directed angles in elementary geometry, *Amer. Math. Monthly* 24 (1917), 101.
2. R. A. Johnson, *Advanced Euclidean Geometry*, Dover Publications, New York, 1960.
3. N. A. Court, *College Geometry*, Barnes & Noble, New York, 1959.
4. H. W. Guggenheimer, *Plane Geometry and Its Groups*, Holden Day, San Francisco, 1967.
5. M. Berger, *Geometry*, Springer-Verlag, Berlin, 1987.
6. H. S. M. Coxeter, *The Real Projective Plane*, 2nd edition, Cambridge University Press, 1961.
7. J. van Yzeren, *Moderne Vlakke Meetkunde*, W. J. Thieme & Cie, Zutphen, 1941.
8. H. S. M. Coxeter and S. L. Greitzer, *Geometry Revisited*, MAA, Washington, DC, 1967.

PROBLEMS

LOREN C. LARSON, *editor*
St. Olaf College

GEORGE GILBERT, *associate editor*
Texas Christian University

Proposals

To be considered for publication, solutions should be received by May 1, 1993.

1408. *Proposed by Syd Bulman-Fleming and Edward T. H. Wang, Wilfrid Laurier University, Waterloo, Ontario, Canada.*

Let S be a commutative, cancellative semigroup and let $P(S)$ denote the semigroup of all nonempty subsets of S under the operation defined by $X \cdot Y = \{xy | x, y \in S\}$ for all $X, Y \in P(S)$. Characterize all those elements of $P(S)$ that are cancellative. That is, find all $A \in P(S)$ with the property that if $X, Y \in P(S)$ and $A \cdot X = A \cdot Y$, then $X = Y$.

1409. *Proposed by Gerald A. Heuer, Concordia College, Moorhead, Minnesota.*

Does there exist a convex pentagon, all of whose vertices are lattice points in the plane, with no lattice point in the interior? (Cf. 1990 Putnam Competition, A3.)

1410. *Proposed by Seung-Jin Bang, Seoul, Republic of Korea.*

Prove that $\lfloor n^{1/3} + (n+1)^{1/3} \rfloor = \lfloor (8n+3)^{1/3} \rfloor$ for every positive integer n.

1411. *Proposed by Miguel Amengual Covas, Cala Figuera (Santanyí), Mallorca, Spain.*

Let C_1 and C_2 be nonconcentric circles in the plane, and consider the set of lines that intersect C_1 and C_2 in equal chords. Show that these lines are all tangent to a single parabola.

1412. *Proposed by Sam Northshield, State University of New York, College at Plattsburgh, Plattsburgh, New York.*

Let g: $\mathbf{N} \to \mathbf{R}$ be a bounded function and f: $\mathbf{N} \to \mathbf{N}$ a function such that $\lim_{n \to \infty} f(n) = \infty$. Show that if $c > 1$ and $\lim_{n \to \infty}(cg(n) - g(f(n)))$ exists, then $\lim_{n \to \infty} g(n)$ exists.

ASSISTANT EDITORS: CLIFTON CORZAT, BRUCE HANSON, RICHARD KLEBER, KAY SMITH, *and* THEODORE VESSEY, *St. Olaf College and* MARK KRUSEMEYER, *Carleton College. We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals should be accompanied by solutions, if at all possible, and by any other information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. An asterisk (*) next to a problem number indicates that neither the proposer nor the editors supplied a solution.*

Solutions should be written in a style appropriate for Mathematics Magazine. Each solution should begin on a separate sheet containing the solver's name and full address.

Solutions and new proposals should be mailed in duplicate to Loren Larson, Department of Mathematics, St. Olaf College, 1520 St. Olaf Ave., Northfield, MN 55057-1098 or mailed electronically via fax: (507) 663-3549 or e-mail: larson@stolaf.edu.

Quickies

Answers to the Quickies are on page 354.

Q797. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada, and dedicated to the late Joseph Konhauser, Macalester College, St. Paul, Minnesota.*

Determine the maximum and minimum values of

$$S = \sin^2 x + \sin^2 y + 2k \sin x \sin y$$

where $x, y \geqslant 0$, $x + y = \alpha$, and α, k are given constants with $0 \leqslant \alpha \leqslant \pi$.

Q798. *Proposed by Howard Morris, Chatsworth, California.*

Can the integers modulo 28 be partitioned into two sets, $\{x_i\}_{i=1}^{14}$ and $\{y_i\}_{i=1}^{14}$ so that $y_i - x_i \equiv i \pmod{28}$?

Q799. *Proposed by Bjorn Poonen, University of California, Berkeley, California.*

Let S_1 and S_2 be nondegenerate n-dimensional simplices (not necessarily regular) in $\mathbf{R}^n$. Let a_1 be the number of vertices of S_1 in the interior of S_2 and let a_2 be the number of vertices of S_2 in the interior of S_1. For fixed n, find the maximum of $a_1 + a_2$.

Solutions

Probabilities of partial sums **December 1991**

1383. *Proposed by Michael Handelsman, Erasmus Hall High School, Brooklyn, New York.*

Integers from 1 to n are randomly generated (as by a computer). Each integer has an equal probability of being selected, and unlimited repetition is permitted. As integers are generated, a running sum is recorded.

Given any integer k, such that $1 \leqslant k \leqslant n$, what is the probability that a sum of *exactly* k will be reached?

I. Solution by Philip D. Straffin, Beloit College, Beloit, Wisconsin.

Let $P(n, k)$ be the probability that a run using the integer n will produce a sum of exactly k. We will show that $P(n, k) = (1/n)(1 + 1/n)^{k-1}$, for $1 \leqslant k \leqslant n$.

First note that $P(n, 1) = 1/n$. Next, for $2 \leqslant k \leqslant n$ there are two kinds of runs that produce a sum of k. The first is a run that starts with 1, followed by a run that produces a sum of $k - 1$. The probability of this is $(1/n)P(n, k-1)$. The second is a run that starts with an integer $1 < m \leqslant k$. Every run of this type has the same probability as a run reaching $k - 1$, which is identical to it except that it starts with $m - 1$, so the total probability of a run of this type is exactly $P(n, k-1)$. Thus

$$P(n,k) = \left(1 + \frac{1}{n}\right)P(n, k-1), \qquad 2 \leqslant k \leqslant n,$$

and the result follows.

II. Solution by Reiner Martin (student), University of California at Los Angeles, Los Angeles, California.

It is well known that the equation $a_1 + \cdots + a_i = k$ has $\binom{k-1}{i-1}$ solutions with positive integers $a_1, \ldots, a_i$. Thus, the probability in question is

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

Also solved by *Robert A. Agnew*, *Gary Amende* and *Tommy Tze (students)*, *Michael H. Andreoli*, *Gene Arnold*, *Pierre Barnouin (France)*, *Donald Batman*, *Arthur Benjamin*, *Robert E. Bernstein*, *Ed Blaise (student)* and *Karl A. Beres*, *Stephen Bronn*, *Centre Problem Solving Group*, *Michael W. Chamberlain*, *Xi Chen (student)*, *Paul Chiou*, *Con Amore Problem Group (Denmark)*, *Robert L. Doucette*, *Drew University Math 101*, *Milton P. Eisner*, *Thomas E. Elsner*, *Jiro Fukuta (Japan)*, *Herbert Gintis*, *Jerrold W. Grossman*, *Ram Prakash Gupta (Virgin Islands; two solutions)*, *David Hankin*, *Russell Jay Hendel*, *Robert High*, *Mack Hill*, *Chuck Hixon*, *Dave Horn*, *R. Daniel Hurwitz*, *Paul Irwin*, *Bruce R. Johnson (Canada)*, *Kiran S. Kedlaya (student)*, *Dovie Kimmins* and *Nell Rayburn*, *Emil F. Knapp*, *John Koker*, *Ed Korntved*, *H. K. Krishnapriyan*, *John W. Krussel*, *Dennis Kunimura*, *Norman F. Lindquist*, *Peter W. Lindstrom*, *Andy Martin*, *Helen M. Marston*, *Ray McClanahan*, *Larry Olson*, *E. D. Onstott* and *N. N. Gurwell*, *John Oman*, *Andrej Panjkov (Australia)*, *Nicholas Passell*, *P. J. Pedler (Australia)*, *Edsel A. Peña*, *Miguel A. Diaz Quiñones (Spain)*, *F. C. Rembis*, *H. D. Ruderman*, *Richard F. Ryan*, *Tony Scalise* and *Vance Cochrane*, *Harvey Schmidt, Jr.*, *Rob Schneiderman (student)*, *Harry Sedinger*, *Heinz-Jürgen Seiffert (Germany)*, *John S. Sumner*, *Patrick Touhey*, *Gilbert Traub*, *University of North Texas Problem Solving Group*, *University of Wyoming Problem Circle*, *Michael Vowe (Switzerland)*, *Robert J. Wagner*, *Jack V. Wales, Jr.*, *William P. Wardlaw*, *Richard Weida*, *Mark R. Woodard*, *Chengmin Yang*, and the *proposer*.

Agnew, *Traub* and *Wales* gave expressions for arbitrary k, and observed, along with *Lindstrom*, that for $k > n$, $P(k) = (1/n)\sum_{j=k-n}^{k-1} P(j)$; that is, $P(k)$ is the average of the previous n probabilities. One consequence of this is that $\max_{1 \leqslant k < \infty} P(k) = P(n)$; that is to say, n is the integer most likely to appear in the running sum. *David Callan*, University of Wisconsin, points out that this problem is rather well known, see Problem 1217, this MAGAZINE, Vol. 59, No. 3, June 1986, pp. 174–175, and the comments following the solution.

Periodic matrices **December 1991**

1384. *Proposed by William P. Wardlaw, U.S. Naval Academy, Annapolis, Maryland.*

A square matrix $\mathbf{M}$ is *periodic* if and only if there is a positive integer r and a nonnegative integer s such that $\mathbf{M}^{r+s} = \mathbf{M}^s$. Characterize those fields K such that every square matrix over K is periodic.

Solution by Reiner Martin (student), University of California at Los Angeles, Los Angeles, California.

Every square matrix over K is periodic if and only if the characteristic of K is positive and K is algebraic over its prime subfield k.

If the characteristic of K is zero, the matrix $2\mathbf{I}$ is not periodic, where $\mathbf{I}$ denotes the identity matrix. If K is not algebraic over k, the matrix $\alpha\mathbf{I}$ is not periodic, where $\alpha \in K$ is transcendental over k.

Now let $\mathbf{A} = (a_{ij})$ be a square matrix over K, where K has positive characteristic and is algebraic over k. Then k is finite, and the field $L \subseteq K$ obtained by adjoining the elements a_{ij} to k is finite. It follows that $\mathbf{A}$ must be periodic, because there are only a finite number of possibilities for distinct powers of $\mathbf{A}$.

Also solved by *D. K. Cohoon*, *David Callan*, *Xi Chen (student)*, *Con Amore Problem Group (Denmark)*, *F. J. Flanigan*, *Furman University Problem Solving Group*, *Stephen I. Gendler and Dipendra Bhattacharya*, *Joel K. Haack*, *David Koster*, *Harvey Schmidt, Jr.*, and the *proposer*. A partial solution was submitted by *Colonel Johnson, Jr.* and the *Jaguar-10 Student Chapter of MAA, Southern University, Baton Rouge, Louisiana*.

A family of irrational numbers **December 1991**

1385. *Proposed by Howard Morris, Chatsworth, California.*

Let $f(x)=\prod_{n=1}^{\infty}(1+x/2^n)$. Show that at the point $x=1$, $f(x)$ and all its derivatives are irrational.

Solution by the Trinity University Problem Group, Trinity University, San Antonio, Texas.

We first express $f(x)$ as a power series $\sum_{n=0}^{\infty}A_n x^n$. It is immediate that $A_0=1$. Also,

$$A_1=\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^j=1=\frac{1}{2-1},$$

and

$$A_2=\sum_{j=1}^{\infty}\sum_{k>j}\frac{1}{2^j}\frac{1}{2^k}=\sum_{j=1}^{\infty}\frac{1}{2^j}\sum_{k=1}^{\infty}\frac{1}{2^{j+k}}=\sum_{j=1}^{\infty}\frac{1}{2^j}\left(\frac{1}{2^j}\sum_{k=1}^{\infty}\frac{1}{2^k}\right)=\sum_{j=1}^{\infty}\left(\frac{1}{2^2}\right)^j A_1.$$

In general,

$$\begin{aligned}A_{n+1}&=\sum_{j=1}^{\infty}\sum_{k_1>j}\sum_{k_2>k_1}\cdots\sum_{k_n>k_{n-1}}\frac{1}{2^j}\frac{1}{2^{k_1}}\cdots\frac{1}{2^{k_n}}\\&=\sum_{j=1}^{\infty}\left(\frac{1}{2^j}\right)^{n+1}\sum_{k_1=1}^{\infty}\sum_{k_2>k_1}\cdots\sum_{k_n>k_{n-1}}\frac{1}{2^{k_1}}\frac{1}{2^{k_2}}\cdots\frac{1}{2^{k_n}}\\&=\sum_{j=1}^{\infty}\left(\frac{1}{2^{n+1}}\right)^j A_n=\frac{1}{2^{n+1}-1}A_n\\&=\frac{1}{2^{n+1}-1}\,\frac{1}{2^n-1}\cdots\frac{1}{2^1-1}.\end{aligned}$$

In what follows, $f(x)$ will be denoted by $f^{(0)}(x)$ and, for $j\geqslant 1$, $f^{(j)}(x)$ will denote the j-th derivative of $f(x)$. Term-by-term differentiation gives

$$f^{(j)}(x)=\sum_{n=j}^{\infty}A_n(n)_j x^{n-j},$$

where $(n)_n=n(n-1)\cdots(n-j+1)$, so

$$f^{(j)}(1)=\sum_{n=j}^{\infty}A_n(n)_j,\qquad j\geqslant 0.$$

Now fix j and assume that $f^{(j)}(1)$ is a rational number p/q, p and q integers. Then,

$$0<\frac{p}{q}-\sum_{n=j}^{k}A_n(n)_j=\sum_{n=k+1}^{\infty}A_n(n)_j \qquad (*)$$

where $\sum_{n=j}^{k}A_n(n)_j$ is a rational number that can be written in the form a/b, with $b=(2-1)(2^2-1)\cdots(2^k-1)$. Multiplying expression $(*)$ by bq gives

$$0 < bp - aq = q \sum_{n=k+1}^{\infty} (2-1)(2^2-1)\cdots(2^k-1)A_n(n)_j$$

$$= q \sum_{n=k+1}^{\infty} \frac{1}{(2^{k+1}-1)\cdots(2^n-1)}(n)_j$$

with $bp - aq$ a positive integer. The right side is less than $q\Sigma(\mathrm{n})_{\mathrm{j}}/(2^{\mathrm{n}}-1)$, which converges, so by taking k large enough we can make the right sum as small as we wish. We have thus produced a positive integer lying between 0 and 1. This is untenable and therefore $f^{(j)}(1)$ must be irrational.

Also solved by *Paul Bracken (Canada), Erhard Braune (Austria), Xi Chen (student), Con Amore Problem Group (Denmark), Kee-Wai Lau (Hong Kong), John S. Sumner* and *Kevin L. Dove*, and the *proposer*.

An application of Casey's power theorem **December 1991**

1386. *Proposed by Jiro Fukuta, Motosu-gun, Gifu-ken, Japan.*

Let ABC be an acute-angled triangle, let H be the foot of the altitude from A, and let D, E, Q be the feet of the perpendiculars from an arbitrary point P in the triangle onto AB, AC, AH, respectively. Prove that

$$|AB \cdot AD - AC \cdot AE| = BC \cdot PQ,$$

where $AB, AD, \ldots$ denote the length of segments $AB, AD, \ldots$.

I. *Solution by the University of Wyoming Problem Circle, University of Wyoming, Laramie, Wyoming.*

Consider the triangle as situated in the Euclidean plane so that points are vectors, and let (U, V) denote the inner product of the vectors U, V.

Since the point D is the projection of P onto the segment AB, it follows that $D - A$ is the orthogonal projection of $P - A$ onto $B - A$. Thus,

$$D - A = \frac{(P-A, B-A)}{(B-A, B-A)}(B-A).$$

Moreover, since D is between A and B, the coefficient is positive. Similarly.

$$E - A = \frac{(P-A, C-A)}{(C-A, C-A)}(C-A),$$

and the coefficient is positive. Therefore,

$$AB \cdot AD = (P-A, B-A) \qquad \text{and} \qquad AC \cdot AE = (P-A, C-A).$$

Consequently,

$$AB \cdot AD - AC \cdot AE = (P-A, B-C) = (P-Q, B-C) + (Q-A, B-C).$$

Since both segments BC and PQ are perpendicular to AQ, it follows that the vectors $B - C$ and $P - Q$ are parallel and that $(Q - A, B - C) = 0$. Thus

$$|AB \cdot AD - AC \cdot AE| = |(P-Q, B-C)| = PQ \cdot BC.$$

II. *Solution by Richard Pfiefer, San Jose State University, San Jose, California.*

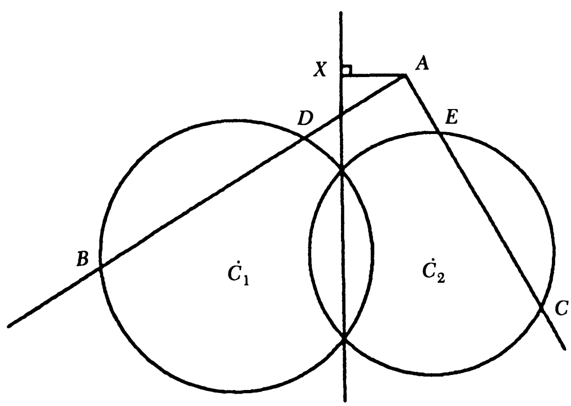
The result is an application of Casey's Power Theorem (see R. A. Johnson, *Advanced Euclidean Geometry*, Dover Publications, Inc., New York, 1960, pp. 86–87): "The difference of the powers of a point with regard to two nonconcentric circles is twice the product of the distance between their centers by the distance from the point to the radical axis of the circles." That is (see figure),

$$AD \cdot AB - AE \cdot AC = 2C_1C_2 \cdot AX.$$

Now let C_1 and C_2 be the midpoints of PB and PC. Then P is on the radical axis, PQ is the distance from A to the radical axis (i.e. $PQ = AX$) and $C_1C_2 = \frac{1}{2}BC$. Applying Casey's Power Theorem yields

$$AD \cdot AB - AE \cdot AC = BC \cdot PQ.$$

Note: Casey's Power Theorem is valid if signed (directed) distances are used, so the notation above obviates the use of absolute values. Also, the triangle ABC need not be acute-angled, and P can be any point in the plane (even a vertex).



$AD \cdot AB - AE \cdot AC = 2C_1C_2 \cdot AX.$

Also solved by *Karen Anewalt, Seung-Jin Bang (Korea), Francisco Bellot* and *María Ascensión López (Spain), Paul Bracken (Canada), Xi Chen (student), Con Amore Problem Group (Denmark), Miguel Amengual Covas (Spain), H. Guggenheimer, Thomas E. Gantner, David Hankin, Francis M. Henderson, John G. Heuver (Canada), R. Daniel Hurwitz, Geoffrey A. Kandall, Kiran S. Kedlaya (student), Václav Konečný, Neela Lakshmanan, Ty Nguyen Le, Helen M. Marston, Ralph Merrill, Roger B. Nelsen, Billy D. Read, Harry D. Ruderman, Volkhard Schindler (Germany), John S. Sumner, Michael Vowe (Switzerland), David Zhu (student), Weiwei Zhu (student),* and the *proposer.*

An application of the Weierstrass approximation theorem **December 1991**

1387. *Proposed by Kenneth Stolarsky, University of Illinois at Urbana-Champaign, Urbana, Illinois.*

Given $\varepsilon > 0$ and a function $f(x)$, continuous on $(-\infty, \infty)$, must there exist a function $g(x)$, continuous on $[0, 1]$, such that

$$\inf_{-\infty < y < \infty} \left(\max_{0 \leqslant x \leqslant 1} |g(x) - f(x+y)| \right) \geqslant \varepsilon?$$

Solution by Adam Riese, Wright State University, Dayton, Ohio.

Let $S = \{f_1, f_2, \dots, f_n, \dots\}$ be a countable set of continuous functions on $[0,1]$ and let S be dense in $C[0,1]$, the set of all continuous functions on $[0,1]$. (By the Weierstrass Approximation Theorem, S can be chosen to be the set of all polynomials with rational coefficients.) Let

$$f(x) = \begin{cases} 0, & \text{if } x \leqslant 1, \\ f_n(x-2n), & \text{if } 2n \leqslant x \leqslant 2n+1 \text{ for some } n>0, \\ (2n-x)f_{n-1}(1) + (x+1-2n)f_n(0), & \text{if } 2n-1 < x < 2n \text{ for some } n>0, \end{cases}$$

where $f_0(1) = 0$, which is a continuous function.

Then for any continuous function $g(x)$ and each $\varepsilon > 0$ there exists n such that $|f_n(x) - g(x)| < \varepsilon$ for all $x \in [0,1]$. Hence $\max_{0 \leqslant x \leqslant 1} |g(x) - f(x+2n)| < \varepsilon$ and thus $\inf_{-\infty < y < \infty} (\max_{0 \leqslant x \leqslant 1} |g(x) - f(x+y)|) = 0$.

Also solved by *Xi Chen (student), Con Amore Problem Group (Denmark), Yuanan Diao, Thomas E. Gantner, Reiner Martin (student), Northern Kentucky University Problem Group, John Sumner* and *Kevin L. Dove, Chengmin Yang*, and the *proposer*. There was one incorrect solution.

Answers

Solutions to the Quickies on page 349.

A797. We rewrite S in the form

$$S = \sin^2 x + \sin^2 y + 2\cos\alpha \sin x \sin y + 2(k - \cos\alpha)\sin x \sin y$$

and it then follows easily that

$$\begin{aligned} S &= \sin^2\alpha + 2(k-\cos\alpha)\sin x \sin y \\ &= \sin^2\alpha + (k - \cos\alpha)\big(\cos(x-y) - \cos(x+y)\big). \end{aligned}$$

Case 1. $k \geqslant \cos\alpha$.

$$\begin{aligned} \max S &= \sin^2\alpha + (k-\cos\alpha)(1-\cos\alpha) = (1+k)(1-\cos\alpha), \\ \min S &= \sin^2\alpha. \end{aligned}$$

Case 2. $k < \cos\alpha$.

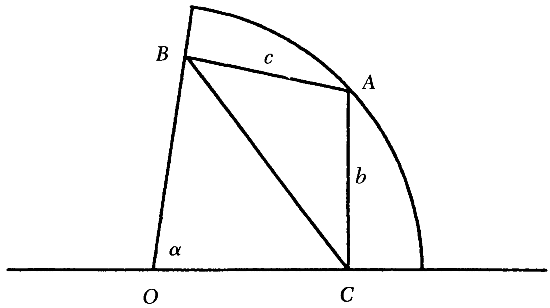
$$\begin{aligned} \min S &= (1+k)(1-\cos\alpha), \\ \max S &= \sin^2\alpha. \end{aligned}$$

This problem was suggested by the following Quickie problem by the late Joe Konhauser (private communication).

From a point A on the circular arc (radius R) of a sector of a nonobtuse angle α, drop perpendiculars to the sides. If the two feet are B and C, determine the extreme values of BC. Since $OBAC$ is cyclic, the circumradius of ABC is the same as that of OBA, which is $OA/2$ since $\angle OBA$ is a right angle. Then since the product of the sides of triangle $BAC = 4 \times$ its circumradius $\times$ its area,

$$bc \cdot BC = 4(R/2)(bc\sin(\pi - \alpha))/2,$$

or $BC = R\sin\alpha =$ constant.



A798. When i is even, x_i and y_i have the same parity. When i is odd they have opposite parity. Since there are 7 odd i between 1 and 14, for a partition to exist there must be an odd number of odd integers modulo 28, a contradiction.

A799. The maximum is $n+1$. The convex hull C of the set V of all $2n+2$ vertices is a polyhedron with vertices in V. Since C is nondegenerate, C has at least $n+1$ vertices, none of which lie in the interior of S_1 or S_2. Thus $a_1 + a_2 \leqslant (2n+2) - (n+1) = n+1$. Equality can be attained when S_1 lies inside of S_2, for instance.

Comments

Readers may be interested in knowing that timely solutions to the Putnam problems are available by way of the USENET newsgroup sci.math. The solutions found therein are due to *Dan Bernstein*, New York University, and are posted within two days of the examination. Readers of the newsgroup contribute to a lively exchange of comments and alternative solutions. The solution to Problem A-4 on the 1991 exam, that appeared in the April issue of this MAGAZINE, can be traced back to Professor Bernstein. He plans to continue this service for future examinations.

1365. *Ernest B. Leach Mayfield Heights, Ohio*, notes that the term *identric mean* was coined in a series of papers [2], [3], and [4]. These papers investigate the various relationships between a family of means of the form

$$E(r,s;x,y) = \left(\frac{(y^s - x^s)/s}{(y^r - x^r)/r} \right)^{1/(s-r)}$$

where $x \neq y$, $x > 0$, $y > 0$, $r \neq s$. The function is extended to be continuous (and even analytic) for all $(r,s;x,y)$ satisfying $x > 0$, $y > 0$. The *identric mean* is defined as $E(r,r;x,y)$, the name being motivated by the *identity* relation $r = s$. Its importance in the theory lies in the fact that $\ln E(r,s;x,y)$ is the integral average of values $\ln E(\sigma,\sigma;x,y)$, for σ in the interval $[r,s]$. Among many other things, it is shown that $E(r,s;x,y)$ increases with an increase of r or s. The solution to **1365** follows, since the problem is equivalent to proving that $E(1/2,1;x,y) < E(1,1;x,y)$. The case $E(1,1;x,y)$ was studied much earlier by *Cisbani* in [1].

ADDITIONAL REFERENCES

1. R. Cisbani, "Contributi alla teoria delle medie I," *Metron*, 13 (2) (1938), 23–34.
2. E. B. Leach and M. C. Sholander, "Extended mean values," *Amer. Math. Monthly*, 81 (1974), 879–883.

3. E. B. Leach and M. C. Sholander, "Extended mean values II," *J. Math. Anal. Appl.* 92 (1983), 207–223.
4. E. B. Leach and M. C. Sholander, "Multivariable extended mean values," *J. Math. Anal. Appl.* 104 (1984), 390–407.

Using the notation established in Solution III of **1365**, *Heinz-Jürgen Seiffert, Germany*, has shown (unpublished) that

$$G\exp\left(\frac{A-L}{L}\right)=I.$$

This result implies many inequalities, such as, $GA \le LI$. He also has the following improvement: Define $M = M(a,b) = (b-a)/(4\arctan(\sqrt{b/a})-\pi)$. Then $L \le M \le I$, with equality if and only if $a=b$.

Q787. *Kiran S. Kedlaya, (student), Georgetown Day High School*, gives the following solution to **Q787**. Suppose $\tan nx$ is defined for all n and $\lim_{n\to\infty}\tan nx = L$. Then $\lim_{n\to\infty}\tan(n+1)x = L$ as well. Taking limits on each side of

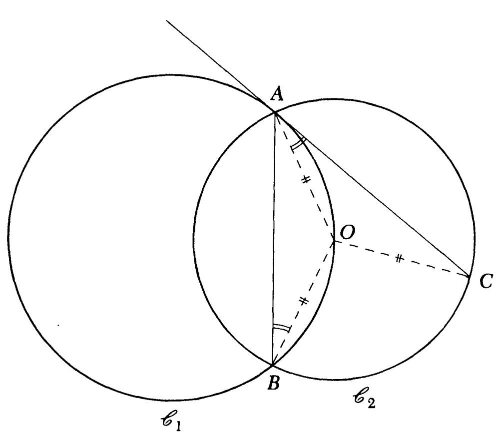
$$\tan x = \frac{\tan(n+1)x - \tan nx}{1+\tan(n+1)x\tan nx}$$

gives $\tan x = (L-L)/(1+L^2) = 0$, so that $x = k\pi$, which is a necessary (and clearly sufficient) condition.

An even more succinct solution is given by *Bruce Reznick, University of Illinois, Urbana*. Let $\{t\}$ denote the fractional part of t. Then $\{n\alpha\}$ is an arithmetic progression (mod 1) and so converges if and only if $\alpha \in \mathbf{Z}$. But $\tan nx = \tan \pi\{(nx)/\pi\}$ and $\pi\{(nx)/\pi\} \in [0,\pi)$ on which tan is one–to–one. Thus, if $\tan nx$ converges then $\{(nx)/\pi\}$ must be 0 for all n.

Proof without Words

If circle $\mathscr{C}_1$ passes through the centre O of circle $\mathscr{C}_2$, the length of the common chord $\overline{AB}$ is equal to the tangent segment $\overline{AC}$.



—R. H. Eddy
Memorial University of Newfoundland
St. John's, Newfoundland
Canada A1C 5S7

3. E. B. Leach and M. C. Sholander, "Extended mean values II," *J. Math. Anal. Appl.* 92 (1983), 207–223.
4. E. B. Leach and M. C. Sholander, "Multivariable extended mean values," *J. Math. Anal. Appl.* 104 (1984), 390–407.

Using the notation established in Solution III of **1365**, *Heinz-Jürgen Seiffert, Germany*, has shown (unpublished) that

$$G\exp\left(\frac{A-L}{L}\right)=I.$$

This result implies many inequalities, such as, $GA \leq LI$. He also has the following improvement: Define $M = M(a,b) = (b-a)/(4\arctan(\sqrt{b/a}) - \pi)$. Then $L \leq M \leq I$, with equality if and only if $a=b$.

Q787. *Kiran S. Kedlaya, (student), Georgetown Day High School*, gives the following solution to **Q787**. Suppose $\tan nx$ is defined for all n and $\lim_{n\to\infty}\tan nx = L$. Then $\lim_{n\to\infty}\tan(n+1)x = L$ as well. Taking limits on each side of

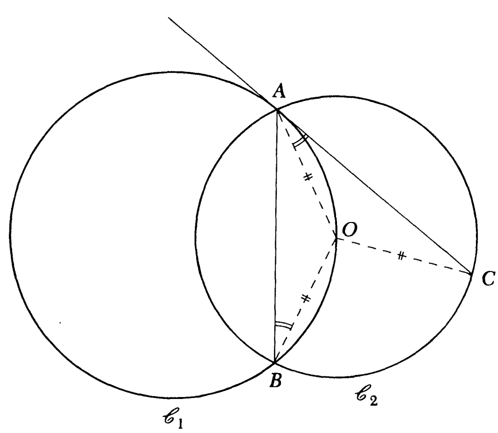
$$\tan x = \frac{\tan(n+1)x - \tan nx}{1+\tan(n+1)x\,\tan nx}$$

gives $\tan x = (L-L)/(1+L^2) = 0$, so that $x = k\pi$, which is a necessary (and clearly sufficient) condition.

An even more succinct solution is given by *Bruce Reznick, University of Illinois, Urbana*. Let $\{t\}$ denote the fractional part of t. Then $\{n\alpha\}$ is an arithmetic progression (mod 1) and so converges if and only if $\alpha \in \mathbf{Z}$. But $\tan nx = \tan \pi\{(nx)/\pi\}$ and $\pi\{(nx)/\pi\} \in [0,\pi)$ on which tan is one–to–one. Thus, if $\tan nx$ converges then $\{(nx)/\pi\}$ must be 0 for all n.

Proof without Words

If circle $\mathscr{C}_1$ passes through the centre O of circle $\mathscr{C}_2$, the length of the common chord $\overline{AB}$ is equal to the tangent segment $\overline{AC}$.



—R. H. Eddy
Memorial University of Newfoundland
St. John's, Newfoundland
Canada A1C 5S7

REVIEWS

PAUL J. CAMPBELL, *editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of the mathematics literature. Readers are invited to suggest items for review to the editors.

Peterson, I., Primality tests: An infinity of exceptions, *Science News* 142 (19 September 1992) 182. Devlin, Keith, There are infinitely many Carmichael numbers, *Focus* 12(4)(September 1992) 1–2. Granville, Andrew, Primality testing and Carmichael numbers, *Notices of the American Mathematical Society* 39(7)(September 1992) 696–700.

Fermat's Little Theorem says that if p is prime, then for any a, p divides $a^p - a$, more customarily written $a^{p-1} \equiv 1 \pmod{p}$. The converse is false, as R.D. Carmichael (1878?–1967) discovered in 1910 (try 561); composites that satisfy the division criterion for all a are called *Carmichael numbers*. They are rare (43 of them less than 1,000,000), but they had been completely characterized by Korselt in 1899: n is Carmichael iff n is squarefree and $p-1$ divides $n-1$ for all primes p that divide n. (Since Korselt did not exhibit any, they are not called *Korselt numbers*.) Carmichael listed 15 of them and footnoted that "This list might be indefinitely extended." His use of "might" seems to imply uncertainty. Other mathematicians were uncertain, too, until now. Thanks to a proof by Red Alford, Andrew Granville, and Carl Pomerance (University of Georgia), we know that *for sufficiently large N, there are more than $N^{2/7}$ Carmichael numbers.* These three expository articles are listed in order of increasing detail and sophistication. Curiously, none gives Carmichael's first name or initials, nor citations to his papers (On composite numbers P which satisfy the Fermat congruence $a^{P-1} \equiv 1 \pmod{P}$, *Amer. Math. Monthly* 19 (1912) 22–27, which expands on an announcement in *Bull. Amer. Math. Soc.* 16 (1910) 237–238).

Bombieri, Enrico, Prime territory: Exploring the infinite landscape at the base of the number system, *The Sciences* (September/October 1992) 30–36.

"What is a prime number, and what makes it so special? ... [W]hat properties of primes are worthy of serious attention ... ? ... [H]unting for the largest perfect number today yields only trophies for the curio shop. ... Is the nonexistence of odd perfect numbers an interesting problem? ... [T]he interest for problems of this kind is 'practically nil.' In my opinion what would be of real interest would be to show that certain problems of that kind are truly undecidable ... ." Bombieri, one of the world's leading number theorists, shows beautifully how the distribution of primes can be analyzed by means of harmonic analysis of waveforms, and goes on to state his opinion that the Riemann Hypothesis—whose truth he accepts on the grounds of parsimony and esthetics—is "the most important unsolved problem in mathematics today." He concludes by alluding to the uses of primes in cryptography. Was he is thrilled at the news of an infinity of Carmichael numbers? We will have to hear from him again.

Apostol, Tom, et al., *A Century of Calculus*, 2 vols., MAA, 1992; xv + 462 pp and xvi + 481 pp, \$36(P) each. ISBN 0-88385-205-5, 0-88385-206-3

The first of these volumes was originally published as *Selected Papers on Calculus* (1969), the selections being from the *Monthly* and this MAGAZINE. The second volume contains reprintings from the years 1969–1991 from the same publications plus the *College Math. J.*

Rabinowitz, Stanley (ed.), *Index to Mathematical Problems 1980–1984*, MathPro Press (P.O. Box 713, Westford, MA 01886), 1992; xii + 532 pp, \$49.95. ISBN 0–9626401–1–5

This innovative and much-needed volume contains all of the more than 5,000 problems (and none of the solutions) from 28 mathematical journals and 16 contests from 1980 through 1984, all originally in English. The problems are grouped by subject matter and accompanied by indexes by journal and problem number, by topic, by keyword, and by author. A well-known problemist and problem editor, Rabinowitz will publish in a subsequent volume English translations of problems from other journals; he hopes to publish additional indexes, working backward and forward five years at a time. The task is monumental, and the mathematical community will be greatly in Rabinowitz's debt.

Gordon, Florence, and Sheldon Gordon (eds.), *Statistics for the Twenty-First Century*, MAA, 1992; xii + 318 pp, \$22(P). ISBN 0–88385–078–8

A more apt title would be *Statistics* Education *for the Twenty-First Century*, as another MAA Notes volume (*Perspectives on Contemporary Statistics*, edited by David C. Hoaglin and David S. Moore, 1992) treats statistics itself. The book at hand concentrates on pedagogical aspects for the "first" statistics course ("the *last* course in statistics for the overwhelming majority of students"). Particularly notable is the article by Ruma Falk and and Clifford Konold, which brands as a grave mistake the now-conventional wisdom that statistics educators should teach as little probability as necessary for learning statistics. This volume also includes bibliographies of materials and case studies and a guide to electronic newsgroups.

Peterson, Ivars, Shuttle liquids: Taking coffee-cup mathematics into orbit, *Science News* 142 (22 August 1991) 124–125. Concus, Paul, and Robert Finn, Math results on capillary surfaces to be tested in space, *SIAM Review* (March 1992) 8–9.

The June 25 flight of the space shuttle Columbia carried an experiment devised by two mathematicians to test the geometry of liquid surfaces in containers of unusual shape. The experiment has practical importance for understanding how liquids behave in space. The theoretical results of Concus (Lawrence Berkeley Laboratory) and Finn (Stanford University) predict that symmetric configurations may be unstable and have higher energy than unexpected asymmetric shapes. How closely did the experiment agree with theory? Concus and Finn won't know until they have analyzed in detail the videotapes of the experiment. Finn, who started out as an engineer before becoming a mathematician, remarked: "I thought I had escaped all that physical reality that I didn't want to have anything to do with, but now it's more fun. I think mathematics should get motivational input from outside—from the rest of science. Otherwise mathematics dies out."

Peterson, Ivars, and Carol Ezzell, Crazy rhythms: Confronting the complexity of chaos in biological systems, *Science News* 142 (5 September 1992) 156–159. Garfinkel, Alan, Mark L. Spano, William L. Ditto, and James N. Weiss, Controlling cardiac chaos, *Science* 257 (28 August 1992) 1230–1235.

Researchers have been able to stabilize cardiac arrhythmias in rabbits by administering electrical stimuli according to a "chaos control" strategy. Their work lends hope that chaos theory may lead to development of "smart" cardiac pacemakers for humans. But are cardiac fibrillations a symptom of a deterministic system naturally exhibiting chaos, or are they simply random? A chaotic system gives the appearance of randomness because it is continually switching from an unstable motion of one period to another unstable motion with a different period. So some of what has been identified by physical scientists as chaotic behavior may in fact be just randomness.

Kerr, Richard A., From Mercury to Pluto, chaos pervades the solar system, *Science* 257 (3 July 1992) 33.

A major achievement of Newtonian physics was explaining and predicting the orbits of the planets. Now we know, from calculations for the next 100 million years, that every one of the planets in the solar system has a chaotic orbit. Resonances between planets "pump up" planetary chaos; but still, the planets appear constrained—over the 100 million years, "they do not fly out of their orbits." The constraints may be secondary resonances.

Maddox, John, Why pebbles float to the surface, *Nature* 358 (13 August 1992) 535.

When you shake a pail containing a mix of sand and pebbles, the pebbles tend to rise to the top. Now there is a simple explanation, based on a geometrical model and simulations by French researchers. After each shaking, the pebbles tend to have small "voids" underneath them, which are filled in (and then some) on the next shaking by sand grains; so the pebbles can move only upward. The process seems to work for a variety of materials, provided that the "pebble" diameter is at least (about) 2.8 times the diameter of the "sand grain."

Olkin, Ingram, Reconcilable differences: Gleaning insight from conflicting scientific studies, *The Sciences* (July/August 1992) 30–36.

Citing numerous examples of its success, statistician Olkin makes the case for *meta-analysis*, the quantitative combining of information from conflicting independent studies of the same phenomenon. Olkin explains the use of *effect size* as an important measure in meta-analysis, cautions against injudicious use, and urges a national registry of experimental results (published or not).

Simon, Barry, Symbolic math software: It's not just for mainframes anymore, *PC Magazine* 11(14)(August 1992) 405–432. Simon, Barry, Comparative CAS reviews, *Notices of the American Mathematical Society* 39(7)(September 1992) 700–710.

Simon, a mathematician at Caltech, offers in *PC Magazine* a comparative review of computer algebra systems (CAS) for PCs (386 or better): Derive, Maple V, Mathematica (both DOS and Windows versions), and Reduce, with a sidebar on MathCAD. The review cites general results (including wrong answers!) from using the packages on a test suite of 20 problems. The specific problems and their solutions are in the second article.

World checkers champ wins against computer, *Wisconsin State Journal* (30 August 1992) 7A. Checkers crown eludes computer, *Science News* 142 (3 October 1992) 217.

Marion Tinsley, semiretired mathematician, has won the World Draughts Championships for the 19th straight year, this time against the computer program Chinook, devised by computer scientist Jonathan Schaeffer (University of Alberta). The score was 4 wins for Tinsley (2 on computer glitches), 2 for Chinook, and 33 draws, over 13 days. Chalk up losses #6 and #7 for Tinsley over the last 42 years. (For more about Chinook and Tinsley, see: The checkers challenge: A checker-playing computer program contends for the world title, by Ivars Peterson, *Science News* 140 (3) (20 July 1991) 40–41.)

Van Gelder, Lawrence, Cosmic meaning, eyelashes and zero, *New York Times* (1 August 1992) The Arts.

This piece is a review of a show that you probably won't ever get to see: "My Mathematics," written and performed by Rose English, with music by Ian Hill. Part of the Serious Fun festival at Lincoln Center in New York, this mostly one-woman show also featured a stallion named "Mathematics." Despite "references to zeros" and invoking "the relationship between nought and zero and nothing," reviewer Van Gelder felt that the performance "seems to be one of those enterprises that speak with more coherence to its creator than to its audience." Hmm ... perhaps an appropriate and perceptive title, after all.

NEWS AND LETTERS

ACKNOWLEDGEMENTS

In addition to our associate editors, the following have assisted the MAGAZINE *by refereeing papers during the past year. We appreciate the time and care they have given.*

Adams, Williams, *University of Maryland, College Park*
Alberts, George E., *Fort George Meade, MD*
Alexander, Jay, *University of Maryland, College Park*
Alexanderson, Gerald L., *Santa Clara University, CA*
Aminzadeh, Mostafa, *Towson State University, MD*
Anand, Kailash, *Concordia University, Montreal, Canada*
Askey, Richard, *University of Wisconsin, Madison*
Banchoff, Thomas, *Brown University, Providence, RI*
Barksdale, James B. Jr., *Western Kentucky University, Bowling Green*
Beals, Richard W., *Yale University, New Haven, CT*
Beineke, Lowell, *Indiana University-Purdue University, Fort Wayne*
Benedetto, John, *University of Maryland, College Park*
Bloch, Norman J., *SUNY College at Brockport*
Boas, Mary L., *Seattle, WA*
Boas, Ralph P., *Seattle, WA*
Booth, Ada, *Santa Clara University, CA*
Botsko, Michael, *St. Vincent College, LaTrobe, PA*
Boules, Raouf, *Towson State University, MD*
Bradberry, Brent, *Lewis-Clark College, Lewiston, ID*
Brenner, Joel L., *Palo Alto, CA*
Broman, Arne, *Chalmers University of Technology, Göteborg, Sweden*
Brualdi, Richard A., *University of Wisconsin, Madison*
Bushaw, Donald W., *Washington State University, Pullman*
Castaldi, Ray, *Towson State University, MD*
Chakerian, G.D., *University of California, Davis*
Chollet, John, *Towson State University, MD*
Comfort, Wistar, *Wesleyan University, Middletown, CT*
Cunningham, Stephen, *California State University, Stanislaus*
Danby, Anthony, *North Carolina State University, Raleigh, NC*
Dauben, Joseph W., *Herbert H. Lehman College, West Bronx, NY*
Delany, James, *California Polytechnical State University, San Obispo*
DeTemple, Duane W., *Washington State University, Pullman*
Devries, David, *Georgia College, Milledgeville*
Donkar, Eli, *Social Security Administration, Baltimore, MD*
Drobot, Vladimir, *San Jose State University, CA*
Drucker, Daniel, *Wayne State University, Detroit, MI*
Drucker, Thomas, *Dickinson College, Carlisle, PA*
Eckert, Ernest J., *University of South Carolina, Aiken*
Eddy, R.H., *Memorial University of Newfoundland, St. John's, Newfoundland, Canada*
Eisenberg, Bennett, *Lehigh University, Bethlehem, PA*
Elliot, P.D.T.A., *University of Colorado, Boulder*
Farris, Frank, *Santa Clara University, CA*
Feroe, John, *Vassar College, Poughkeepsie, NY*
Fisher, J. Chris, *University of Regina, Saskatchewan, Canada*
Gallian, Joseph A., *University of Minnesota-Duluth*
Gardiner, Anthony, *University of Birmingham, U.K.*
Ghahramani, Saeed, *Towson State University, MD*
Gillman, Leonard, *University of Texas, Austin*
Goldstein, Jerome A., *Louisianna State University, Baton Rouge*
Goldstein, Larry, *Brookeville, MD*
Goodson, Geoffrey, *Towson State University, MD*
Gorini, Catherine, *Maharishi International University, Fairfield, IA*
Grant, John, *Towson State University, MD*
Grossman, Steve, *Baltimore, MD*
Grünbaum, Branko, *University of Washington, Seattle*
Gulick, Denny, *University of Maryland, College Park*
Guy, Richard, *University of Calgary, Alberta, Canada*
Halmos, Paul R., *Santa Clara University, CA*
Harrell, Ronald E., *Allegheny College, Meadville, PA*
Hatcher, Rhonda, *Texas Christian University, Fort Worth*
Henderson, James, *Colorado College, Colorado Springs*
Hilton, Peter, *SUNY at Binghamton*
Hindman, Neil, *Howard University, Washington, DC*
Hochwald, Scott, *University of North Florida, Jacksonville*
Honsberger, Ross, *University of Waterloo, Ontario, Canada*
Horak, Martin, *Towson State University, MD*
Horn, Roger A., *The Johns Hopkins University, Baltimore, MD*
Hutchinson, Joan, *Smith College, Northampton, MA*
Ilgenfritz, Ernie, *Towson State University, MD*
Itzkowitz, Gerald, *Queens College, Flushing, NY*
Jamison, Robert E., *Clemson University, SC*
Jech, Thomas, *Pennsylvania State University, University Park*
Jepsen, Charles H., *Grinnell College, IA*
Johnston, Elgin, *Iowa State University, Ames*
Jones, Sharon, *Towson State University, MD*

Jur, Barbara, *Macomb County Community College, Warren, MI*
Karr, Alan F., *The Johns Hopkins University, Baltimore, MD*
Katz, Irving, *George Washington University, Washington, DC*
Kelly, Leroy M., *Michigan State University, East Lansing, MI*
Kimberling, Clark H., *University of Evansville, IN*
Klamkin, Murray, *University of Alberta, Edmonton, Canada*
Krieger, Henry, *Harvey Mudd College, Claremont, CA*
Krusemeyer, Mark, *Carleton College, Northfield, MN*
Kung, Sidney H., *Jacksonville University, FL*
Kurepa, Alexandra, *Texas Christian University, Fort Worth*
LaCampagne, Carole, *Washington, DC*
Lagarias, Jeffrey, *AT&T Bell Labs, Murray Hill, NJ*
Lawrence, Louise, *Towson State University, MD*
Lay, David C., *University of Maryland, College Park*
Leinbach, Carl, *Gettysburg College, PA*
Liu, Andrew, *University of Alberta, Edmonton, Canada*
Mader, Adolf, *University of Hawaii, Honolulu*
Markham, Elizabeth, *St. Joseph's College, West Harford, CT*
Maurer, Stephen B., *Swarthmore College, PA*
McCleary, John, *Vassar College, Poughkeepsie, NY*
McEachin, Raymond, *Indiana University-Purdue University, Fort Wayne*
Merrill, Samuel III, *Wilkes College, Wilkes-Barre, PA*
Meskin, Stephen, *Columbia, MD*
Michaels, John G., *SUNY at Brockport*
Moore, Thomas L., *Grinnell College, IA*
Moritz, Roger, *Alfred University, NY*
Morrison, John, *Towson State University, MD*
Moser, William O.J., *McGill University, Montreal, Quebec, Canada*
Nelsen, Roger B., *Lewis & Clark College, Portland, OR*
Nievergelt, Yves, *Eastern Washington University, Cheney*
Owens, Robert, *Lewis & Clark College, Portland, OR*
Page, Warren, *New York City Technical College, Brooklyn*
Park, Efton, *Texas Christian University, Fort Worth*
Peeples, Joanne V., *El Paso Community College, TX*
Penn, Howard Lewis, *U.S. Naval Academy, Annapolis, MD*
Pfiefer, Richard, *San Jose State University, CA*
Porter, Gerald, *University of Pennsylvania, Philadelphia*
Ralston, Anthony, *SUNY at Buffalo*
Richman, Dan, *Baltimore, MD*
Rigby, John F., *University of Wales, College of Cardiff, U.K.*
Ross, Peter, *Santa Clara University, CA*
Rota, Gian-Carlo, *Cambridge, MA*
Rothman, Neal J., *Indiana University-Purdue University, Indianapolis*
Sagan, Hans, *North Carolina State University, Raleigh*
Sandefur, James T., *Georgetown University, Washington, DC*
Sangwine-Yager, Jane, *St. Mary's College, Davis, CA*
Schaumberger, Norman, *Hofstra University, Hempstead, NY*
Schecter, Murray, *Lehigh University, Bethlehem, PA*
Scheinerman, Ed, *The Johns Hopkins University, Baltimore, MD*
Segal, Sanford, *University of Rochester, NY*
Senechal, Marjorie, *Smith College, Northampton, MA*
Seydel, Kenneth, *Skyline College, San Bruno, CA*
Shallit, Jeffrey, *University of Waterloo, Ontario, Canada*
Simon, Julie, *Antioch College, Yellow Springs, OH*
Small, Donald, *U.S. Military Academy, West Point, NY*
Smart, Rod, *University of Wisconsin, Madison*
Smith, Bryan, *University of Puget Sound, Tacoma, WA*
Smith, David A., *Duke University, Durham, NC*
Snell, J. Laurie, *Dartmouth College, Hanover, NH*
Sohrab, Houshang, *Towson State University, MD*
Spanier, Jerome, *Claremont Graduate School, CA*
Stein, Sherman K., *University of California, Davis*
Stern, Leonid, *Towson State University, MD*
Stone, Arthur H., *Northeastern University, Boston, MA*
Stone, Dorothy Maharam, *Northeastern University, Boston, MA*
Stromquist, Walter, *Daniel H. Wagner Assoc., Berwyn, PA*
Taylor, Alan, *Union College, Schenectady, NY*
Teles, Elizabeth J., *Montgomery College, Tacoma Park, MD*
Tingley, Daryl, *University of New Brunswick, Fredericton, Canada*
Tucker, Alan, *SUNY at Stony Brook*
Tucker, Thomas W., *Colgate University, Hamilton, NY*
Wagon, Stan, *Macalester College, St. Paul, MN*
Wakin, Shirley, *Southern Conecticut State University, New Haven*
Weiss, Asia Ivic, *York University, Ontario, Canada*
Wilf, Herbert, *University of Pennsylvania, Philadelphia*
Zimmerman, Jay, *Towson State University, MD*
Zorn, Paul, *St. Olaf College, Northfield, MN*

Statement of Ownership, Management and Circulation
(Required by 39 U.S.C. 3685)

1A. Title of Publication: MATHEMATICS MAGAZINE

1B. PUBLICATION NO.: 0 0 2 5 5 7 0 X

2. Date of Filing: October 2, 1992

3. Frequency of Issue: bi-monthly except July and August

3A. No. of Issues Published Annually: Five

3B. Annual Subscription Price: Library $72.00; Indiv. Member $17.00

4. Complete Mailing Address of Known Office of Publication *(Street, City, County, State and ZIP+4 Code) (Not printers)*
1529 Eighteenth St., N.W., Washington, D.C. 20036-1385

5. Complete Mailing Address of the Headquarters of General Business Offices of the Publisher *(Not printer)*
1529 Eighteenth St., N.W., Washington, D.C. 20036-1385

6. Full Names and Complete Mailing Address of Publisher, Editor, and Managing Editor *(This item MUST NOT be blank)*

Publisher *(Name and Complete Mailing Address)*
The Mathematical Association of America, 1529 Eighteenth St., N.W., Washington, D.C. 20036-1385

Editor *(Name and Complete Mailing Address)*
Martha J. Siegel, Towson State Univ., Dept. of Mathematics, Baltimore, MD 21204

Managing Editor *(Name and Complete Mailing Address)*
Harry Waldman, MAA, 1529 Eighteenth St.,N.W., Washington, D.C. 20036-1385

7. Owner *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given. If the publication is published by a nonprofit organization, its name and address must be stated.) (Item must be completed.)*

Full Name	Complete Mailing Address
The Mathematical Association of America, Inc.	1529 Eighteenth St., N.W. Washington, D.C. 20036-1385

8. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages or Other Securities *(If there are none, so state)*

Full Name	Complete Mailing Address
None	

9. For Completion by Nonprofit Organizations Authorized To Mail at Special Rates *(DMM Section 424.12 only)*
The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes *(Check one)*

(1) [X] Has Not Changed During Preceding 12 Months
(2) [ ] Has Changed During Preceding 12 Months
(If changed, publisher must submit explanation of change with this statement.)

10. Extent and Nature of Circulation *(See instructions on reverse side)*	Average No. Copies Each Issue During Preceding 12 Months	Actual No. Copies of Single Issue Published Nearest to Filing Date
A. Total No. Copies *(Net Press Run)*	20,697	18,292
B. Paid and/or Requested Circulation 1. Sales through dealers and carriers, street vendors and counter sales	0	0
2. Mail Subscription *(Paid and/or requested)*	17,030	15,379
C. Total Paid and/or Requested Circulation *(Sum of 10B1 and 10B2)*	17,030	15,379
D. Free Distribution by Mail, Carrier or Other Means Samples, Complimentary, and Other Free Copies	1,027	653
E. Total Distribution *(Sum of C and D)*	18,057	16,032
F. Copies Not Distributed 1. Office use, left over, unaccounted, spoiled after printing	2,640	2,260
2. Return from News Agents	0	0
G. TOTAL *(Sum of E, F1 and 2—should equal net press run shown in A)*	20,697	18,292

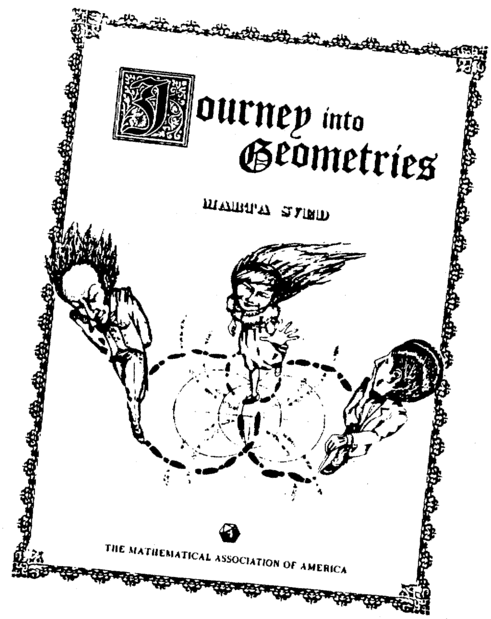
11. I certify that the statements made by me above are correct and complete

Signature and Title of Editor, Publisher, Business Manager, or Owner
Kay Lamont
Membership Manager

PS Form 3526, January 1991 *(See instructions on reverse)*

JOURNEY INTO GEOMETRIES

Marta Sved



This charming book introduces us to topics in hyperbolic geometry in a delightfully informal style. Early in the 19th century, Janos Bolyai created "non-Euclidean" geometry, discovered independently by two other mathematicians of Bolyai's day, Gauss, and Lobachevsky. At the time these concepts were too revolutionary to make a serious impact. However, later developments in relativity theory and twentieth century perceptions made hyperbolic geometry an integral part of geometry, logically as perfect as classical geometry, yet still strangely surprising.

JOURNEY INTO GEOMETRIES can be read at two levels. It can be studied as an informal introduction to post-Euclidean geometry, brought to life in dialogues between three fictitious figures: a somewhat grown up Alice, Lewis Carroll and their visitor from the Twentieth century, Dr. Whatif. It also can serve as background material for university students, for the material presented in the text is extended by carefully selected problems. The background required is minimal, standard high school geometry, yet the serious student, aided by problems attached to each chapter, should acquire a deeper understanding of the subject.

192 pp., Paperbound, 1991
ISBN 0-88385-500-3

List: $21.00 MAA Member: $14.00

Catalog Number JOG

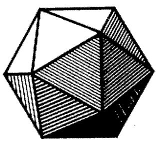
ORDER FROM:

Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC. 20036
(FAX) (202) 265-2384

Prepaid orders sent postage & handling free. Visa and Mastercard orders accepted. (Please give the card number and expiration date on credit card orders) We will bill for orders over $10.00.

PROBLEMS FOR MATHEMATICIANS: Young and Old

Paul R. Halmos



This is a book of problems for mathematicians at all levels. Halmos says: "I wrote this book for fun. It was fun indeed—the book almost wrote itself. It consists of some of the many problems that I started saving and treasuring a long time ago. Problems came up in conversations with friends, and in correspondence, and in books and in lectures. I enjoyed them, thought about them, tried to solve them, tried to change them, and tried to think of new ones, and then I tried to organize and write down the ones I was fondest of—and this book is the result."

The problems come complete with their statements, hints, and solutions. The purpose of the statements is to stimulate thought. The reader is asked to think of extensions and improvements of the results asked for. The hints are intended to get the reader to look in a possibly profitable direction. The solutions may sometimes be "wrong," or "partially wrong," and then corrected. The solutions make no pretense of being the best, the shortest, the most elegant or even complete, but their purpose is to have the reader solve the problem, and to enjoy doing so.

Some of the problems can be solved by high school students. Others require the maturity of a professional mathematician, who can be a second year graduate student or someone who has been earning a living by thinking about mathematics for a long time. All of them are challenging and fun.

1991, Paperbound,
ISBN 0-88385-321-3

List: $24.00 MAA Member: $16.00

Catalog Number DOL-12

ORDER FROM:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036
(202) 387-5200 Fax (202) 265-2384

Name________________

Address________________

City________ State____ Zip____

Qty.	Title	Price

Payment ❑ Check ❑ VISA/MASTERCARD

Credit Card No.__________ Total $____

Signature__________ Exp. Date____

CONTENTS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, D.C. 20036