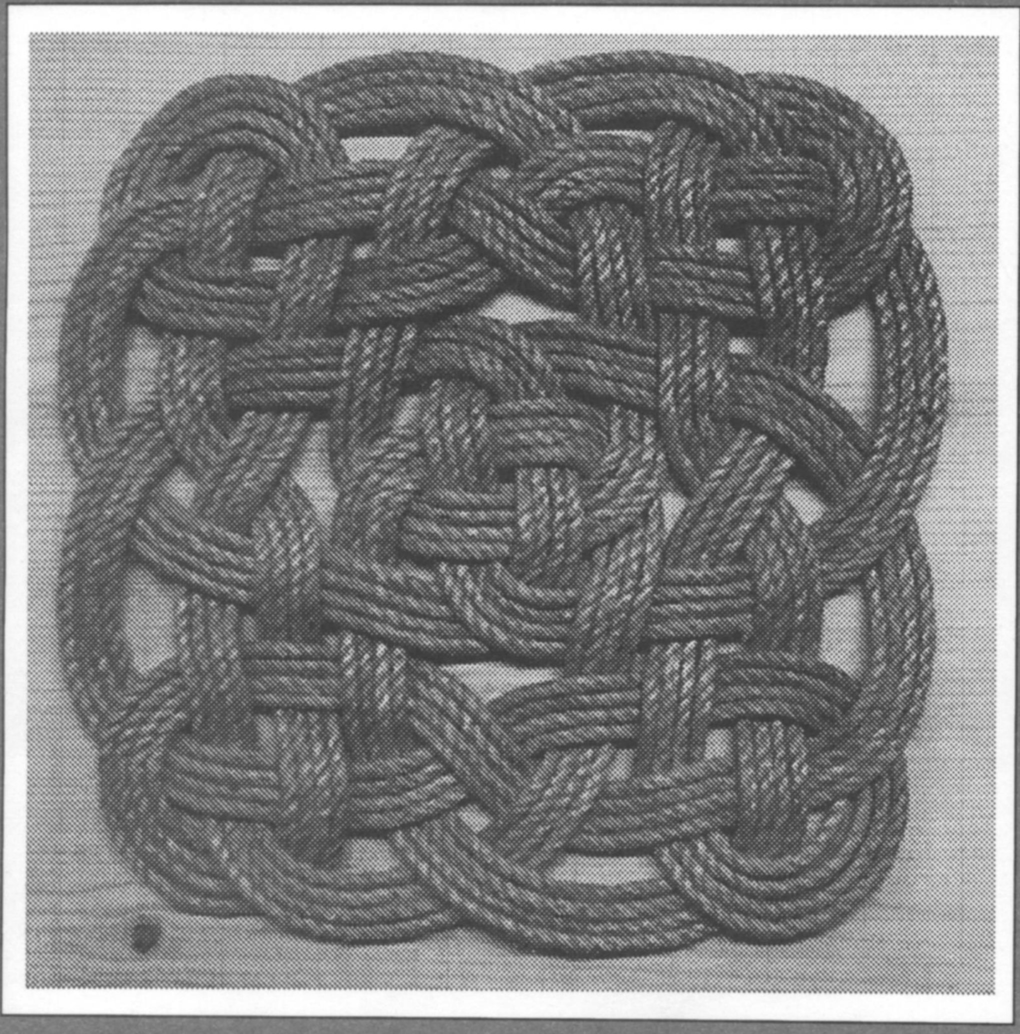
Vol. 72, No. 1, February 1999

MATHEMATICS MAGAZINE



Wormloop (see pp. 32–38)

- Dynamics of Circle Homeomorphisms
- Chaos and One-to-Oneness
- When and Why Do Water Levels Oscillate in Three Tanks?

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

Mathematics Magazine aims 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. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 71, pp. 76–78, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Send new manuscripts to Paul Zorn, Editor, Department of Mathematics, St. Olaf College, 1520 St. Olaf Avenue, Northfield, MN 55057-1098. Manuscripts should be laser-printed, with wide line-spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit three 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. Copies of figures should be supplied on separate sheets, both with and without lettering added.

Cover illustration: Wormloop. The cover shows a rope mat designed by Nils Kr. Rossing in 1995. The mathematical expression for the mat was found by a two-dimensional Fourier analysis, and can be mathematically described by 9 Fourier components, or 9 rotating vectors. Some vectors rotate clockwise and some counterclockwise.

The distance between successive spectral components is 4, which is reflected in the main structure of the mat: 4 small loops and 4 large loops. The offset of the dominant component is 3. This is reflected in the number of turns the string makes around the center of the mat when sweeping through the total trajectory.

The technical name of the rosette is Alternating Eye Rosette. Most of the mats within the eye rosette family have eyes turned inward. This special mat has eyes turned alternately inward and outward. For this reason, the mat gives a Celtic impression, which justifies the name: Wormloop.

AUTHORS

Keith Kendig received his Ph.D. from UCLA in algebraic geometry, under Basil Gordon, and subsequently spent two years at the Institute for Advanced Study in Princeton. There, Hassler Whitney sparked in him a lifelong love of examples and counterexamples. One subset of examples and metaphors comprises "tanklish," which he teaches in his differential equations classes. Currently a professor in the mathematics department at Cleveland State University, he wrote the text *Elementary Algebraic Geometry,* published by Springer-Verlag. Keith is a semi-professional cellist who, in his spare time, plays chamber music and dabbles in composing classical music.

Rodney Nillsen studied mathematics, science, and literature at the University of Tasmania, graduating in 1967. He obtained a Ph.D. in 1971 from The Flinders University of South Australia. After a period at the Royal University of Malta and the University College of Swansea, Wales, he joined the mathematics department at the University of Wollongong in 1974. His primary research interests are in harmonic analysis. Since giving a course on chaos, he has become more involved in this area, which he has found to be of great interest to undergraduate students.

Jim Walsh studied two years in Togo, West Africa as a U.S. Peace Corps volunteer after receiving his bachelor's degree from the University of Connecticut. He has taught in a public high school, earned a master's degree in education, and, in 1991, received his doctorate in mathematics from Boston University. Currently an associate professor at Oberlin College, his research interests are in low-dimensional dynamical systems, with occasional forays into fields as disparate as queueing theory and numerical weather prediction. He is convinced that many topics from dynamical systems can be successfully incorporated into the undergraduate curriculum; he offers his article as a case in point.

Vol. 72, No. 1, February 1999

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.

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

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

Printed in the United States of America

ARTICLES

The Dynamics of Circle Homeomorphisms: A Hands-on Introduction

JAMES A. WALSH
Oberlin College
Oberlin, OH 44074-1019

Introduction

The dynamics of circle homeomorphisms is a deep, beautiful, and surprisingly accessible topic for students in an advanced calculus or introductory real analysis course. The remainder of this article should be considered a proof of this claim. It is structured as a sequence of connected exercises providing the reader with an introduction to the theory. These exercises could be interspersed throughout a semester course in junior-level real analysis, or used collectively as a capstone experience (solutions to the exercises are on the web **[25]** or available by writing the author). We begin with historical background.

Henri Poincaré **[22]** introduced the study of the dynamics of circle homeomorphisms in his attempt to classify solutions to ordinary differential equations (or *flows*) defined on the two-dimensional torus $\mathbb{T}^2$. We think of $\mathbb{T}^2$ as being obtained by identifying points (x_1, y_1) and (x_2, y_2) in $\mathbb{R}^2$ if $(x_1, y_1) = (x_2, y_2) + (m, n)$ for some integer pair (m, n). Flows defined on $\mathbb{T}^2$ thus correspond to vector fields $V: \mathbb{R}^2 \to \mathbb{R}^2$, $V(x, y) = (V_1(x, y), V_2(x, y))$, which are 1-periodic in each coordinate as in FIGURE 1 ([**5**, Ch. 17], [**11**, §6.1], [**16**, §1.5, §14.2]). Poincaré was interested in the role the topology of $\mathbb{T}^2$ plays in determining the long-term behavior of solutions.

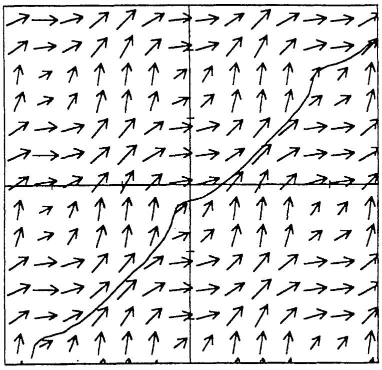


FIGURE 1
The vector field $V(x, y) = (1.1 + \sin(2\pi y), 1.1 - \cos(2\pi x))$ is invariant under integer translations.

Beginning with the simplest case, Poincaré considered differential equations having no equilibrium points (points (x_0, y_0) for which $V(x_0, y_0) = (0, 0)$). These flows exist on $\mathbb{T}^2$ because the Euler characteristic is zero [**10**, §3.5]. Within this class of differential equations, he restricted his study to vector fields with $V_1(x, y) > 0$ and $V_2(x, y) > 0$ for all $(x, y) \in \mathbb{R}^2$—all vectors point "northeast" in such vector fields (see FIGURE 1).

Now consider a meridianal circle $\mathbb{T}$ on the torus and let $\theta \in \mathbb{T}$ (see FIGURE 2). Note that, due to the class of vector fields under consideration, the trajectory beginning at θ must wrap around the torus and intersect $\mathbb{T}$ on the "other side." Poincaré thus

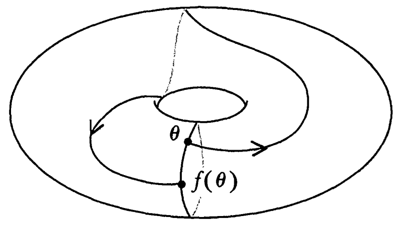


FIGURE 2
Defining the return map $f: \mathbb{T} \to \mathbb{T}$.

defined a map $f: \mathbb{T} \to \mathbb{T}$ which assigns to $\theta \in \mathbb{T}$ the first point of intersection of the trajectory of θ with $\mathbb{T}$. Assuming the vector field V has continuously differentiable components, the existence and uniqueness theorem for ODE's (see, e.g., [**21**, §2.2]) implies the map f is an orientation-preserving (circle) homeomorphism. The qualitative study of ODE's defined on $\mathbb{T}^2$ is thus reduced to the study of iteration of the return map f.

The map f is an example of what is now called a *Poincaré return map*. More generally, such maps allow for the qualitative study of flows in $\mathbb{R}^n$ by considering the Poincaré return map on an appropriately chosen hyperplane in $\mathbb{R}^n$ ([**9**, §1.5], [**11**, §6.1], [**16**, §0.3], [**21**, §3.4], [**24**, §5.8]). What now seems such a natural technique is one of Poincaré's many deep and significant contributions to the field of dynamical systems. For an insightful discussion of Poincaré's work in dynamical systems ("creation of" is perhaps more apt) see the introduction in [**23**].

The study of the dynamics of circle homeomorphisms provides a host of wonderful exercises at the advanced undergraduate level. In the following sections we present an introduction to the subject, but rather than include proofs of statements we provide exercises (with hints where appropriate) which the reader is encouraged to complete. So, with pencil and paper in hand, please read on!

Preliminaries

Much of the following can be found in [**12**]. Let $\mathbb{T}$ denote the real numbers mod 1, i.e., $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We think of $\mathbb{T}$ as a circle of circumference one due to the identification of 0 mod 1 and 1 mod 1 in $\mathbb{R}/\mathbb{Z}$. The value $1/4 \in \mathbb{T}$, for example, represents one-fourth of a "turn" around $\mathbb{T}$. Let $f: \mathbb{T} \to \mathbb{T}$ be an orientation-preserving circle homeomorphism, so that each of f and f^{-1} is continuous, and f preserves the (cyclic) order of points on $\mathbb{T}$. The simplest type of circle homeomorphism is a *rigid rotation* $r_\omega(\theta) = \theta + \omega$, where ω is a fixed real number mod 1.

Given $\theta \in \mathbb{T}$, define $f^0(\theta) = \theta$ and, for integers $i > 0$, $f^i(\theta) = f(f^{i-1}(\theta))$. For $i < 0$, let $f^i(\theta) = f^{-1}(f^{i+1}(\theta))$. We seek to understand the behavior of *orbits* of f

$$o(\theta, f) = o(\theta) = \{f^i(\theta) : i \in \mathbb{Z}\}.$$

The simplest type of orbit is a *periodic orbit*, one for which $f^n(\theta) = \theta$ for some $n \in \mathbb{Z}$. Such a θ-value is called a *periodic point* of period n. Periodic orbits for f correspond to periodic trajectories on $\mathbb{T}^2$ for the class of differential equations considered by Poincaré in **[22]**.

Exercise 1. Consider rigid rotation $r_\omega : \mathbb{T} \to \mathbb{T}$.

(a) Let $\omega = p/q \in \mathbb{Q}$ (throughout, we assume $\gcd(p, q) = 1$). Show that all orbits of r_ω are periodic with the same period.

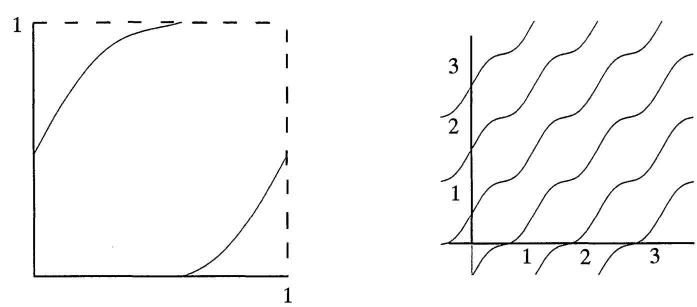
(b) Suppose $\omega \notin \mathbb{Q}$. Show that all orbits of r_ω are dense in $\mathbb{T}$. Hint: For $\theta \in \mathbb{T}$, show that $o(\theta)$ is an infinite set. Use the compactness of $\mathbb{T}$ or the Bolzano-Weierstrass Theorem to conclude that $o(\theta)$ has a limit point in $\mathbb{T}$. Given $\epsilon > 0$, deduce the existence of integers m and n with $r_\omega^n(\theta)$ and $r_\omega^m(\theta)$ less than ϵ apart. Now use the fact that r_ω is rigid rotation.

For $x \in \mathbb{R}$, let $\langle \cdot \rangle$ denote the fractional part of x. Let $\pi : \mathbb{R} \to \mathbb{T}$, $x \mapsto \langle x \rangle$, be the projection map from the reals onto $\mathbb{T}$. Note that π wraps any interval of the form $[x, x+1)$ once around $\mathbb{T}$ since it identifies x and $x+1$. A *lift* of f is a function $F : \mathbb{R} \to \mathbb{R}$ such that $\pi F(x) = f\pi(x)$ for all $x \in \mathbb{R}$, i.e., such that the diagram

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\ F\ } & \mathbb{R} \\ \pi \downarrow & & \downarrow \pi \\ \mathbb{T} & \xrightarrow{\ f\ } & \mathbb{T} \end{array}$$

commutes. Given that $f : \mathbb{T} \to \mathbb{T}$ is continuous, one can choose a continuous lift $F : \mathbb{R} \to \mathbb{R}$ (**[5**, §17.1], [**17**, §5.6]). We will assume that all lifts in this article are continuous.

The graph of a lift can be drawn as follows. Consider the graph of f as a subset of $\mathbb{T} \times \mathbb{T}$; the graph can be drawn in the unit square with opposite edges identified, as in FIGURE 3(a). Now tile the plane with integer translates of this square. Each curve in



The graph of $f : \mathbb{T} \to \mathbb{T}$

(a)

Graphs of lifts $F : \mathbb{R} \to \mathbb{R}$

(b)

FIGURE 3

FIGURE 3(b) is then the graph of a lift of f. We see that f has infinitely many lifts, any two of which differ by an integer. Each lift is strictly increasing on $\mathbb{R}$, and $F(x+1) = F(x) + 1$ for all $x \in \mathbb{R}$. (You'll soon be invited to prove these facts!)

To understand the behavior of orbits of f, we often study orbits of F:

$$o(x, F) = o(x) = \{F^i(x) : i \in \mathbb{Z}\},$$

which can be thought of as orbits of f "laid out" in the real line. For example, if $o(\theta, f)$ has passed $\pi(0) \in \mathbb{T}$ p times after q iterates, and if $x \in [0, 1)$, $\pi(x) = \theta$ and F is the lift satisfying $F(0) \in [0, 1)$, then $p \le F^q(x) < p + 1$.

A lift of a circle homeomorphism is particularly useful as it allows us to study iteration of a map defined on $\mathbb{R}$—for which we have the full arsenal of results from real analysis at our disposal.

Exercise 2. Show that for all $k \in \mathbb{Z}$, $R_\omega(x) = x + \omega + k$ is a lift of the rigid rotation r_ω. (For the remainder of this paper, we will set $k = 0$, so $R_\omega(x) = x + \omega$.)

Exercise 3. Show that a lift of an orientation-preserving circle homeomorphism is strictly increasing. Hint: Show F is strictly increasing on $[n, n+1)$ for all $n \in \mathbb{Z}$. Then use the continuity of F.

Exercise 4. Let F be a lift of an orientation-preserving circle homeomorphism f. Show that, for all integers n, $\pi F^n(x) = f^n \pi(x)$. That is, F^n is a lift of f^n for all $n \in \mathbb{Z}$.

Exercise 5. Show that any two lifts of a circle homeomorphism differ by a fixed integer k. Hint: To show that the *same* k works for all x, use the fact that continuous functions take connected sets to connected sets.

Exercise 6. (a) Show that a lift F of an orientation-preserving circle homeomorphism must satisfy $F(x+1) = F(x) + 1$ for all $x \in \mathbb{R}$. Hint: Given $x \in \mathbb{R}$, show $F(x+1) = F(x) + k$ for some $k \in \mathbb{Z}$. If $k \neq 1$, use the intermediate value theorem to contradict the fact f is a homeomorphism.
(b) Show $\forall x \in \mathbb{R}$, $\forall n, k \in \mathbb{Z}$, $F(x+k) = F(x) + k$ and $F^n(x+k) = F^n(x) + k$.

Exercise 7. Let F be a lift of an orientation-preserving circle homeomorphism.

(a) Suppose there exist $x \in \mathbb{R}$ and integers p and q such that $F^q(x) = x + p$. Show that $\pi(x)$ is a periodic point of period q for f. Such an x is called a *p/q-periodic point*.
(b) Suppose $\theta \in \mathbb{T}$ is a periodic point of period q. Show that there is an integer p such that, for all $x \in \pi^{-1}(\theta) = \{x \in \mathbb{R} : \pi(x) = \theta\}$, $F^q(x) = x + p$. Hint: Use the fact that if $x, y \in \pi^{-1}(\theta)$, then $x = y + k$ for some $k \in \mathbb{Z}$. Then use Exercise 6(b).

Knowing that $x \in \mathbb{R}$ is a p/q-periodic point yields information beyond the fact $\pi(x)$ is a period q point for f. As indicated in FIGURE 4, p counts the number of times the

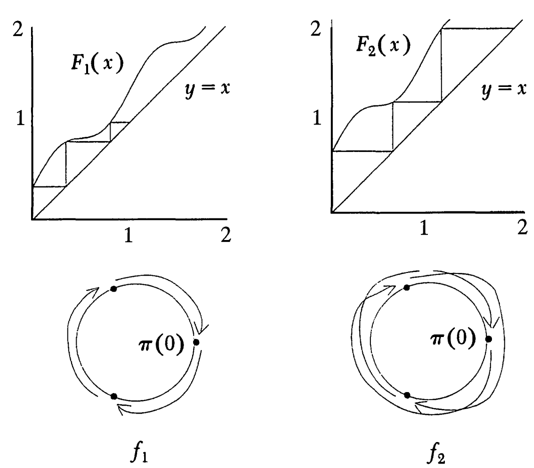


FIGURE 4
$\pi(0)$ is a period 3 point for f_1 and f_2. 0 is a $1/3$-periodic point for F_1 and 0 is a $2/3$-periodic point for F_2.

orbit of $\pi(x)$ traverses $\mathbb{T}$ every q iterates. The horizontal and vertical lines in FIGURE 4 (the "cobweb" diagram) represent a graphical interpretation of iteration [**8**, §1.3].

The *extended orbit* of $x \in \mathbb{R}$ under F is defined by

$$eo(x, F) = eo(x) = \{F^i(x) + j : i, j \in \mathbb{Z}\}.$$

Exercise 8. (a) Show that $eo(x) = \pi^{-1}(o(\pi(x), f))$. (b) Let $\omega \notin \mathbb{Q}$. Show that $eo(x, R_\omega)$ is dense in $\mathbb{R}$.

The rotation number

DEFINITION. Let F be a lift of an orientation-preserving circle homeomorphism $f: \mathbb{T} \to \mathbb{T}$. The *rotation number* $\rho(f)$ is defined, for any $x \in \mathbb{R}$, as the fractional part of

$$\lim_{n \to \infty} \frac{F^n(x) - x}{n}. \tag{1}$$

Remarks. The fractional part is chosen so that $\rho(f)$ is independent of the lift F (see Exercise 5). Henceforth, when the limit in (1) is used we will assume the fractional part has been taken.

The rotation number measures the average distance a point x travels per iterate of F, or, projecting onto $\mathbb{T}$, the average rotation per iterate of f.

Exercise 9. Show that, for all $x \in \mathbb{R}$, $\rho(r_\omega) = \omega$.

PROPOSITION 1. *Let F be a lift of an orientation-preserving circle homeomorphism. The rotation number $\rho(f)$ exists and its value is independent of x. That is, $\rho(f)$ is well-defined.*

Exercise 10. Show that $\rho(f)$, if it exists, is independent of x by completing the following steps:

(a) Let $k \in \mathbb{Z}$. Use Exercise 6 to show

$$\lim_{n \to \infty} \frac{F^n(x+k) - (x+k)}{n} = \lim_{n \to \infty} \frac{F^n(x) - x}{n}.$$

Hence to show $\rho(f)$ is the same for $x_1 \neq x_2$, we may assume $x_1, x_2 \in [0, 1)$, so that

$$x_2 - 1 \leq x_1 \leq x_2 + 1. \tag{2}$$

(b) Show

$$\lim_{n \to \infty} \frac{F^n(x_2) - x_2}{n} \leq \lim_{n \to \infty} \frac{F^n(x_1) - x_1}{n} \leq \lim_{n \to \infty} \frac{F^n(x_2) - x_2}{n}.$$

(Use each of inequality (2) and Exercises 3 and 6 twice.)

Exercise 11. Show that $\rho(f)$ exists by completing the following steps ([**8**, §1.14]).

(a) Show that if there is a p/q-periodic point x, then $\rho(f) = p/q$. Hint: Use x to compute $\rho(f)$. For a fixed integer q show there is a constant M such that $|F^r(y) - y| \leq M$ for all $y \in \mathbb{R}$ and all $0 \leq r < q$.

(b) Suppose there are no p/q-periodic points. (i) Use the intermediate value theorem to show that, for any $n > 0$, there is an integer k_n such that $k_n < F^n(x) - x < k_n + 1$ for all $x \in \mathbb{R}$. (ii) Apply this inequality m times with x-values 0, $F^n(0)$,

$F^{2n}(0), \ldots, F^{(m-1)n}(0)$, respectively. Add the resulting m inequalities to find $mk_n < F^{mn}(0) < m(k_n + 1)$. (iii) Show that

$$\left|\frac{F^{mn}(0)}{mn} - \frac{F^n(0)}{n}\right| < \frac{1}{n} \quad \text{and} \quad \left|\frac{F^{mn}(0)}{mn} - \frac{F^m(0)}{m}\right| < \frac{1}{m}.$$

Deduce that

$$\left|\frac{F^n(0)}{n} - \frac{F^m(0)}{m}\right| < \frac{1}{n} + \frac{1}{m},$$

and hence that $\left\{\frac{F^n(0) - 0}{n}\right\}_{n=1}^{\infty}$ is a Cauchy sequence. Conclude that $\rho(f)$ exists.

Rational rotation numbers

We have seen that if F has a p/q-periodic point then the rotation number equals p/q. Surprisingly, the converse is also true.

THEOREM 1 **[22]**. *Suppose F is a lift of an orientation-preserving circle homeomorphism. Then $\rho(f) = p/q$ if and only if F has a p/q-periodic point.*

This theorem provides an example of a true success story in dynamical systems theory: simply compute a limit to gain significant insight into the dynamics of the given map. If $\rho(f) \in \mathbb{Q}$, f has a periodic point (and, indeed, the long-term behavior of all orbits is completely determined—see Theorem 2). If $\rho(f) \notin \mathbb{Q}$, f has no periodic points. A natural jump to make in this case would be to conclude that all orbits are dense in $\mathbb{T}$. We will see, however, that this is not always the case.

Exercise 12. Prove Theorem 1. Hint: Exercise 11(a) provides one direction. For the other, suppose F has no p/q-periodic point. Argue that, for some $\epsilon > 0$ and for all $x \in \mathbb{R}$, either $F^q(x) > x + p + \epsilon$ or $F^q(x) < x + p - \epsilon$ (use the intermediate value theorem, compactness of $[0, 1]$, and Exercise 6). Now find a contradiction.

THEOREM 2. *Let F be a lift of an orientation-preserving circle homeomorphism with $\rho(f) = p/q$. Then, for all $x \in \mathbb{R}$, either*
(*i*) *x is a p/q-periodic point, or*
(*ii*) *there is a p/q-periodic point $x_0 \in \mathbb{R}$ with $|F^n(x) - F^n(x_0)| \to 0$ as $n \to \infty$.*

Case (ii) of Theorem 2 implies that asymptotically the orbit of $\pi(x)$ tends to a periodic orbit that traverses $\mathbb{T}$ p times every q iterates of f.

Exercise 13. Let $G : \mathbb{R} \to \mathbb{R}$ be continuous and strictly increasing. For $x \in \mathbb{R}$, show that either $|G^n(x)| \to \infty$ or $G^n(x) \to x_0$ as $n \to \infty$, where x_0 is a fixed point of G. Hint: Use properties of monotonic sequences of real numbers.

Exercise 14. Prove Theorem 2. Hint: Use Theorem 1 to conclude that $G(x) = F^q(x) - p$ has a fixed point x_0. Show that G has infinitely many fixed points via Exercise 6. Now use Exercise 13.

Irrational rotation numbers

The dynamics of a circle homeomorphism with irrational rotation number involves more delicate and much deeper mathematics (**[7]**, **[12]**, **[13]**, **[14]**, **[27]**), so only part of the story will be told here. Yet even this partial investigation leads to several surprising results.

We assume F is a lift of an orientation-preserving circle homeomorphism with $\rho(f) = \alpha \notin \mathbb{Q}$. Recall that all orbits under rigid rotation r_α are dense in $\mathbb{T}$. Must the same condition also hold for f? To answer this, we need the notion of a conjugacy.

DEFINITION. Two orientation-preserving circle homeomorphisms f and g are *conjugate* if there exists an orientation-preserving circle homeomorphism $h: \mathbb{T} \to \mathbb{T}$ such that $hf(\theta) = gh(\theta)$ for all $\theta \in \mathbb{T}$, i.e., such that the diagram

$$\begin{array}{ccc} \mathbb{T} & \xrightarrow{\ f\ } & \mathbb{T} \\ h\big\downarrow & & \big\downarrow h \\ \mathbb{T} & \xrightarrow{\ g\ } & \mathbb{T} \end{array}$$

commutes.

Remark. The definition of topological conjugacy (see, e.g., **[24**, §2.6**]**) requires that h be a homeomorphism satisfying $hf = gh$. We require in addition that h is orientation-preserving; this assures that conjugate orientation-preserving circle homeomorphisms have the same rotation number (**[12]**, **[16**, §11.1**]**).

Exercise 15. (a) Show that two conjugate orientation-preserving circle homeomorphisms have the same dynamics. That is, show $h(o(\theta, f)) = o(h(\theta), g)$ for all $\theta \in \mathbb{T}$. Hence, if θ is a periodic point of period n for f, for example, then $h(\theta)$ is a periodic point of period n for g. Likewise, $h^{-1}(o(\theta, g)) = o(h^{-1}(\theta), f)$.

(b) Show that f has a dense orbit if and only if g has a dense orbit. (This is an example of the fact that topological properties of orbits are also preserved by the conjugacy.)

We can now rephrase our question as follows: Given $\rho(f) = \alpha \notin \mathbb{Q}$, under what condition(s) is f conjugate to r_α? If the conjugacy exists, then all orbits of f are dense in $\mathbb{T}$. This problem is surprisingly difficult; Poincaré, for example, was unable to resolve it **[2]**. But we can make progress towards its solution via the following propositions.

PROPOSITION 2. *Let F be a lift of an orientation-preserving circle homeomorphism with $\rho(f) = \alpha \notin \mathbb{Q}$. Fix $x_0 \in \mathbb{R}$ and define $\overline{H}: eo(x_0) \to \mathbb{R}$ by*

$$\overline{H}\big(F^i(x_0) + j\big) = i\alpha + j \quad \textit{for all} \quad i, j \in \mathbb{Z}.$$

If $x_1,\ x_2 \in eo(x_0)$ and $x_1 < x_2$, then $\overline{H}(x_1) < \overline{H}(x_2)$. That is, $\overline{H}$ preserves the "$<$" ordering on $eo(x_0)$.

Proof idea: Suppose the conclusion is false. Then there exist $i,\ j,\ k,\ l \in \mathbb{Z}$ satisfying $F^i(x_0) + j < F^k(x_0) + l$ and $i\alpha + j \geq k\alpha + l$, or $(i - k)\alpha \geq l - j$. Note that $i \neq k$.

Case 1. Assume $i - k > 0$. Then $\alpha \geq (l - j)/(i - k)$. By Exercises 3 and 6, $F^i(x_0) + j < F^k(x_0) + l$ implies $F^{(i-k)}(x_0) - x_0 < l - j$. Now show this implies $\alpha \leq (l - j)/(i - k)$, and deduce a contradiction.

Exercise 16. Complete the argument for the $i - k < 0$ case.

Hence if $\rho(f) = \alpha \notin \mathbb{Q}$, the ordering of points in $eo(x_0, F)$ is the same as the ordering of points in $eo(0, R_\alpha)$. This would seem to kindle hope that a conjugacy between F and R_α exists, with an appropriate extension of $\overline{H}$ to the real line as the choice for the conjugacy. Although this is not always possible, we can say the following:

PROPOSITION 3. *Let F be a lift of an orientation-preserving circle homeomorphism $f:\mathbb{T}\to\mathbb{T}$ with $\rho(f)=\alpha\notin\mathbb{Q}$. There exists a map $H:\mathbb{R}\to\mathbb{R}$ such that*

(*i*) *H is continuous and increasing;*
(*ii*) *$\forall x\in\mathbb{R}$, $H(x+1)=H(x)+1$, and H is a lift of a map $h:\mathbb{T}\to\mathbb{T}$ such that the diagram*

$$\begin{array}{ccc} \mathbb{T} & \xrightarrow{\ f\ } & \mathbb{T} \\ h\big\downarrow & & \big\downarrow h \\ \mathbb{T} & \xrightarrow{\ r_\alpha\ } & \mathbb{T} \end{array}$$

commutes.

Remark. By "increasing" we mean nondecreasing; thus the map H need not be a conjugacy, as it may not be strictly increasing.

Proof. Fix $x_0\in\mathbb{R}$, and define $\overline{H}:eo(x_0)\to\mathbb{R}$ as in Proposition 2. Extend $\overline{H}$ to a map $H:\mathbb{R}\to\mathbb{R}$, $y\mapsto\sup\{\overline{H}(x):x\in eo(x_0),\ x<y\}$.

Exercise 17. Prove that $H:\mathbb{R}\to\mathbb{R}$ is increasing. Hint: Use Proposition 2.

Exercise 18. Prove that $H:\mathbb{R}\to\mathbb{R}$ is continuous. Hint: Suppose H is not continuous at $y_0\in\mathbb{R}$. Use Exercises 17 and 8(b) to derive a contradiction.

Exercise 19. Prove that for all $x\in\mathbb{R}$, $H(x+1)=H(x)+1$. Hint: First show that $eo(x_0,F)+1=eo(x_0,F)$, and that $H:eo(x_0,F)\to\mathbb{R}$ satisfies $H(x+1)=H(x)+1$ for all $x\in eo(x_0,F)$.

To complete the proof of Proposition 3, note that for $i,\ j,\ k,\ l\in\mathbb{Z}$,

$$\begin{aligned} H\big(F\big(F^i(x_0)+j\big)\big) &= H\big(F^{i+1}(x_0)+j\big)=(i+1)\alpha+j=H\big(F^i(x_0)+j\big)+\alpha \\ &= R_\alpha\big(H\big(F^i(x_0)+j\big)\big), \end{aligned}$$

so that $H\circ F=R_\alpha\circ H$ on $eo(x_0)$. Now for $y\in\mathbb{R}$,

$$\begin{aligned} HF(y) &= \sup\{\overline{H}(x):x\in eo(x_0),\ x<F(y)\} \\ &= \sup\{\overline{H}(F(x)):x\in eo(x_0),\ x<y\} \quad \text{(Exercise 6)} \\ &= \sup\{R_\alpha(\overline{H}(x)):x\in eo(x_0),\ x<y\} \\ &= R_\alpha H(y). \quad \square \end{aligned}$$

If the map $H:\mathbb{R}\to\mathbb{R}$ in Proposition 3 is not a homeomorphism, there must be an interval $J\subset\mathbb{R}$ on which H is constant, i.e., $H(J)=\{y_0\}$ for some $y_0\in\mathbb{R}$.

Exercise 20. (a) Show that for such an interval J, H is constant on $F^i(J)+j$ for all $i,\ j\in\mathbb{Z}$. Hint: Proposition 3.

(b) Show that H is locally constant on an open, dense subset of $\mathbb{R}$. Thus H is an example of a *Cantor function*; see FIGURE 5 for an approximation of the graph of a Cantor function arising in the next section. Hint: Use Proposition 3 and Exercise 8(b).

Exercise 21. Suppose f is an orientation-preserving circle homeomorphism with $\rho(f)=\alpha\notin\mathbb{Q}$. Show that if f has a dense orbit, then f is conjugate to r_α, and hence all orbits of f are dense in $\mathbb{T}$.

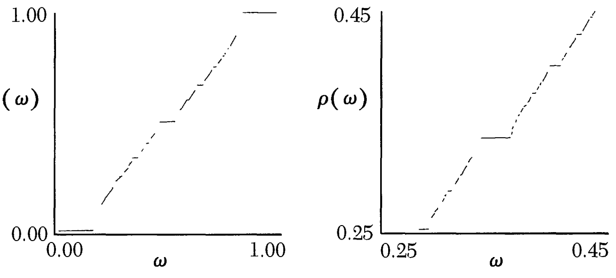


FIGURE 5
A plot of $\rho(\omega) = \rho(f_{\omega,1})$.

We see via Exercise 21 that the question of whether the map H is a homeomorphism depends upon the existence of a dense orbit for f. It is indeed curious to find that, in the absence of a dense orbit for f, the mapping H is a pathological Cantor function.

Not until 1932 did A. Denjoy provide a sufficient condition for determining when a circle homeomorphism with irrational rotation number α is conjugate to rigid rotation r_α. Surprisingly, the degree of smoothness of f is the deciding factor. For proofs of Denjoy's theorem see **[5]**, **[7]**, **[9]**, **[12]**, **[16]**, **[18]**, **[19]**, **[24]**.

Denjoy's Theorem **[7]**. *Let* $f:\mathbb{T}\to\mathbb{T}$ *be an orientation-preserving circle homeomorphism with* $\rho(f)=\alpha\notin\mathbb{Q}$. *If* f *is a* C^2*-diffeomorphism, then* f *is conjugate to* r_α.

Remarks. The requirement that f be twice continuously differentiable is sufficient but not necessary: the conjugacy exists if f is C^1 with f' of bounded variation (**[5]**, **[7]**, **[9]**, **[16]**). See **[9]** for a discussion of the role played by the bounded variation of f'. There also exist C^1 orientation-preserving circle diffeomorphisms with irrational rotation number that are not conjugate to r_α (**[7]**, **[8]**, **[12]**, **[13]**, **[27]**).

We note that a natural analogue of rotation number exists for maps and flows defined on higher dimensional tori. For an introduction to *rotation vectors* and the role they play in understanding the dynamics of such maps and flows, see **[26]**.

An example

We now link ideas from the preceding sections by considering the two-parameter *standard family* of circle maps

$$F:\mathbb{R}\to\mathbb{R},\ F_{\omega,\epsilon}(x) = x+\omega+\frac{\epsilon}{2\pi}\sin(2\pi x),\quad \omega,\epsilon\geq 0. \tag{3}$$

Note that if $\epsilon = 0$ we recover R_ω, so family (3) represents a perturbation of rigid rotation. The standard family arises in many physical systems as a model for periodically forced nonlinear oscillators (**[3]**, **[6]**, **[15]**, **[20]**).

Exercise 22. Show that for $\epsilon\leq 1$, $F_{\omega,\epsilon}$ is a lift of an orientation-preserving circle homeomorphism $f_{\omega,\epsilon}:\mathbb{T}\to\mathbb{T}$.

We would like to investigate how $\rho(f_{\omega,\epsilon})$ changes as the parameters ω and ϵ are varied. To this end, temporarily fix ϵ and let $F_\omega(x) = F_{\omega,\epsilon}(x)$. Define a function $\rho:[0,1)\to\mathbb{R}$, $\rho(\omega)=\rho(f_\omega)$.

Exercise 23. (a) Show that $\omega_1<\omega_2$ implies $\rho(\omega_1)\leq\rho(\omega_2)$, so that $\rho(\omega)$ is an increasing function. Hint: Exercise 3.

(b) Show $\rho(0)=0$ and $\lim_{\omega\to 1}\rho(\omega)=1$. Hint: Find a fixed point for F_0 and a 1/1-periodic point for $F_{1-\delta}$ for δ sufficiently small.

(c) Show $\rho:[0,1)\to\mathbb{R}$ is a continuous function of ω. Hint: Let $\omega_0\in[0,1)$ and $\epsilon>0$ be given. Pick $n\in\mathbb{Z}$ with $2/n<\epsilon$, and $k\in\mathbb{Z}$ satisfying $k-1<F^n_{\omega_0}(0)<k+1$. Using Exercises 3 and 6, show $m(k-1)<F^{mn}_{\omega_0}(0)<m(k+1)$. Pick $\delta>0$ so that $|\omega-\omega_0|<\delta$ implies $k-1<F^n_\omega(0)<k+1$, and show $|F^{mn}_\omega(0)-F^{mn}_{\omega_0}(0)|<2m$. Now deduce $|\rho(\omega)-\rho(\omega_0)|<\epsilon$.

Suppose $\rho(f_{\omega_0})=p/q$ for some ω_0. By Theorem 1, the function $G(x)=F^q_{\omega_0}(x)-p$ has a fixed point x_0. Let $\lambda=G'(x_0)$. If $\lambda\neq 1$ (so that the graph of G is not tangent to the line $y=x$ at $x=x_0$), the implicit function theorem yields a $\delta>0$ so that for $\omega\in(\omega_0-\delta,\omega_0+\delta)$, $F^q_\omega(x)-p$ has a fixed point. That is, for $\omega\in(\omega_0-\delta,\omega_0+\delta)$, F_ω has a p/q-periodic point and $\rho(\omega)=p/q$. If $\lambda=1$ there again exists such a δ, but the argument is more complicated ([**8**, §1.14], [**16**, §11.1]). In either case we have that if $\rho(\omega_0)=p/q$, then $\rho(\omega)=p/q$ on some interval of ω-values containing ω_0.

V.I. Arnol'd ([**1**], [**16**, §11.1]) showed that adding an arbitrarily small constant to an orientation-preserving circle homeomorphism with $\rho(f)\notin\mathbb{Q}$ changes the rotation number. In summary, we have that $\rho(\omega)$ is a continuous, increasing function, with $\rho(0)=0$ and $\rho(1)=1$. Moreover, to each rational in $[0,1)$ corresponds an interval on which ρ is constant. Since the rationals are dense in $[0,1)$, we see that $\rho(\omega)$ (surprisingly) provides another example of a Cantor function (see FIGURE 5).

To see how $\rho(f_{\omega,\epsilon})$ varies as a function of both parameters ω and ϵ, set $A_r=\{(\omega,\epsilon):\rho(f_{\omega,\epsilon})=r\}$. (The A_r are level sets of the function $\rho:(\omega,\epsilon)\mapsto\rho(f_{\omega,\epsilon})$.)

Exercise 24. (a) Show that $A_r=\{(r,0)\}$ for $\epsilon=0$. (b) Show that $A_{r_1}\cap A_{r_2}=\emptyset$ for $\epsilon\leq 1$ and $r_1\neq r_2$.

We have seen that $\rho(f_{\omega_0,\epsilon})=p/q$ implies the existence of an interval of ω-values for which $\rho(f_{\omega,\epsilon})=p/q$. The sets $A_{p/q}$ therefore have nonempty interior, and are known as *p/q-resonance zones* or *Arnol'd tongues*. For $r\notin\mathbb{Q}$, A_r will have empty interior due to Arnol'd's result. In this case A_r is a curve extending from $\epsilon=0$ to $\epsilon=1$ in the ω,ϵ-plane [**4**]. Note that in general no two of the A_r regions can intersect (Exercise 24(b)). We have then an amazingly intricate *bifurcation plot* for the standard family in the ω,ϵ-plane. A computer-generated plot of the boundaries of certain resonance zones for $0\leq\epsilon\leq 3$ is shown in FIGURE 6. Note that for $\epsilon>1$ these

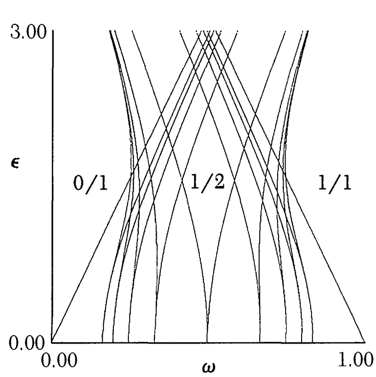


FIGURE 6

Boundaries of p/q-resonance zones for the standard family $F_{\omega,\epsilon}$ for $0\leq\epsilon\leq 3$. The boundaries of the 0/1, 1/6, 1/5, 1/4, 1/3, 1/2, 2/3, 3/4, 4/5, 5/6, and 1/1 horns are pictured from left to right.

regions can intersect ($F_{\omega,\epsilon}$ is no longer a homeomorphism); it turns out that for rationals r_1 and r_2 and $(\omega,\epsilon) \in A_{r_1} \cap A_{r_2}$, $F_{\omega,\epsilon}$ is chaotic in the sense that it has positive topological entropy [**4**].

Acknowledgment. Special thanks to friend and mentor G.R. Hall for permitting his lecture notes as the basis for this article. Thanks also to the referees for helpful comments.

REFERENCES

1. V.I. Arnol'd, Small denominators I : on the mapping of a circle into itself, *AMS Transl. Ser.* 2 46 (1965), 213–284.
2. G.D. Birkhoff, The work of Poincaré on differential equations, *Bull. Amer. Math. Soc.* 40 (1934), 363–366.
3. T. Bohr, P. Bak, and M.H. Jensen, Transition to chaos by interaction of resonances in dissipative systems. II. Josephson junctions, charge-density waves, and standard maps, *Phys. Rev. A* 30 (1984), 1970–1981.
4. P. Boyland, Bifurcations of circle maps, Arnol'd tongues, bistability, and rotation intervals, *Commun. Math. Phys.* 106 (1986), 353–381.
5. E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, NY, 1955.
6. P. Coullet, C. Tresser, and A. Arneodo, Transition to turbulence for doubly periodic flows, *Phys. Lett. A* 77 (1980), 327–331.
7. A. Denjoy, Sur les courbes définies par les equations différentielles à la surface du tore, *J. de Math. Pures et App.* 11 (1932), 333–375.
8. R.L. Devaney, *An Introduction to Chaotic Dynamical Systems* (2nd ed.), Benjamin-Cummings, New York, NY, 1986.
9. J. Guckenheimer and P.J. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York, NY, 1983.
10. V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, Englewood Cliffs, NJ, 1974.
11. J.K. Hale and H. Koçak, *Dynamics and Bifurcations*, Springer-Verlag, New York, NY, 1991.
12. G.R. Hall, The Dynamics of Circle and Annulus Maps, Regional Geometry Institute Lecture Notes, Boston University, 1990, to appear at `http://math.bu.edu/people/rockford`. Hard copy available by writing Hall at Boston University Department of Mathematics, 111 Cummington St., Boston, MA 02215.
13. G.R. Hall, A C^∞ Denjoy counterexample, *Ergod. Th. & Dynam. Sys.* 1 (1981), 261–272.
14. M. Herman, Sur la conjugaison différentiable des diffeomorphismes du cercle à des rotations, *Inst. Hautes Etudes Sci. Publ. Math.* 49 (1979), 5–233.
15. F. Hoppensteadt and J.P. Keener, Phase locking of biological clock, *J. Math. Biol.* 15 (1982), 339–349.
16. A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, UK, 1995.
17. W.S. Massey, *Algebraic Topology: An Introduction*, Springer-Verlag, New York, NY, 1987.
18. W. de Melo and S. van Strien, *One-dimensional Dynamics*, Springer-Verlag, New York, NY, 1993.
19. Z. Nitecki, *Differentiable Dynamics*, MIT Press, Boston, MA, 1971.
20. R. Perez and L. Glass, Bistability period doubling bifurcations and chaos in a periodically forced oscillator, *Phys. Lett. A* 90 (1982), 441–443.
21. L. Perko, *Differential Equations and Dynamical Systems*, Springer-Verlag, New York, NY, 1991.
22. H. Poincaré, Sur les courbes définies par les equations différentielles, *J. de Math. Pures et App.* t. **I** (1885), 167–244.
23. H. Poincaré, *New Methods of Celestial Mechanics* (ed. by D. Goroff), American Institute of Physics, Woodbury, NY, 1993.
24. C. Robinson, *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*, CRC Press, Boca Raton, FL, 1995.
25. See `http://www.oberlin.edu/~math/people/faculty/walsh/solutions.html`
26. J.A. Walsh, Rotation vectors for toral maps and flows: a tutorial, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* 5 (1995), 321–348.
27. J-C. Yoccoz, Il n'y a pas de contre-exemple Denjoy analytique, *C. R. Acad. Sci. Paris Sr. I Math.* 298 (1984), 141–144.

Chaos and One-to-Oneness

RODNEY NILLSEN
University of Wollongong
Wollongong, NSW 2522
Australia

1. Introduction

The term "chaos" was introduced in a mathematical context by T. Li and J. Yorke **[6]**. This paper considers chaos in the setting of an interval I of real numbers and a function f that maps I into I. Such a pair (I,f) is called a *dynamical system*, or simply a *system*. Given a system (I,f), the function $x \mapsto f(f(x))$ also maps I into I and is denoted by $f^{[2]}$. More generally, the function $x \mapsto f(f(\dots f(f(x)\dots)))$, where f appears n times, maps I into I, is denoted by $f^{[n]}$, and is called the n^{th} iterate of f, $f^{[1]}$ being f itself.

It is sometimes useful to think of I as the set of states of a physical system. Then $f(x)$ may be thought of as the state of the system that results from an initial state x after one unit of time. Thus, $f^{[n]}(x)$ is the state of the system after n time units, given that the initial state was x. A sequence of states $x, f(x), f^{[2]}(x), \dots, f^{[n]}(x), \dots$ describes the evolution of the system over time.

Whether the system (I,f) behaves chaotically is related to the behavior of the sequence of iterates $f, f^{[2]}, f^{[3]}, \dots, f^{[n]}, \dots$. Roughly speaking, "chaos" in the system means that the iterates of f "churn up" the points of I. For chaos to occur, this "churning up" process should not occur separately in disjoint parts of I.

A function f is *one-to-one* if $f(x) = f(y)$ only if $x = y$. We will prove that, for a one-to-one function f on I, chaotic behavior cannot occur if f is continuous, but that a type of chaotic behavior may occur if f has even *one* discontinuity. Although it is not the main concern of this paper, we note in passing that continuous functions that are *not* one-to-one—even seemingly harmless ones—may also lead to chaos.

2. Chaos, continuity, and one-to-one functions

If $f(x) = x$ then x is called a *fixed point* of f. If $f^{[n]}(x) = x$ for some $n \in \mathbb{N}$ ($\mathbb{N}$ denotes the natural numbers), then x is called a *periodic point* of f. In this case, the least $k \in \mathbb{N}$ such that $f^{[k]}(x) = x$ is called the *period* of the periodic point x. Note that a point is periodic if and only if it is a fixed point of some iterate of f, and that a fixed point is periodic with period 1.

We call the system (I,f) *transitive* if, for all $\varepsilon > 0$ and $x, y \in I$, there exist $z \in I$ and $n \in \mathbb{N}$ such that $|x - z| < \varepsilon$ and $|f^{[n]}(z) - y| < \varepsilon$. Intuitively, transitivity means that we can start arbitrarily close to any state but, over time, we can attain any state to within any desired degree of approximation.

We say the system (I,f) is *sensitive to initial conditions* if there is a number $\delta > 0$ such that, given any $\varepsilon > 0$ and any $x \in X$, there is $y \in X$ and a positive integer n such that $|x - y| < \varepsilon$ and $|f^{[n]}(x) - f^{[n]}(y)| > \delta$. Sensitivity to initial conditions is often taken as a hallmark of chaos, since it embodies the idea that even minute differences in initial data can lead to a marked divergence among possible evolutions of the system. In practice, such a system's behavior cannot be predicted by improving the accuracy of the initial data.

A widely used definition is that of Devaney [**3**, p. 119], according to which the system (I, f) is chaotic if

(i) every point of I is the limit of a sequence whose terms are periodic points of f (in which case the periodic points of f are said to be *dense* in I);
(ii) the system is transitive;
(iii) the system is sensitive to initial conditions.

(Sometimes we refer to the *function* f as being transitive, or sensitive to initial conditions, or chaotic, when the system (I, f) has the corresponding property or properties.) The three conditions above are not independent of each other. Banks, Davis, Stacey, Brooks, and Cairns [**1**] have proved in the general metric space context that, when the function is continuous, (i) and (ii) imply (iii). Glasner and Weiss have expressed the same result in a different but essentially equivalent form [**5**, Corollary 1.4]. Block and Coppel [**2**, p. 156] prove that for any continuous function mapping an interval I into itself, transitivity implies that there is a dense set of periodic points, a result also proved later by Vellekoop and Berglund [**10**]. Hence on intervals, as noted in [**10**], transitivity by itself implies sensitivity to initial conditions. It is also an immediate consequence of another result of Glasner and Weiss [**5**, Lemma 1.2] that, in a compact metric space, if the function is continuous and transitive but not one-to-one, then the system is sensitive to initial conditions.

We say a function f on I is *strictly increasing* (resp. *decreasing*) if $x < y$ implies $f(x) < f(y)$ (resp. $f(x) > f(y)$). The following result summarizes the properties of periodic points, transitivity, and sensitivity to initial conditions for continuous, one-to-one functions on an interval.

THEOREM 1. *Let I be a closed, bounded interval of real numbers having more than one point, and let f be a continuous and one-to-one function that maps I into itself. Then*:

(a) *Either f is strictly increasing or f is strictly decreasing on I.*
(b) *The system (I, f) is not transitive and is not sensitive to initial conditions.*
(c) *If f is strictly increasing, every periodic point of f is a fixed point of f. Moreover, if $f(x) \neq x$ for some $x \in I$, I does not have a dense set consisting of periodic or fixed points of f.*
(d) *If f is strictly decreasing, there is exactly one fixed point of f and all other periodic points of f have period 2. Moreover, if $f^{[2]}(x) \neq x$ for some $x \in I$, then I does not have a dense set consisting of periodic points of f.*

Proof. Assertion (a) is a standard result in real analysis (see, e.g., [**4**], [**9**]).

We prove (b) first in the case that f is strictly increasing. If $f(x) = x$ for all $x \in I$, then it is clear that neither transitivity nor sensitivity to initial conditions occurs. Suppose there is $x \in I$ with $f(x) \neq x$. Then either $f(x) > x$ or $f(x) < x$; we consider the former case. (Similar arguments apply if $f(x) < x$.) As f is continuous, we may assume that x is not an endpoint of I. Let $J = \{y \in I : y \geq x\}$. If $y \in J$, then $f(y) \geq f(x) > x$, so $f(y) \in J$. Thus f maps J into J, and $f^{[n]}$ maps J into J for all $n \in \mathbb{N}$. Now let $z \in I$ be such that $z < x$. Then there is $\varepsilon > 0$ such that $x - z > \varepsilon$. It follows that

$$|f^{[n]}(y) - z| > \varepsilon \quad \text{for all} \quad y \in J \quad \text{and all} \quad n \in \mathbb{N}. \tag{1}$$

So, transitivity fails.

Still assuming $f(x)>x$, we now show that (I,f) is not sensitive to initial conditions. Since $f(x)>x$, and since f is strictly increasing, it follows that $f^{[2]}(x)=f(f(x))>f(x)>x$, that $f^{[3]}(x)>f^{[2]}(x))>f(x)>x$, and that, for all $n\in\mathbb{N}$,

$$f^{[n+1]}(x)>f^{[n]}(x)>f^{[n-1]}(x)>\cdots>f^{[2]}(x)>f(x)>x.$$

Thus, in this case, $\{f^{[n]}(x)\}$ is an increasing sequence in I, and as I is closed and bounded, this sequence must converge to a limit $f^*(x)$. Note that $f^*(x)\in I$ and $f^*(x)>x$.

We will show that $y<f(y)$ for all $y\in(x,f^*(x))$. Otherwise, $f(y)\le y$, and we have $x<f(x)<f(y)\le y<f^*(x)$. Because f is strictly increasing we have, for all $n\in\mathbb{N}$,

$$f^{[n]}(x)<f^{[n+1]}(y)\le f^{[n]}(y)\le\cdots\le f(y)\le y<f^*(x).$$

Hence $f^*(x)=\lim_{n\to\infty}f^{[n]}(x)\le y<f^*(x)$, which is impossible. Thus $f(y)>y$, and it follows that, for all $y\in(x,f^*(x))$,

$$y<f(y)<f(f^*(x))=\lim_{n\to\infty}f^{[n+1]}(x)=f^*(x).$$

Now let $\delta>0$. There is $\delta'>0$ such that $\delta'<\delta$ and $(f^*(x)-\delta',f^*(x))\subseteq(x,f^*(x))$. Let $K=(f^*(x)-\delta',f^*(x))$. If $y\in K$, then $x<y<f(y)<f^*(x)$, so f maps K into K. Thus, $f^{[n]}$ maps K into K for all $n\in\mathbb{N}$. We now can say that

$$\left|f^{[n]}(z_1)-f^{[n]}(z_2)\right|<\delta'<\delta \quad \text{for all} \quad z_1,z_2\in K \quad \text{and all} \quad n\in\mathbb{N}.$$

This shows that the system is not sensitive to initial conditions, and completes the proof of (b) for the case when f is strictly increasing. We will prove (b) for strictly decreasing f after the proof of (d) below.

Next we prove (c). Let x be a periodic point of f with period p, say. Then if $f(x)>x$, as above, it would follow that

$$x<f^*(x)=\lim_{n\to\infty}f^{[n]}(x)=\lim_{n\to\infty}f^{[np]}(x)=x,$$

which is not possible. Thus, if x is periodic and $f(x)\ge x$, then $f(x)=x$ and x is actually a fixed point. Similarly, if x is periodic and $f(x)\le x$, then $f(x)=x$ and x is again a fixed point. Thus every periodic point of f is fixed. So, if the set of periodic points is dense, then the fixed points are dense. Hence, by continuity, $f(x)=x$ for all $x\in I$. This proves (c).

To prove (d), assume that f is strictly decreasing. Let the interval I be $[u,v]$. If $f(u)=u$, then $u=f(u)>f(v)\ge u$, which is impossible. Hence, $f(u)>u$; a similar argument shows that $f(v)<v$. Thus, the function $x\mapsto f(x)-x$ is continuous, strictly positive at u, and strictly negative at v. Therefore, there is a point in I at which this function vanishes; this point is a fixed point for f.

Now if x_1 and x_2 are distinct fixed points of f with $x_1<x_2$, then $x_1=f(x_1)>f(x_2)=x_2$, which is impossible. Hence, f has exactly one fixed point. To complete the proof, observe that because f is strictly decreasing, $f^{[2]}$ is strictly increasing on I. The remainder of (d) follows by applying the conclusion of (c) to the function $f^{[2]}$.

Finally, consider (b) in the case that f is strictly decreasing. Let $I=[u,v]$ and let x_0 be the unique fixed point of f. Let $L=[u,x_0]$ and $M=[x_0,v]$. As f is strictly decreasing, we see easily that f maps L into M and M into L, so that $f^{[2]}$ maps L into L and M into M. Also, $f^{[2]}$ is strictly increasing as f is strictly decreasing. Hence, the conclusion in equation (1) may be applied to $f^{[2]}$ and L in place of f and

I, and shows there is a closed subinterval L' of L, a point $z \in L$ with $x_0 \neq z$, and a number $\varepsilon > 0$ such that $|f^{[2n]}(y) - z| > \varepsilon$, for all $y \in L'$ and all $n \in \mathbb{N}$. Let M' be the closed subinterval $f(L')$ of M. Then $z \notin M'$, so there is $\varepsilon' > 0$ such that $|w - z| > \varepsilon'$ for all $w \in M'$. It follows that

$$|f^{[n]}(y) - z| > \min(\varepsilon, \varepsilon') \quad \text{for all} \quad y \in L' \quad \text{and all} \quad n \in \mathbb{N},$$

so f is not transitive. Using these ideas, it is straightforward to adapt the proof of the non-occurrence of sensitivity in the case of strictly increasing f to the case of strictly decreasing f. Details are left to the reader. □

The results of **[2, 10]** say, in the context of Theorem 1, that transitivity implies sensitive dependence. Thus, in Theorem 1, the failure of transitivity could have been deduced from the failure of having sensitive dependence upon initial conditions. Theorem 1 shows that when f is continuous and one-to-one, the system (I, f) is not chaotic, as it is not transitive and not sensitive to initial conditions. Although some of the conclusions of Theorem 1 are almost certainly known, the author is unaware of any reference.

3. Fractional parts, discontinuity, and chaos

Theorem 1 requires that the function be continuous at *every* point. We now weaken the assumption of continuity—but in the least possible way. We shall see that if the function has even one point of discontinuity, chaotic behavior may occur. We will construct a family of one-to-one functions, each function in the family mapping $[0, 1)$ onto itself and having one point of discontinuity. Each function in the family behaves chaotically (in a sense weaker than that of Devaney).

Every real number x may be written uniquely as the sum of an integer and a number $\text{frac}(x)$ in $[0, 1)$; $\text{frac}(x)$ is called the *fractional part* of x. It is easy to show that

$$\text{frac}(\text{frac}(x) + y) = \text{frac}(x + y). \tag{2}$$

Now, for each $\alpha \in \mathbb{R}$, define the function $f_\alpha : [0, 1) \to [0, 1)$ by

$$f_\alpha(x) = \text{frac}(x - \alpha). \tag{3}$$

FIGURE 1 shows the function $f_{0.3}$.

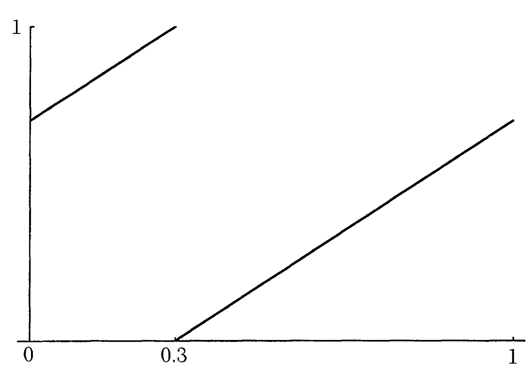


FIGURE 1
Graph of $f_{0.3}$

An alternative description of f_α is

$$f_\alpha(x)=\begin{cases} x+1-\operatorname{frac}(\alpha), & \text{if } 0\le x<\operatorname{frac}(\alpha);\\ x-\operatorname{frac}(\alpha), & \text{if } \operatorname{frac}(\alpha)\le x<1.\end{cases} \tag{4}$$

Clearly, each function f_α is continuous and increasing on each of the intervals $[0,\operatorname{frac}(\alpha))$ and $[\operatorname{frac}(\alpha),1)$, but when α is not an integer f_α is discontinuous at $\operatorname{frac}(\alpha)$. Thus, when α is not an integer, the graph of f_α looks just the same as the graph of $f_{0.3}$ in FIGURE 1, except that the discontinuity is at $\operatorname{frac}(\alpha)$ instead of 0.3. Also, f_α is one-to-one and maps $[0,1)$ onto $[0,1)$. It is not difficult to check, using equations (2) and (3), that for real α and β,

$$f_\alpha\circ f_\beta=f_{\alpha+\beta}\,. \tag{5}$$

Another geometric way of thinking of f_α involves the unit circle in the complex plane. Let $\mathbb{T}$ denote the set of complex numbers of modulus 1. The function $x\mapsto\exp(2\pi ix)$ is a one-to-one correspondence between $[0,1)$ and $\mathbb{T}$. Using this correspondence, we can think of the function f_α on $[0,1)$ as the function on $\mathbb{T}$ that is given by rotation through the angle $-2\pi\alpha$. Thus, if we replace α by $-\beta$ and β by $-\alpha$, the identity $f_\alpha\circ f_\beta=f_{\alpha+\beta}$ becomes $f_{-\beta}\circ f_{-\alpha}=f_{-(\alpha+\beta)}$; we may think of this as saying "if a point is rotated through an angle $2\pi\alpha$, and then is further rotated through an angle $2\pi\beta$, the total effect is to rotate the point through an angle $2\pi(\alpha+\beta)$." The interpretation of f_α as a rotation on $\mathbb{T}$ does not preserve continuity of functions. For, it is clear that a rotation through $2\pi\alpha$ on $\mathbb{T}$ is continuous, but the corresponding function f_α is not continuous on $[0,1)$ (unless α is an integer). This phenomenon arises because, although the function $x\mapsto\exp(2\pi ix)$ is continuous from $[0,1)$ onto $\mathbb{T}$, its inverse is *not* continuous.

The proof of the following theorem uses the definition of f_α. However, it is worthwhile to consider some of the conclusions and parts of the proof in light of the interpretation of f_α as a rotation on the unit circle. For example, it is obvious that a rotation on $\mathbb{T}$ is one-to-one and maps $\mathbb{T}$ onto $\mathbb{T}$.

THEOREM 2. *Let* $\alpha\in\mathbb{R}$. *Then:*

(a) f_α *is one-to-one on* $[0,1)$, *and maps* $[0,1)$ *onto* $[0,1)$.
(b) f_α *has a fixed point in* $[0,1)$ *if and only if* α *is an integer, in which case* $f_\alpha(x)=x$ *and every point is a fixed point.*
(c) *If* α *is not an integer then* f_α *has a discontinuity at* $\operatorname{frac}(\alpha)$, *but* f_α *is continuous and increasing on each of the intervals* $[0,\operatorname{frac}(\alpha))$ *and* $[\operatorname{frac}(\alpha),1)$.

Proof. The result is immediate from equation (3) or (4).

LEMMA. *Let* $\alpha\in\mathbb{R}$. *Then* f_α *has a periodic point if and only if* α *is rational. In this case, if* $\alpha\neq 0$ *and* $\alpha=\pm p/q$, *where* p,q *are in* $\mathbb{N}$ *and have no common factor, every point of* $[0,1)$ *is a periodic point of* f_α *with period* q.

Proof. Assume that $\alpha=p/q$ where $p,q\in\mathbb{N}$ have no common factor. Then if $x\in[0,1)$ and $n\in\mathbb{N}$, $f_\alpha^{[n]}(x)=x\Leftrightarrow f_{n\alpha}(x)=x\Leftrightarrow n\alpha$ is an integer, by part (b) of Theorem 2 and equation (5). It follows that f_α has a periodic point if and only if α is rational. So assume that $\alpha=p/q$ where p and q are integers having no common factor. Then

$$f_\alpha^{[n]}(x)=x\Leftrightarrow f_{n\alpha}(x)=x\Leftrightarrow f_{np/q}(x)=x\Leftrightarrow np/q \text{ is an integer}\Leftrightarrow n \text{ is divisible by } q.$$

If n *is* divisible by q, then $f_\alpha^{[n]}(x) = f_{np/q}(x) = x$ for all x, so in this case every point of $[0, 1)$ is periodic with period no greater than n. Since the smallest possible value of n that is divisible by q is q itself, every point of $[0, 1)$ has period q. □

The following theorem, the main result in this paper, is in contrast to Theorem 1. It asserts that if the one-to-one function f has even one point of discontinuity, then the system (I, f) may have transitivity and sensitivity to initial conditions, thus exhibiting a form of chaotic behavior.

THEOREM 3. *Let α be an irrational number. Then the dynamical system* $([0, 1), f_\alpha)$ *has no periodic points, but it is transitive and it is sensitive to initial conditions.*

Proof. The fact that f_α has no periodic points is immediate from the Lemma. Now we shall repeatedly use the fact that, since α is irrational, the set $\{\text{frac}(n\alpha): n = 1, 2, 3, \dots\}$ is dense in $[0, 1)$ (see, e.g., [**3**, pp. 211–212] or [**8**, p. 75]).

Now we prove transitivity. Let $\varepsilon > 0$. Let $x, y \in [0, 1)$, with $x \leq y$. Let $\eta > 0$ be such that $\eta < \varepsilon$ and $1 - y - \eta > 0$. Then $0 < 1 + x - y - \eta < 1$, so there is $n \in \mathbb{N}$ such that $1 + x - y - \eta < \text{frac}(n\alpha) < 1 + x - y + \eta$. This implies that $x < 1 + x - y - \eta < \text{frac}(n\alpha)$ and that $y - \eta < x - \text{frac}(n\alpha) + 1 < y + \eta$. Then, recalling the description of $f_{n\alpha}$ given in equation (4), we obtain

$$f_{n\alpha}(x) = x - \text{frac}(n\alpha) + 1 \in (y - \eta, y + \eta) \subseteq (y - \varepsilon, y + \varepsilon).$$

Thus, as $f_{n\alpha} = f_\alpha^{[n]}$, we have $|f_\alpha^{[n]}(x) - y| < \varepsilon$. (The conclusion follows in a similar way if $x > y$.) This shows transitivity.

To prove sensitivity, let $x \in [0, 1)$ and let $0 < \varepsilon < 1$. Let $y \in [0, 1)$ be such that $x < y < x + \varepsilon$. Choose $n \in \mathbb{N}$ such that $x < \text{frac}(n\alpha) < y$. It follows that $|x - y| < \varepsilon$ and

$$\begin{aligned} |f_\alpha^{[n]}(x) - f_\alpha^{[n]}(y)| &= |f_{n\alpha}(x) - f_{n\alpha}(y)| \\ &= |1 - \text{frac}(n\alpha) + x - y + \text{frac}(n\alpha)| \\ &= |1 + x - y| \\ &= 1 - (y - x) \\ &\geq 1 - \varepsilon, \end{aligned}$$

which completes the proof. □

Remarks. The failure to have a dense set of periodic points means that f_α does not satisfy Devaney's definition of chaos, but f_α *does* produce a dense set of "approximately periodic" points in the sense now described. Let $\varepsilon > 0$ and let $x \in [0, 1)$. Then as f_α is transitive by Theorem 3, there is $y \in [0, 1)$ and $n \in \mathbb{N}$ such that $|x - y| < \varepsilon/2$ and $|x - f_\alpha^{[n]}(y)| < \varepsilon/2$, so that $|y - f_\alpha^{[n]}(y)| < \varepsilon$. So, if in a general system (I, f) an ε-periodic point is defined to be a point y such that $|f^{[n]}(y) - y| \leq \varepsilon$ for some $n \in \mathbb{N}$, the argument shows that for each $\varepsilon > 0$, the set of ε-periodic points of $([0, 1), f_\alpha)$ is dense in $[0, 1)$. (Periodic points themselves may be regarded as "0-periodic" points.) Since f_α is transitive and has sensitivity to initial conditions, it therefore seems reasonable to regard the system $([0, 1), f_\alpha)$ as chaotic, although in a slightly weaker sense than that of Devaney.

4. Notions of chaos

There are other definitions of chaos apart from the one used by Devaney. For example, Stephen Wiggins [**12**, p. 57] omits any consideration of the periodic points and defines a certain type of dynamical system to be chaotic if it is transitive and has sensitive dependence on initial conditions. Theorem 3 shows that when α is irrational, the system (I, f_α) satisfies this definition of chaos.

Other notions of chaos have involved the concept of a scrambled set. Given a system (I, f), a subset S of I is said to be *scrambled* if

(i) for all $x, y \in S$ with $x \neq y$,

$$\limsup_{n\to\infty} |f^{[n]}(x) - f^{[n]}(y)| > 0 \quad \text{and} \quad \liminf_{n\to\infty} |f^{[n]}(x) - f^{[n]}(y)| = 0;$$

(ii) for any $x \in S$ and any periodic point z in I,

$$\limsup_{n\to\infty} |f^{[n]}(x) - f^{[n]}(z)| > 0.$$

When f is continuous and has a periodic point of period 3, Li and Yorke **[6]** showed that the system (I, f) has an uncountable scrambled set (they did not use this term). The existence of an uncountable scrambled set has sometimes been taken as a definition of chaos (see **[2**, pp. 143–145] for relevant results and a discussion). In the case of the system $([0,1), f_\alpha)$ when α is irrational, there is no scrambled set. For, if $x, y \in [0,1)$ with $x < y$, and if $n \in \mathbb{N}$,

$$\begin{aligned} |f_\alpha^{[n]}(x) - f_\alpha^{[n]}(y)| &= |f_{n\alpha}(x) - f_{n\alpha}(y)| \\ &= \begin{cases} 1 - |x-y| & \text{if} \quad x < \operatorname{frac}(n\alpha) \le y; \\ |x-y| & \text{otherwise.} \end{cases} \end{aligned}$$

It follows that

$$\liminf_{n\to\infty} |f_\alpha^{[n]}(x) - f_\alpha^{[n]}(y)| = \min(|x-y|, |1-|x-y||) > 0,$$

which proves the non-existence of a scrambled set in this case. This conclusion emphasizes that when α is irrational, we should consider the system (I, f_α) to be chaotic in a weaker sense than is often the case.

Another notion used in chaos is that of *topological entropy*. This is an element of $[0,\infty]$ associated with a dynamical system (I, f) when f is continuous; we can think of it as measuring the "long term loss of information" in the system. Chaos and entropy are linked in that the topological entropy of (I, f) is positive if and only if (I, f) is chaotic, where chaos is taken in the sense of **[2**, p. 33] (this result is proved in **[2**, p. 218]). Now when α is irrational, f_α is discontinuous, and so there are complications about assigning an entropy to $([0,1), f_\alpha)$. However, f_α may be thought of as a rotation on the circle group $\mathbb{T}$ and any such rotation is known to have zero topological entropy **[11**, p. 179]. Also, f_α may be assigned a *measure-theoretic entropy* **[11**, Chapter 4] which, by an adaptation of the proof of Theorem 4.24 in **[11**, pp. 100–101], is seen to be 0. Then, Theorem 8.6 in **[11**, p. 188] makes it clear that a natural way of assigning topological entropy to $([0,1), f_\alpha)$ would give the value 0. So, from this viewpoint, the case of the system (I, f_α) for irrational α shows that for a function with a single discontinuity, a form of chaos may occur while the "topological entropy" is zero, which contrasts with the result mentioned above for continuous f. A detailed comparison of related notions of chaos, including a very useful table, may be found in the recent article by Martelli, Dang, and Seph **[7]**.

For a dynamical system (I, f), all notions of chaos mentioned so far have involved properties of the *evolution* or *asymptotic properties* of the iterates of the function f.

However, instead of concentrating upon asymptotic behavior, we might consider asking: *Given a dynamical system* (I, f), *how chaotic has an initial state of the system become after n steps*? This latter question is imprecise, but it raises the point that it may be reasonable to consider "chaos" as a particular state, or subset of states, of a given system, rather than in terms of how the whole system evolves over time.

One way of thinking about chaos in relation to the function f_α is to regard the point $\mathrm{frac}(\alpha)$, the single discontinuity of f_α, as a "point of chaos." Since f_α has a jump discontinuity from 1 to 0 at $\mathrm{frac}(\alpha)$, f_α "attempts" to cover all of $[0,1)$ at $\mathrm{frac}(\alpha)$ in the sense that f_α tries to jump from 1 to 0 near that point. Now after n iterations, as $f_\alpha^{[n]} = f_{\mathrm{frac}(n\alpha)}$, the "point of chaos" has moved to another *single* point, namely $\mathrm{frac}(n\alpha)$, and it has not spread to become a multiplicity of "chaotic" points. In this sense, chaos has not spread throughout the system, but has simply changed its position. However, it is a famous result of H. Weyl (see, e.g., **[8]**) that if α is irrational, the sequence $\{\mathrm{frac}(n\alpha)\}$ is "uniformly distributed" in the sense that for each subinterval J of $[0,1)$, the asymptotic proportion of terms of $\{\mathrm{frac}(n\alpha)\}$ that fall into J is equal to the length of J. That is, if $\#A$ denotes the number of elements in a finite set A and $|J|$ denotes the length of J, then for each subinterval J of $[0,1)$,

$$\lim_{n\to\infty}\left(\frac{\#\{j : 1 \le j \le n \quad \text{and} \quad \mathrm{frac}(j\alpha) \in J\}}{n}\right) = |J|.$$

Thus, although chaos in the case of f_α occurs in different parts of the system with successive iterations, we can still say that all parts of the system are "uniformly affected" in the sense that the sequence of points where chaos occurs with successive iterations is uniformly distributed as described by Weyl's theorem.

Acknowledgment. I am most grateful for the interest and enthusiasm of the students to whom I gave lectures at the University of Wollongong in 1994 and 1995 on "Topology and Chaos"; and which led me to write the present article. I am indebted to Michel Lapidus and Keith Tognetti for helpful discussions about chaos, and to Peter Stacey for drawing my attention to the work of Glasner and Weiss. I would like to thank the referees for their many helpful and detailed comments on an earlier version of this article.

REFERENCES

1. J. Banks, G. Davis, P. Stacey, J. Brooks, and G. Cairns, On Devaney's definition of chaos, *Amer. Math. Monthly*, 99 (1992), 332–334.
2. L. S. Block and W. A. Coppel, *Dynamics in One Dimension*, Lecture Notes in Mathematics, Volume 1513, Springer-Verlag, Berlin, 1991.
3. R. Devaney, *A First Course in Chaotic Dynamical Systems*, Addison-Wesley, Reading, MA, 1992.
4. J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, NY, 1960.
5. E. Glasner and B. Weiss, Sensitive dependence on initial conditions, *Nonlinearity*, 6 (1993), 1067–1075.
6. T. Y. Li and J. A. Yorke, Period three implies chaos, *Amer. Math. Monthly*, 82 (1975), 985–992.
7. M. Martelli, M. Dang and T. Seph, Defining chaos, *Math. Mag.*, 71(1998), 112–122.
8. I. Niven, *Irrational Numbers*, Carus Mathematical Monographs, No. 11, MAA, Washington, DC, 1967.
9. M. Spivak, *Calculus*, Benjamin, New York, NY, 1967.
10. M. Vellekoop and R. Berglund, On intervals, transitivity = chaos, *Amer. Math. Monthly*, (101) 1994, 353–355.
11. P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, Berlin, 1982.
12. S. Wiggins, *Chaotic Transport in Dynamical Systems*, Interdisciplinary Applied Mathematics Series, Volume 2, Springer-Verlag, Berlin, 1992.

When and Why Do Water Levels Oscillate in Three Tanks?

KEITH M. KENDIG
Cleveland State University
Cleveland, OH 44115

In his book *I Want to be a Mathematician*, Paul Halmos offers a one-word description of how he studies mathematics: *examples*. He goes on to say:

> ...it's examples, examples, examples that, for me, all mathematics is based on, and I always look for them. I look for them first, when I begin to study, I keep looking for them, and I cherish them all [**1**, pp. 61–62, 64].

1. The example of one tank

In elementary differential equations, there's one particular example that is as educational as it is simple: a leaking tank of water. In its barest form, this is just a tin can, open at the top, with water draining out of a hole at the bottom. It has a much deeper side, though, because it models a whole host of familiar real-world phenomena. For example, "water" can be heat flowing out of a cooling cup of coffee; or it can be electric charge draining from a capacitor; or even carbon C-14 leaking (decaying) to normal C-12, familiar in radioactive dating. The leaking tank is obligingly versatile, too. For example, one can run a movie of it backwards; this corresponds to exponential growth—from population growth to increasing principal in a savings account. Or one can take an empty tank and push it part way, bottom first, into a large body of water such as a lake, causing water to flow *into* the tank through the hole; this models, for instance, an abandoned cold soda warming up to room temperature, or an ocean liner reaching cruise speed. If the submerged tank's water level starts higher than the lake's, it mimics a parachutist slowing down to terminal speed.

With the right assumptions, the differential equation governing a leaking tank directly reflects the physical setup. To this end, we begin by supposing that the tank is cylindrical with base area 1; this allows us to track the tank's water volume using the water height $y(t)$. Next, this height will always approach a steady-state level, and for simplicity we choose the origin of the y-axis to be this level. Finally, if k denotes the cross-sectional area of its drain, we choose units so that the water level y satisfies $y' = -ky$; this equation and the physical setup are now direct translations of each other: at any time, the flow rate equals drain area times water height. The solution of this equation is $y(t) = A_0 e^{-kt}$, where A_0 is the initial water level. Notice that the larger k is, the faster the water approaches its steady-state level.

Interestingly, the physical dimensions of a tank translate to, and even clarify, various physical notions. As just two examples, the thermal equivalents of water volume, water height, tank base area, and drain area are: heat, temperature, specific heat, and thermal conductivity. Electrical counterparts are: charge, voltage, capacitance, and electric conductance (the reciprocal of resistance). Learning to translate to and from "tank-lish" helps one unify parts of the physical world; one can profitably get hooked on it.

2. Two tanks

Tanks are obliging in yet another way: one can connect them together, providing a pictorial means of introducing first order linear systems to students. That is actually how this article arose—in the classroom, instead of starting off with an abstract system of differential equations, I began with a simple picture: two tanks coupled together. This easily translates to a system of two coupled differential equations, and this physical connection extends not only to the solution, but even to the eigenvalues and eigenvectors. This approach engaged students much more easily, left them with more understanding, and frequently piqued their curiosity. Here's the picture:

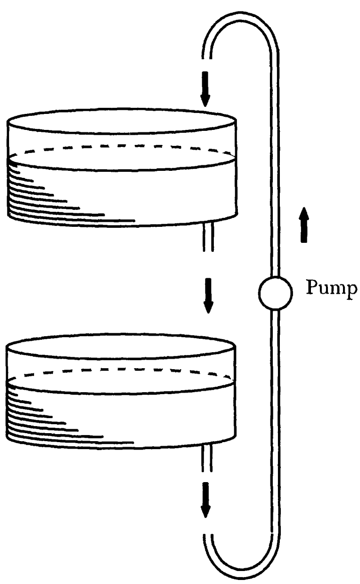


FIGURE 1
Two coupled tanks

Water drains from the top tank into the bottom one; the pump returns to the top tank water that has (gravity-) drained from the bottom tank. Let y_i and k_i be the water height and drain hole area for tank T_i. One can directly translate from picture to equations if, as before, both tanks are cylindrical with common base area 1, with remaining units chosen so that, for each T_i in isolation, the water level y_i satisfies $y_i' = -k_i y_i$. One can now write down the system of differential equations by "contemplation of the figure": water draining from T_1 contributes a negative rate of height increase $y_1 = -k_1 y_1$ in T_1, while feedback from T_2 into T_1 gives a positive contribution. The net rate for T_1 is therefore $y_1' = -k_1 y_1 + k_2 y_2$. Similar reasoning gives y_2', and we have the following system of equations corresponding to FIGURE 1:

$$\begin{aligned} y_1' &= -k_1 y_1 + k_2 y_2; \\ y_2' &= +k_1 y_1 - k_2 y_2. \end{aligned} \tag{1}$$

The solution to this system is easily checked to be

$$\begin{aligned} y_1(t) &= A_1 + Be^{-(k_1+k_2)t}; \\ y_2(t) &= A_2 - Be^{-(k_1+k_2)t}. \end{aligned} \tag{2}$$

The initial levels are $A_1 + B$ and $A_2 - B$, and the final levels are A_1 and A_2 (evident by choosing $t = 0$ and $t = \infty$). The exponent $-(k_1 + k_2)$ is one of the two eigenvalues. Physically, the magnitude $(k_1 + k_2)$ of this eigenvalue is the sum of the drain hole areas, and determines how fast the associated eigenvector varies. This eigenvector is the varying part of the vector solution (y_1, y_2)—that is, it's $(Be^{-(k_1+k_2)t}, -Be^{-(k_1+k_2)t})$. For any t, its components represent respective differences of current water levels from steady-state water levels; of course, a large $(k_1 + k_2)$ means the levels approach their steady-states quickly. Notice that since the system is conservative (no water is lost to the environment), the total amount of water in the system is constant through time; adding these two amounts, say by adding the two equations in (2), gives the constant $A_1 + A_2$. Notice that the amount above steady state in one tank is always equal to the amount below steady state in the other one.

An eigenvector linearly independent from the above is the *non*-varying part (A_1, A_2) of (y_1, y_2); physically, its components represent respective steady state levels. Each constant A_i is actually multiplied by $e^{0t} = 1$; the associated eigenvalue is therefore 0 and, as before, this eigenvalue determines how fast the associated eigenvector varies—of course in this case, it doesn't vary at all.

Though we consider mostly conservative systems, one can easily draw a picture for a nonconservative system by introducing an additional hole at the bottom of a tank, letting water leak to the environment. "Eigen-data" still continue to have physical meaning; for example, both water levels now go to zero, and from this one correctly predicts that both eigenvalues must be negative. One can graphically explore these physical connections using the PhasePort view in the program PHASER, included in [**3**].

3. Three tanks

Adding a third tank appears harmless enough:

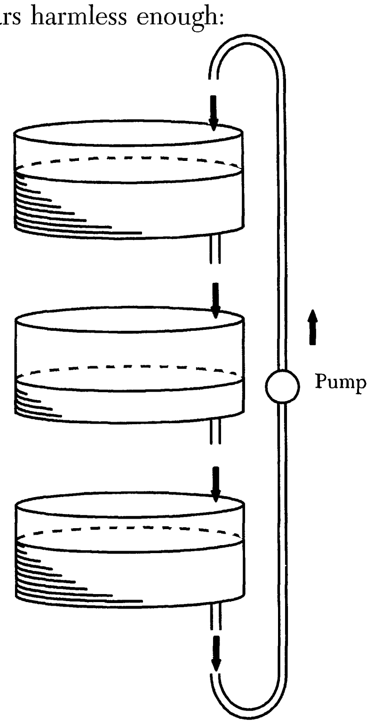


FIGURE 2
When will water levels oscillate in these three coupled tanks?

It seems reasonable that each level should eventually rise or fall to its final height, as with one and two tanks. The levels *can* do that. It is somewhat harder to believe (but true) that *more often than not, the levels oscillate forever*! To find examples of each kind of behavior, let's observe that unless the levels are at steady state, the amounts of water in each tank have no say in whether oscillation will occur—that's the job of the drain sizes. (This nicely highlights the difference between "initial" and "system" constants: precisely what the water levels happen to be initially—or at any other single time—doesn't influence the system's behavior, while system constants like drain hole sizes, do.) Here are two examples, both based on our picture. In one, the water levels vary monotonically; in the other, the levels oscillate forever. Can you guess which is which?

(a) One drain has diameter 1 inch, the other two have diameter $\frac{1}{2}$ inch.

(b) All three drains are pinholes of diameter .01 inch.

(The answer appears in the next paragraph.) The full story of what happens, and under what conditions, is not widely known. It turns out, however, that the answer can be neatly encapsulated in a simple geometric picture: an equilateral triangle with its inscribed circle (FIGURE 3 in the next section). Each point in the triangle corresponds to some three-tank setup looking like FIGURE 2. Moving around in the triangle corresponds to changing relative drainhole sizes. Being inside the circle, or on it, or outside it, or near a triangle vertex or midpoint (and so on) all have physical significance. By the time you've finished this article, you'll be able to "read" FIGURE 3 to make qualitatively accurate predictions.

Surprisingly, three tanks share something basic with a damped mass-spring, the mass-spring actually supplying some handy vocabulary: the motion of the mass is either underdamped, critically damped, or overdamped. For quite different physical reasons (which we consider a little later), water levels in three tanks also vary in an either underdamped, critically damped, or overdamped fashion. Scenario (a) is an example of critically damped motion, where no oscillation takes place. In (b), the motion is underdamped, so the levels oscillate—this, despite the snail's pace of the level changes! It turns out that in (b), slowly increasing the size of any one of the three pinholes increases the damping (even though the overall water circulation speeds up). The system finally becomes critically damped and then overdamped, where oscillation dies out completely. These facts suggest some questions:

- If a drain size is given for each tank, is there a quick way of telling if the levels will oscillate?
- If the drainhole size is randomly chosen in each tank, what's the probability the levels will oscillate?
- Does the probability change if the tanks themselves have different diameters?
- Physically, why can oscillation occur in three tanks, but never two?
- Do the answers change if the tanks are arranged horizontally, so that the natural flow is two-way?

4. The particulars

The same kind of reasoning used on FIGURE 1 leads to these differential equations for FIGURE 2:

$$\begin{aligned} y_1' &= -k_1 y_1 + 0 y_2 + k_3 y_3, \\ y_2' &= +k_1 y_1 - k_2 y_2 + 0 y_3, \\ y_3' &= +0 y_1 + k_2 y_2 - k_3 y_3. \end{aligned} \tag{3}$$

Let $\mathbf{K}$ be the coefficient matrix of this system, and $\lambda_1, \lambda_2, \lambda_3$, the three roots of the minimal polynomial $\det(\mathbf{K} - \lambda\mathbf{I})$. This polynomial is cubic and factors; the minimal equation works out to be

$$\lambda\left(\lambda^2 + (k_1 + k_2 + k_3)\lambda + (k_2k_3 + k_1k_3 + k_1k_2)\right) = 0. \tag{4}$$

In this form, the minimal equation reveals some basic information about its roots:

- Exactly one root is zero, say λ_1.
- If the quadratic's roots are complex, then the common real part of both roots must be negative. This is because the sum of the roots (which are conjugate) is $-(k_1 + k_2 + k_3)$; this sum, being the negative of the quadratic's middle coefficient, is therefore negative.
- If the quadratic's roots are real, then $-(k_1 + k_2 + k_3) < 0$ once again shows that the sum of the roots is negative, so certainly at least one of them must be negative. This in turn means that both of them must be negative, since the product of the roots is the quadratic's constant term $k_2k_3 + k_1k_3 + k_1k_2$, which is positive.

If the λ_i are distinct and $\mathbf{v}_i$ are corresponding linearly independent eigenvectors, then the general solution is

$$\mathbf{y}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2 + c_3 e^{\lambda_3 t}\mathbf{v}_3 \tag{5}$$

where $\mathbf{y}(t)$ denotes the solution vector $(y_1(t), y_2(t), y_3(t))$. As $t \to \infty$, $\mathbf{y}(t)$ approaches $\mathbf{y}(\infty)$, which is just a scalar times the eigenvector $\mathbf{v}_1$. Actually, $\mathbf{y}(t)$ still approaches $\mathbf{v}_1$ in the repeated-roots case of $\lambda_2 = \lambda_3$; we leave this as an exercise for the reader. The eigenvector $\mathbf{v}_1$ has the same physical meaning as in two tanks: its components are in the same ratios as the final water heights. These heights occur when the flow into each tank matches the flow out, which (by transitivity) means all flow rates are the same, say c. The steady-state flow out of T_i is thus $c = k_i y_i(\infty)$, so $y_i(\infty) = c/k_i$. Writing

$$\mathbf{y}(\infty) = c\left(\frac{1}{k_1}, \frac{1}{k_2}, \frac{1}{k_3}\right),$$

we see that the steady state levels are in the ratios $\frac{1}{k_1} : \frac{1}{k_2} : \frac{1}{k_3}$.

How about the imaginary parts of the quadratic factor? It is precisely when these are nonzero that oscillation can occur—nonzero imaginary parts lead to sine and/or cosine terms in the solution. Therefore if the system is not at steady state, the levels oscillate if and only if the discriminant of the quadratic factor is negative.

After simplification, this discriminant is

$$k_1^2 + k_2^2 + k_3^2 - 2(k_2k_3 + k_1k_3 + k_1k_2). \tag{6}$$

If this is negative and if water levels are not in the ratio $\frac{1}{k_1} : \frac{1}{k_2} : \frac{1}{k_3}$, oscillation occurs. (This is a quick way of checking for oscillation, and answers the first question posed earlier.)

Notice that the discriminant is homogeneous (every term has degree two), so its zero set in (k_1, k_2, k_3)-space is also homogeneous. (Actually, homogeneous polynomials define homogeneous sets, and vice versa; see Theorem 2.6 in **[2]**, for example.) This set turns out to be a circular cone. Three-tank systems corresponding to points inside the cone oscillate; systems defined by points outside, don't.

This cone is very natural: besides being circular, it is tangent to each of the three coordinate planes in (k_1, k_2, k_3)-space, and its line of symmetry is the 1-space through $(1,1,1)$.

The tangency part is easy to establish—any 2-space is tangent to the cone if and only if it intersects the cone in one line (rather than in two or none). For instance, the cone intersects the (k_1, k_2)-coordinate plane in those points satisfying both the cone equation and the plane equation. Substituting the plane's equation $k_3 = 0$ into the equation of the cone gives $k_1^2 + k_2^2 - 2k_1k_2 = 0$—that is to say, $(k_1 - k_2)^2 = 0$. Therefore $k_1 = k_2$, meaning that the cone is tangent to the (k_1, k_2)-plane along the subspace through $(1,1,0)$. (The reasoning is the same for the other two coordinate planes.)

We can easily see the other properties if we use what is one of my favorite classroom exercises: find the equation of any circular cone with vertex at the origin, given its line of symmetry and the common angle the cone makes with that line. Here's the solution: Let $\mathbf{c}$ be a fixed vector defining the central line; $\mathbf{x}$, a typical point on the cone; and α, the common angle. The equation can now be written in vector form—just equate algebraic and geometric versions of the inner product:

$$\mathbf{c} \cdot \mathbf{x} = \|\mathbf{c}\| \cdot \|\mathbf{x}\| \cdot \cos(\alpha).$$

To get the familiar homogeneous equation of degree two, square both sides and simplify. (Done!)

In our case, we can determine $\cos(\alpha)$ by choosing, say, $\mathbf{x} = (1,1,0)$ and $\mathbf{c} = (1,1,1)$. This gives

$$(1,1,0) \cdot (1,1,1) = \sqrt{2} \cdot \sqrt{3} \cdot \cos(\alpha),$$

from which we find $\cos(\alpha) = \sqrt{\frac{2}{3}}$. With this $\mathbf{c}$ and $\cos(\alpha)$, our vector equation becomes the discriminant in (6) set equal to zero, so our cone is indeed circular with line of symmetry through $(1,1,1)$.

Actually, since all $k_i > 0$, we need look only at points in the first octant of (k_1, k_2, k_3)-space. Moreover, all points lying on any particular one-subspace determine systems whose behaviors are identical "up to a time factor" (which means that starting with the same initial conditions, their "movies" are identical except that they run at different speeds). If we choose the k_i's so that $k_1 + k_2 + k_3 = 1$, then the set of points inside the equilateral triangle with vertices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ becomes a natural parameter-space for the behavior of these systems. (This is just one instance of using a parameter space to "catalogue" the behavior of a system; there are many others. A really famous one is the Mandelbrot set in the complex plane, which indexes Julia sets; also, the part of the Mandelbrot set on the real axis parametrizes final-state values of the logistic (population) equation. See [**4**, Ch. 14.2] for an account.)

From the geometric information we've obtained about the discriminant cone, we see that it intersects this triangular region in a circle (the plane $k_1 + k_2 + k_3 = 1$ is perpendicular to the cone's line of symmetry), and from the cone's tangency, we see the circle is tangent to each side of the triangle at its midpoint. For any point randomly chosen in the triangular region, the system oscillates if and only if it is in the open disk bounded by the circle. The oscillation is fastest at the center. (This center may be regarded as a circle of zero radius.) The frequency is constant on any circle about this center, and decreases with increasing radius; it finally becomes zero on the inscribed circle, where the motion is critically damped. In the region outside this circle (shown dotted in FIGURE 3), the motion is overdamped.

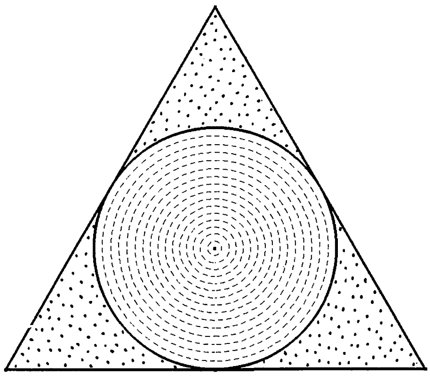


FIGURE 3
Our "Rosetta Stone": the basic parameter space

The system is always damped, even at the center; correspondingly, water levels always approach steady-state. Since any point in the triangle physically corresponds to relative drainhole areas of T_1, T_2, and T_3, one can now make an intuitive connection between drainhole sizes and what the water levels must do. For instance, the closer a point is to a vertex, the larger the corresponding drainhole is relative to the others. (Since we're taking $k_1 + k_2 + k_3 = 1$, these triples are actually barycentric coordinates on the triangle.) Physically, points near the vertices of the triangle depict one large and two small drains; reading off Figure 3, we see that the levels can never oscillate in this case. As another example, points close to the midpoint of a triangle side correspond to one small and two large drains. Here, levels may or may not oscillate, as is evident from the figure.

Notice that if we throw a dart at the triangle, the ratio of the disk's area to that of the triangle gives the probability that the dart has chosen an oscillating system, so we can now answer the second question we asked earlier. If for simplicity we take the circle's radius to be 1, the circle's area is then π, and the triangle's area is $3\sqrt{3}$. Thus *if we choose a k_i for each tank T_i by randomly selecting a point in parameter space, the probability of oscillation is $\pi/(3\sqrt{3}) \approx 60.46\%$.*

5. What makes water levels oscillate?

From a physical standpoint, why can oscillation occur in three tanks, but never in two? In a mass-spring setup, oscillation comes from an obvious source: momentum, which carries the mass beyond the equilibrium point. But in our tanks, only the water levels (and tank constants) were used in predicting the fate of the levels an instant later—momentum never entered the picture. So the question remains: *Where does the oscillation in three tanks come from?*

A short answer is: three tanks, as distinct from two, provide room for time-lag, or delay. We illustrate with a simple but extreme example, where a lot of delay is introduced: consider a large number n of tanks, all cyclically coupled (say counterclockwise). Let the first tank T_1 start with a unit volume of water, and let all others start empty. We can approximate what happens to the water in any tank T_i by a one-way random walk—in a fixed time element Δt, a small volume element of fixed size moves counterclockwise to the next tank with a certain probability (which is directly proportional to k_i); otherwise, the element remains in the same tank. If the three k_i are identical, then these probabilities are all the same, say p. For $m < n$, the distribution of the water after the m^{th} Δt is given by the binomial expansion of $[(1-p)+p]^m$. As time progresses, the initial "density spike" represented by the one full tank, smears out to a discrete version of a bell-curve, its peak slowly moving counterclockwise around the circle of tanks. The advancing tail rotates twice as fast as the peak; the other tail stays at T_1.

By the time the leading edge of water finally makes it once all the way around (which has taken quite a while—this is the time lag), the first tank has become nearly empty. But now this tank begins to pick up water from the wave that has just hit it, and its level begins to rise. This process continues: the evolving bell-curve continues to wind around, overlapping itself and smearing out more and more with time. Overlapping occurs for $m \geq n$; $[(1-p)+p]^m$ continues to give the amount of water in each tank as long as one writes the expansion in a similarly overlapping fashion, adding the i^{th} term to tank $T_{1+(i-1) \bmod (n)}$. As time goes on, the peak becomes flatter, the rising and falling of the crest in tank T_1 becomes less, and all levels approach steady-state. Because the moving peak of the bell-curve advances at a constant rate, the frequency of oscillation is the same in each tank.

Why, then, do the levels fail to oscillate in scenario (a)? The larger drain reduces the time lag which that tank contributes in (b). With a still larger drain, one could say that "water leaves almost as soon as it enters." The three tanks then approximate a two-tank system in which the large-drain tank has simply become a part of the connecting pipe and, as the solution (2) shows, a two-tank system never oscillates.

We also asked earlier about two-way or non-oriented flow, created by placing all the water tanks on the same horizontal plane (as on a table top) or replacing them by canisters of gas. This translates into a two-way random walk; now both tails move and the central peak remains stationary. This means there's no "rotating peak" to create both rise and fall in tanks as the peak passes through. Algebraically, the coefficient matrix of the set of differential equations is easily checked to be symmetric, and this leads to a nice physical application of an old friend from linear algebra: "The eigenvalues of a real symmetric matrix are real." In the solution, sines and cosines come only from imaginary parts; this theorem therefore tells that in any non-oriented coupling of tanks, the water levels never oscillate.

6. Looking through invariant eyeglasses

We assumed at the start that our tanks all have unit base. Therefore any k, being the area of the drainhole, is also the ratio of this area to the base area. For tanks of arbitrary base area, let's define k to actually *be* this ratio; this is advantageous because then k, all by itself, determines how a tank behaves in a system. For example, altering any tank in a system by multiplying both its base area and drain area by the same (positive) constant not only leaves k the same, it also leaves the water flow unchanged. Reason: for a given volume of water in any tank, multiplying base area by M results in multiplying water height by $\frac{1}{M}$; but since drain area is multiplied by M, the flow rate is multiplied by $\frac{1}{M} \cdot M = 1$. If in a system we replace any tank by one altered as above, the system would never detect a difference; so although water height changes, water volume doesn't. Using $V' = -kV$ in place of $y' = -ky$ then yields equations that are invariant in the sense that they're independent of the cylinders' bases. Rewriting equations (1) and (3) in terms of volumes V_i and ratios k_i means that everything we've said holds for tanks individually having arbitrary base area, with k_i meaning ratio and the words "level" and "height" replaced by volume; this answers another of the questions posed earlier.

This leads to further results that may surprise the unwary. Here are two examples:

- In scenario (a), the large drain has double the diameter and therefore four times the area of either small one. The discriminant (6) turns out to be zero, which is why the damping is critical. Keeping the large drain the same while expanding that tank's base leads immediately to oscillation. When the base area reaches four times its original, the behavior is exactly that of (b), up to a time factor. In FIGURE 3, this translates to starting near a vertex, and moving along its angle bisector to the circle's center. On further increasing the base, the point moves beyond the center toward the opposite side; the frequency decreases, approaching but never reaching zero.
- In scenario (b), suppose we keep the pinholes the same but reduce any one tank's base area. This translates in FIGURE 3 to moving the point corresponding to scenario (b); it approaches that tank's vertex k_i along the angle bisector. At some stage, oscillation must vanish.

As with two-tank systems, one can graphically explore three tank systems (both conservative and nonconservative) using the program PHASER in **[3]**; choose Equation "linear 3d" in type "3D Differential" to draw phase portraits.

7. Further questions

1. Show that the vector of steady-state levels in n conservative, cyclically-coupled tanks is a scalar times $\left(\frac{1}{k_1}, \frac{1}{k_2}, \dots, \frac{1}{k_n}\right)$, and that this is an eigenvector corresponding to $\lambda = 0$.
2. For two tanks (conservative or nonconservative), show that the limiting ratio of water levels as $t \to \infty$, defines one eigenvector; and the limiting ratio as $t \to -\infty$, defines the other eigenvector. In what sense do the two corresponding eigenvalues measure the levels' rate of change? (One can approach this problem analytically, working with the solution; here again, it's instructive to use PHASER to plot the phase portrait.)
3. What happens if the three tank system isn't conservative? It turns out that even the smallest leak to the environment makes oscillation to steady state impossible. "But how," one might wonder, "could three Titanic-sized tanks behave so differently as a result of, say, a single, micron-sized hole?" Actually, for a long time, they don't. Eventually, however, the oscillation fades away, leaving monotonic decrease. To get an idea of how (and why) this happens, let's introduce a small leak of size k into each tank. (This symmetry simplifies the algebra.)
   (a) Show that the minimal equation is just equation (4), with λ replaced by $\lambda + k$.
   (b) Conclude from this that k's size affects neither the presence nor the frequency ω of any sine or cosine term appearing in the solution.
   (c) Show that the water level in any tank has the form $y(t) = Ae^{-\alpha t} + Be^{-\beta t}\sin(\omega t + \phi)$, where $\beta > 0$.
   (d) Any level which oscillates to steady state has infinitely many extrema. But from (c), show that the condition $y'(0) = 0$ becomes

$$B\sqrt{\beta^2 + \omega^2}\sin(\omega t + \phi) = \alpha Ae^{(\beta - \alpha)t}.$$

   Since $(\beta - \alpha) > 0$, there are only finitely many solutions of this, hence only finitely many extrema.

For leaks of arbitrary size, the crux of the argument still comes down to showing $\beta > \alpha$, though the minimal polynomial is not as tractable. Also, part (b) doesn't hold in general, but that changes only minor details of the argument.

Do you think the levels could oscillate in one tank while varying monotonically in another?

4. Introductory applications of systems of differential equations sometimes include "mixing problems." Although it doesn't look as if any useful mixing is taking place in either two or three tank setups, knowing what happens to a single liquid allows one to directly solve apparently complicated mixing problems (as in making an alloy, for example). One need only determine the behavior of each liquid separately, as if the others weren't there. Then, as long as the contents of each tank are kept well stirred, one can add the individual results (using the superposition principle) to find the composition in each tank at any time. Reducing mixing problems to only one liquid works whether the system is conservative or nonconservative. As an application, recall textbook examples of two or three tanks of brine connected in series, with fluid flowing from one to the next. All flow rates are typically some common constant, with pure water entering the first tank and brine flowing to the environment from the last tank. From the superposition principle, one can look at the "liquids" being mixed as water and pure salt, and the water can just be ignored. That is, one can think of the pure salt as a flowing fluid, gravity-draining from one tank to the next. Show that the eigenvalues for salt flow are $-k_i$, and that each pure salt level has at most one extremum (a maximum). Since the brine heights stay constant, the concentrations can never oscillate: any moving peak created by an initially high amount of salt, never makes it back to its original location where it could otherwise increase the salt level and initiate oscillation.

REFERENCES

1. Paul Halmos, *I Want to be a Mathematician*, Springer-Verlag, New York, NY, 1985.
2. K. Kendig, *Elementary Algebraic Geometry*, Springer-Verlag, New York, NY, 1977.
3. H. Koçak, *Differential and Difference Equations through Computer Experiments*, Springer-Verlag, New York, NY, 1989.
4. H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals*, Springer-Verlag, New York, NY, 1992.

NOTES

Woven Rope Friezes

FRANK A. FARRIS
Santa Clara University
Santa Clara, CA 95053

NILS KRISTIAN ROSSING
SINTEF Telecom and Informatics
N – 7046 Trondheim
Norway

Introduction The rope mat shown in FIGURE 1 was woven by Nils Kristian Rossing following equations that produce frieze symmetry. To analyze the pattern, think of the curve the rope follows rather than the rope itself and recognize that, since rope knows more physics than mathematics, the rope does not quite follow the beautifully

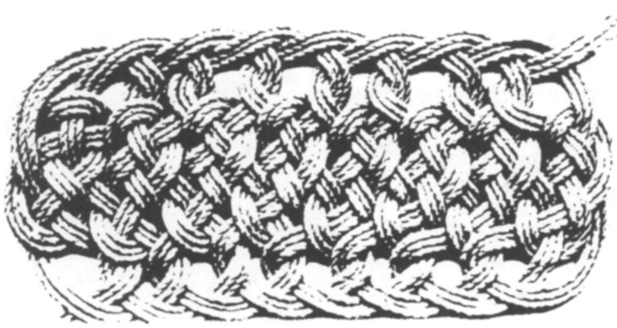


FIGURE 1
A woven rope mat with G_2 symmetry

symmetric curve shown. (See FIGURE 2.) This curve, meant to be interpreted as continuing indefinitely in both directions, has translational symmetry and rotational symmetry in the form of half-turns about the points indicated, leading one to identify its symmetry group as the *frieze group*, generally called G_2.

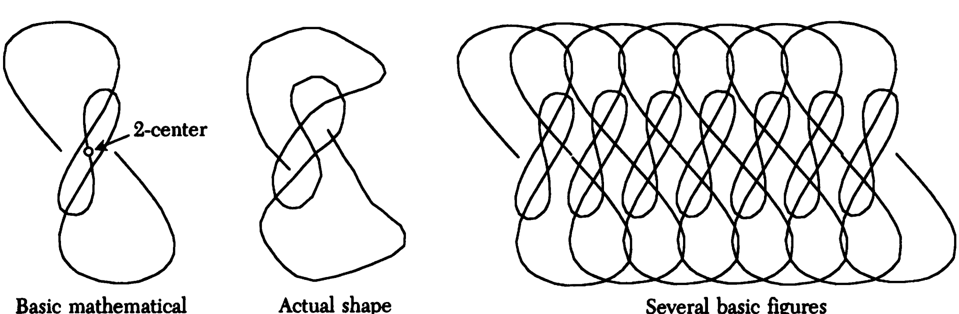


FIGURE 2

Here we present a complete set of recipes showing how to construct smooth curves with any desired frieze symmetry; we provide examples woven by Rossing for many of the pattern types, and invite readers to make others.

We review the concept of frieze symmetry, develop the formulas for parametric equations with given symmetries, and pose some open questions raised by our analysis. We also provide a Java applet to allow easy experimentation; this is now available on the World Wide Web along with high-resolution photographs of the ropes; please see `http://math.scu.edu/~ffarris/frieze.html` or `http://www.maa.org/pubs/mathmag.html`.

Frieze symmetry For us, a frieze (or frieze pattern) will be a set of points in the plane invariant under a translation, and hence an infinite cyclic group of translations. To relate this concept to the woven rope above, we idealize the path of the rope as a smooth curve, treating places where the rope passes over itself as self-intersections of the curve. We recognize that this treatment misses important features of the physical object pictured, such as whether the rope goes over or under itself at crossings; these considerations are expanded below and relegated to future discussion.

The set of all Euclidean motions, or isometries, that leave a frieze invariant is necessarily a group, called the symmetry group of the frieze. We say that two friezes have the same symmetry type if their groups are isomorphic. It is well-known (see, for instance, Cederberg **[1]**) that there are exactly seven isomorphism classes of symmetry groups containing a single infinite cyclic group of translations. Without going into detail, we assemble a few facts about these groups.

Suppose G is the invariant group of a frieze, which by definition contains an infinite cyclic subgroup of translations. Call the generator of the translation group τ, a translation of length K along a line l, so that K is the smallest distance one can translate the pattern and find that it falls into coincidence with itself. We find it convenient to think of l as the horizontal direction. The line l may not be uniquely defined if G has no elements other than multiples of τ, in which case we call the group G_1, or if the only other elements are reflections about lines perpendicular to l, when the group is called G_4.

Every other type of isometry possible in a frieze group leaves invariant a line parallel to the direction of translation, which we call the axis of the frieze, and name l. There are limited possibilities: The group G_2 contains a half-turn about a point of l, as well as all the half-turns generated by composing it with translations, but no other symmetries. We call the fixed point of a half-turn a 2-center of the frieze. G_3 contains no half-turns, but a reflection through l, which we will call a horizontal reflection. G_5 has half-turns as well as the horizontal reflection; since these together generate vertical reflections, this group has G_2, G_3, and G_4 as subgroups.

A glide reflection is the composition of a reflection and a translation. The group G_7 contains only a glide reflection along l in addition to its translation subgroup. G_6 has a glide reflection and half-turns. Examples of friezes with each type of symmetry are shown in FIGURE 3.

Frieze curves: general theory The rope frieze above was constructed by first finding a continuous curve in the plane, using a formula like $c(t) = (x(t), y(t))$, $-\infty < t < \infty$.

We suppose the curve is invariant under translation along the x-axis of a distance K, and that K is the smallest such distance. We assume that the curve is parametrized consistently with the translational invariance, so that

$$c(t+2L) = (x(t+2L), y(t+2L)) = (x(t)+K, y(t)).$$

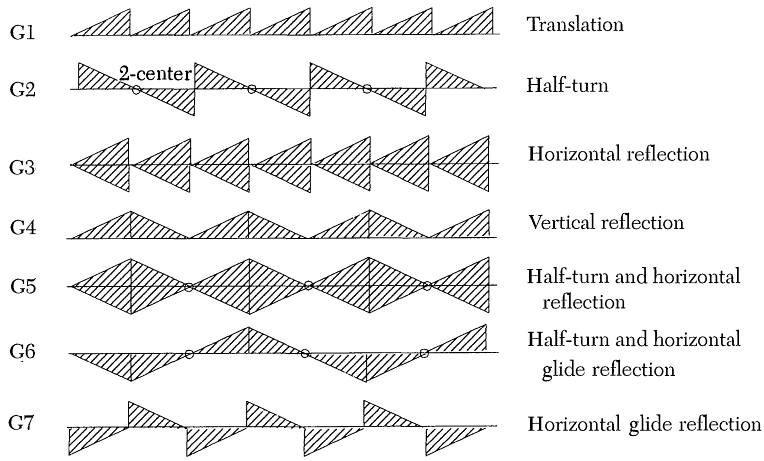


FIGURE 3

This can always be achieved by a reparametrization, using the equation above, recursively, to define a new parametrization outside a base interval $[-L, L]$. Note that we name the period $2L$ for a convenient match with the period 2π in most of our examples.

Before carrying out the Fourier analysis for such a curve, we make some observations about symmetry. There are several natural ways to decompose the functions $x(t)$ and $y(t)$. We define the even and odd parts of these two functions as usual, but also identify parts that we call the *glide-positive* and *glide-negative* parts:

$$x(t) = x_e(t) + x_o(t) = \frac{x(t) + x(-t)}{2} + \frac{x(t) - x(-t)}{2}$$

$$y(t) = y_e(t) + y_o(t) = \frac{y(t) + y(-t)}{2} + \frac{y(t) - y(-t)}{2}$$

$$x(t) = x_{g^+}(t) + x_{g^-}(t) = \frac{x(t) + x(t+L)}{2} + \frac{x(t) - x(t+L)}{2}$$

$$y(t) = y_{g^+}(t) + y_{g^-}(t) = \frac{y(t) + y(t+L)}{2} + \frac{y(t) - y(t+L)}{2}.$$

It turns out that every type of symmetry possible for curves of this type can be achieved by setting one or more parts of these decompositions to zero. To elaborate, we name some particular Euclidean motions of the plane and give their coordinate formulas.

TABLE 1. Some useful symmetries

Symbol	Description	Equation
τ	translation	$\tau(x, y) = (x + K, y)$
ρ	half-turn about origin	$\rho(x, y) = (-x, -y)$
σ_v	vertical mirror	$\sigma_v(x, y) = (-x, y)$
σ_h	horizontal mirror	$\sigma_h(x, y) = (x, -y)$
γ	glide reflection	$\gamma(x, y) = (x + \frac{K}{2}, -y)$

Note that our decompositions and choice of symmetries favor the origin. We comment on this later, but this is done without loss of generality.

The decomposition of $x(t)$ and $y(t)$ into even and odd parts makes identification of G_2 symmetry easy: if the even part of each function is zero, the curve $c(t)$ will be symmetric about the origin. Since it is also translation invariant, it will have G_2 symmetry.

G_4 symmetry is also easy to recognize in this set-up. If the even part of $x(t)$ and the odd part of $y(t)$ both vanish, G_4 symmetry will result. We elaborate this point with the equation

$$(x(-t), y(-t)) = (-x(t), y(t)) = \sigma_v(x(t), y(t)).$$

G_6 and G_7 are also simple cases. For G_7 simply require that x_{g^-} and y_{g^+} both vanish. Then

$$(x(t+L), y(t+L)) = \left(x(t) + \frac{K}{2}, -y(t)\right) = \gamma(x(t), y(t)).$$

G_6 symmetry results when the additional conditions for G_2 invariance are imposed.

A problem arises when we look for G_3 symmetry by this method. The desired equation would be:

$$(x(-t), y(-t)) = (x(t), -y(t)) = \sigma_h(x(t), y(t)),$$

and one might expect to satisfy this by choosing the odd part of $x(t)$ and the even part of $y(t)$ to be zero. However, a simple computation shows that $x(t)$ cannot be an even function; this is inconsistent with the translation equation. Intuitively, the curve cannot simultaneously move periodically to the right and appear equally on both sides of the horizontal axis.

We have solved this problem for woven ropes by using two strands to make a frieze. To create a pattern with a horizontal mirror, we introduce a second strand that mirrors the first. This can be done in several ways; to achieve G_3 symmetry, the original curve can have either G_1 or G_7 symmetry. Our classification of symmetries by the shape of the path the rope follows does not permit a distinction between these two types.

We arrive at G_5 realizing that a pair of curves will be necessary to create a frieze with a horizontal mirror. We would like the pattern created by both paths together to have G_5 symmetry, which means we need invariance under ρ and σ_v. Since $\rho = \sigma_v \sigma_h$ and $\sigma_v = \sigma_h \rho$, as long as we have one of these symmetries, we must have the other. Since there are two strands, we may construct either one to have one of these symmetries and the composite pattern will have G_5 symmetry. Thus G_5 patterns may be made from pairs of curves having G_2, G_4, or G_6 symmetry.

We summarize these results in Table 2.

It is natural to ask whether we have left something out. We have not considered every possible pair of conditions. For instance, what if $x_{g^-} \equiv 0$ and $y_{g^-} \equiv 0$? One can

TABLE 2. Symmetry recipes for frieze curves

Type	Conditions on components	min # strands
G_1	no components vanish	1
G_2	$x_e \equiv y_e \equiv 0$	1
G_3	G_1 or G_7 conditions in mirrored strands	2
G_4	$x_e \equiv 0$, $y_o \equiv 0$	1
G_5	G_2, G_4, or G_6 conditions in mirrored strands	2
G_6	$x_e \equiv y_e \equiv x_{g^-} \equiv y_{g^+} \equiv 0$	1
G_7	$x_{g^-} \equiv y_{g^+} \equiv 0$	1

easily check that such a function would have period L, rather than $2L$. The smallest translation of such a pattern would have length $\frac{K}{2}$. If one wanted to construct friezes of that period, presumably one would have started with that condition from the start. Also, some of the components cannot vanish. As noted above $x_o \equiv 0$ contradicts the translation equation. The same is true for x_{g^+}.

Frieze curves: Fourier analysis As a first step in the Fourier analysis, note that a curve satisfying the equation of translational symmetry can be thought of as having one component causing it to move to the right, and another, more interesting component causing decorative twists and turns. We sort out the interesting part by subtracting the known motion to the right (at a rate of K units of distance per $2L$ units of time), and define:

$$m(t) = (x(t), y(t)) - \left(\frac{Kt}{2L}, 0\right) = (h(t), v(t)),$$

where $h(t)$ and $v(t)$ indicate the horizontal and vertical components of the decorative portion of $c(t)$.

Now $m(t)$ is continuous and periodic of period $2L$, and so has a convergent Fourier series of the form

$$m(t) = \left(\sum_{n=0}^{\infty} a_n \sin\left(\frac{n\pi t}{L}\right) + b_n \cos\left(\frac{n\pi t}{L}\right), \sum_{m=0}^{\infty} c_m \sin\left(\frac{m\pi t}{L}\right) + d_m \cos\left(\frac{m\pi t}{L}\right)\right).$$

It becomes simple to translate the requirements above into conditions on the coefficients of the Fourier series. For example, $x(t)$ and $y(t)$ will both be odd functions if only sine terms appear; the recipe for constructing G_2 friezes is simply stated as $b_n = d_m = 0$.

The frieze in our first example has formula

$$c(t) = \left(\frac{13t}{2\pi} - \sin(2t) - 5\sin(3t) - 5\sin(4t), \sin(t) + 2\sin(2t)\right).$$

Here $L = \pi$ and $K = 13$. This was found by experimentation, once we knew that only sine terms could be included.

For symmetry type G_4, we see that one needs sines in the horizontal terms and cosines in the vertical terms. The rope frieze in FIGURE 4 was designed from this

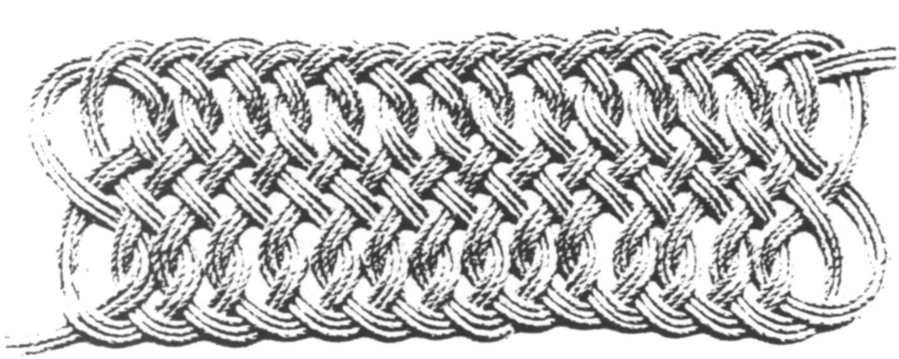


FIGURE 4
A woven G_4 frieze

recipe, using only one term of each Fourier series for simplicity. The equation is

$$c(t) = \left(\frac{0.8t}{2\pi} + \cos(2t), -0.8\sin(t)\right)$$

Types G_6 and G_7 are only slightly more tricky. One must use even values of n (the coefficients in the horizontal terms) and odd values of m (those in the vertical terms). Restricting both series to sine terms only will add the symmetry of the half-turn, yielding a frieze curve of type G_6.

One thing that may disturb the reader about G_6 patterns is that they have vertical mirrors without obeying the G_4 recipes. The vertical mirror in G_4 recipes was set up through the origin, through our privileged choice of coordinates. The vertical mirror in the G_6 patterns is about the line $x = \frac{L}{2}$.

FIGURE 5 shows a pattern woven from two ropes, with G_6 symmetry in each strand, producing G_5 symmetry overall. One thing to note at this point is that a formula for

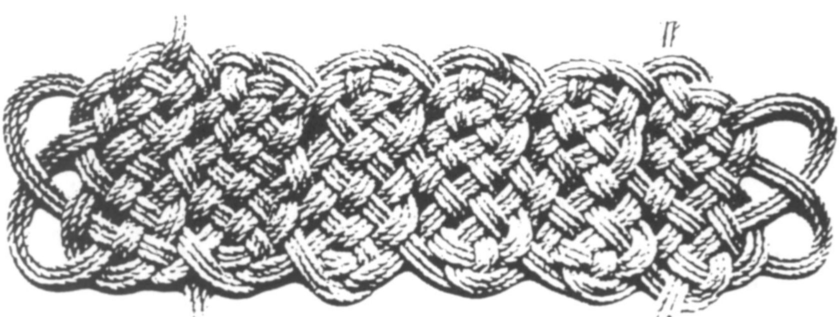


FIGURE 5
A G_6 pattern woven from two ropes

the pattern above could be detected by spectral analysis, using software like MATLAB. The equations used to produce a single strand are as follows:

$$x(t) = \frac{13t}{2\pi} - 1.5\sin(2t) - 3.5\sin(4t);$$

$$y(t) = -0.3\sin(t) - 3.2\sin(3t).$$

Table 3 gives a summary of all these recipes. We hope readers will experiment and make friezes of their own.

TABLE 3. Symmetry recipes for frieze curves

Type	Strands	Recipe and remarks
G_1	1	No additional requirements; use general parity
G_2	1	sine terms only
G_3	2	No additional requirements; use general parity
G_4	1	sines in $x(t)$, cosines in $y(t)$, general parity
G_5	2	G_2, G_4 or G_6 requirements in one strand, mirrored
G_6	1	n is even; m is odd; sines only
G_7	1	n is even; m is odd; both sines and cosines appear

Frieze ropes: further directions It may be unsatisfying to some readers that the instructions above only show how to construct curves with desired frieze symmetries. Surely the fact that these curves are realized physically as woven ropes is the most attractive thing about our illustrations. For empirically minded readers, we offer a way to experiment with these formulas in search of their own patterns to weave into ropes. For those interested in theory we expand a bit on the questions we have not yet answered.

To experiment making your own designs, you could use any computer algebra system, such as *Mathematica* or *DERIVE*, but we found this a bit cumbersome. Farris has written a JAVA applet allowing you to try out the formulas. An additional strand mirroring the first is available at a click of the mouse. To use this applet, direct your Java-enabled web browser to `http://math.edu/~ffarris/frieze.html` or `http://www.maa.org/pubs/mathmag.html`.

Probably the most interesting theoretical question involves the difference between curves and ropes. In all our examples we have simply assumed that the rope crosses itself in an alternate fashion, going over and then under every time. We call this the simple coding pattern. It is certainly possible to use more complex coding patterns, for instance, two over, one under, and so on. Such different coding patterns can be implemented very nicely in the simplest patterns, however for the friezes shown here it is difficult to attain handsome results.

On the other hand, almost every pattern we have encountered can be woven with the simple coding pattern. Experience teaches one to recognize patterns for which the simple coding pattern will not work. Perhaps a theoretical result is possible: the classification of curves for which the simple coding pattern is suitable.

There are further questions about the relationship between coding patterns and symmetry. In several examples, a half-turn has the effect of negating the coding pattern of the symmetry. A finer classification of frieze ropes might use the two-color frieze groups, outlined, for instance, in Grünbaum and Shephard **[3]**.

Finally, the techniques we use here are certainly adaptable for the construction of curves with wallpaper symmetry, at least if infinitely many strands are used. We would be delighted to see examples of wallpapers discovered using Fourier analysis and woven from rope.

How we came to write this note Our article had an interesting genesis. Rossing, a Norwegian radio engineer, responded to Farris's *Mathematics Magazine* article **[2]**, saying that he had already woven rotationally symmetric patterns like those described in the article. He had combined Fourier analysis with the old sailors' craft of weaving rosettes from rope, and found mathematical expressions for many existing rosettes. The tools Rossing developed had enabled him to produce new rosettes that he then wove out of rope **[4]**.

Farris proposed that they carry out the analysis to find the most general formula for patterns of each type of frieze symmetry, and see whether this would lead to the discovery of interesting designs suitable for rope work. Once the formulas were found, we experimented and found many examples for Rossing to weave with rope. As we have not actually met in person, we find this an Internet success story.

REFERENCES

1. Judith N. Cederberg, *A Course in Modern Geometries*, Springer-Verlag, New York, NY, 1989.
2. Frank A. Farris, Wheels on wheels on wheels—surprising symmetry, this MAGAZINE 69 (1996), 185–189.
3. Branko Grünbaum and G. C. Shephard, *Tilings and Patterns*, W. H. Freeman, New York, NY, 1986.
4. Nils Kr. Rossing, The old art of rope work and modern signal processing, *Knotting Matters*, *Newsletter of the International Guild of Knot Tyers*, December, 1996.

Compounding Evidence from Multiple DNA-Tests

SAM C. SAUNDERS
Scientific Consulting Services, Inc.
P.O. Box 458
Kirkland, WA 98083

N. CHRIS MEYER
DANE W. WU
Pacific Lutheran University
Tacoma, WA 98447

1. Introduction The sensational trial of a national football hero, O.J. Simpson, for the brutal murder of his former wife, Nicole, and her friend Ron Goldman, has become one of the most controversial trials in American history. It ranks with the 1925 trial of John Scopes, for teaching evolution in biology, as a premier example of a divisive and controversial ruling by the American juridical system (cf. [**5**]).

The American system of jurisprudence requires the jury to hold the defendant innocent until proved guilty beyond a reasonable doubt. If jurors are confused about the degree of incrimination of certain pieces of evidence, then doubt (reasonable or not) will ensue. The outcome of the O.J. trial hinged upon the jury's understanding of the probabilities of DNA blood-matching, a technique to establish guilt which was tested in this trial as never before. Some experts claim that DNA can only be used to exonerate persons accused of a crime, but never to establish guilt (see, e.g., [**7**]). The jury's acquittal was controversial. One columnist ([**2**]) ascribed it to their innumeracy. Others said the prosecution could not explain what they, themselves, did not understand. Justice requires that the probabilities of such events as matching fingerprints or matching DNA from blood samples be understandable. In this note we do not concern ourselves with the chemical problems of DNA matching. Neither do we consider the possibility of blood-sample contamination or of a law enforcement conspiracy to fake evidence. (These complications are discussed in [**4**].) Here we treat only some of the subtleties of conditional probabilities related to DNA matching.

2. Conditional probabilities and guilt Let M be the event of a DNA match between the defendant's blood and blood found at the crime scene, let I be the event the defendant is innocent, and let I' be the complement of I—the event the defendant is guilty. There are reliable measures for $P(M \mid I)$, the conditional probability of M given I. If the defendant is innocent, the blood must be someone else's; hence $P(M \mid I)$ is the probability of a match between blood samples of two different individuals.

Human DNA signatures are so distinctive that the probability of a random match is very low. Depending upon the amount of DNA obtained, it is often found that $P(M \mid I)$ is between 10^{-8} and 10^{-10}. But it is $P(I \mid M)$, the likelihood of innocence given the evidence, that should be considered by the jurors, not $P(M \mid I)$. These two conditional probabilities are related by Bayes' little theorem:

$$P(I \mid M) = \frac{P(I)}{P(M)} P(M \mid I). \qquad (1)$$

We see from (1) that $P(I \mid M) \neq P(M \mid I)$ unless $P(I) = P(M)$. The erroneous assumption that $P(I \mid M) = P(M \mid I)$ is sometimes called the *prosecutors' fallacy*. (For a discussion of the confusion this fallacy causes, see **[3]**.)

If in formula (1) we substitute for $P(M)$ from the theorem of total probability, we obtain Bayes' big theorem:

$$P(I \mid M) = \frac{P(I)P(M \mid I)}{P(I)P(M \mid I) + P(I')P(M \mid I')}.$$

It follows, since $(1 + x^{-1})^{-1} < x$ for $x > 0$, that

$$P(I \mid M) < \frac{P(I)P(M \mid I)}{P(I')P(M \mid I')} \approx \frac{P(I)P(M \mid I)}{P(I')}. \tag{2}$$

The approximate equality follows because $P(M \mid I')$, the probability of a match given that the defendant is guilty, is virtually unity. Here $P(I)$ and $P(I')$ are the *a priori* probabilities of innocence and guilt, before the evidence has been introduced. Thus we see from (2) that the prosecutor's fallacy is not so much a fallacy as ignorance of a factor of proportionality (the prior odds, $P(I)/P(I')$) in an upper bound on the true probability of innocence given a DNA match.

3. The compounding evidence In many cases, such as the O.J. trial, more than one DNA matching is involved. Our calculations thus far only deal with one piece of evidence, M. The jury needs to consider the combined effect of all the evidence (cf. **[6]**). Suppose events $M_1, \ldots, M_k$ are all pieces of evidence introduced against the defendant. We want to calculate $P(I \mid \bigcap_{i=1}^{k} M_i)$, the probability of innocence given all the evidence. The theorem below gives an upper bound for this probability. First, we introduce some needed definitions.

For a given event I, we denote $P_I(\cdot) = P(\cdot \mid I)$. Two events M_1 and M_2 are said to be *conditionally independent* with respect to I if $P_I(M_2 \mid M_1) = P_I(M_2)$. Note that the events M_1 and M_2 can be interchanged in this definition and that the conditional independence becomes mutual independence if I is a sure event.

The ratio $P(M_2 \mid M_1)/P(M_2)$ gives a measure of how strongly associated M_2 and M_1 are. If the ratio is less than 1 then $P(M_2 \mid M_1) < P(M_2)$, which means that M_2 is less likely to occur given M_1. Likewise, M_2 is more likely to occur given M_1 if the ratio is greater than 1. The ratio is 1 if and only if M_1 and M_2 are mutually independent.

With this in mind, we say that two events M_1 and M_2 are *more strongly associated conditionally* given I' than given I if

$$\frac{P_{I'}(M_2 \mid M_1)}{P_{I'}(M_2)} \geq \frac{P_I(M_2 \mid M_1)}{P_I(M_2)};$$

this is equivalent with

$$\frac{P_{I'}(M_1 M_2)}{P_{I'}(M_1)P_{I'}(M_2)} \geq \frac{P_I(M_1 M_2)}{P_I(M_1)P_I(M_2)}. \tag{3}$$

M_1 and M_2 are said to be *equally associated conditionally* under I and I' if equality holds in (3) (see **[1]**). We leave to the reader the proof of the following fact: If M_1 and M_2 are conditionally independent with respect to both events I and I', then they are equally associated conditionally under I and I'.

The following simple example illustrates equal association. Consider throwing two fair dice, say (X, Y), with events $M_1 = \{(x, y): y = 3, 4, \text{ or } 5\}$, $M_2 = \{(x, y): x = 1 \text{ or } 2\}$, and $I = \{(x, y): x + y = 7\}$. In this case M_1 and M_2 are equally associated conditionally under I' and I, since

$$\frac{P_I(M_1M_2)}{P_I(M_1)P_I(M_2)} = \frac{P_{I'}(M_1M_2)}{P_{I'}(M_1)P_{I'}(M_2)} = 1.$$

If we modify event M_2 as $M_2 = \{(x, y): x = 1 \text{ or } 2\} \cup \{(3,3)\}$, then M_1 and M_2 are more strongly associated conditionally given I' than given I, since, in this case,

$$\frac{P_I(M_1M_2)}{P_I(M_1)P_I(M_2)} = 1 \quad \text{and} \quad \frac{P_{I'}(M_1M_2)}{P_{I'}(M_1)P_{I'}(M_2)} = \frac{12}{11}.$$

In general, events $M_1, \dots, M_k$ are said to be *more strongly associated conditionally* given I' than given I if

$$\frac{P_{I'}\left(\bigcap_{i=1}^{k} M_i\right)}{\prod_{i=1}^{k} P_{I'}(M_i)} \geq \frac{P_I\left(\bigcap_{i=1}^{k} M_i\right)}{\prod_{i=1}^{k} P_I(M_i)}. \tag{4}$$

We have equal conditional association when the quality in (4) holds. (This definition is related to the concept of association; see **[1]** for more details.)

THEOREM. *If events $M_1, \dots, M_k$ are more strongly associated conditionally given I' than given I, then*

$$P\left(I \mid \bigcap_{i=1}^{k} M_i\right) \leq \frac{P(I)}{P(I')} \prod_{i=1}^{k} \frac{P(M_i \mid I)}{P(M_i \mid I')}. \tag{5}$$

Proof. From Bayes' big theorem we obtain

$$P\left(I \mid \bigcap_{i=1}^{k} M_i\right) = \frac{P(I)P\left(\bigcap_{i=1}^{k} M_i \mid I\right)}{P(I)P\left(\bigcap_{i=1}^{k} M_i \mid I\right) + P(I')P\left(\bigcap_{i=1}^{k} M_i \mid I'\right)}$$

$$\leq \frac{P(I)P\left(\bigcap_{i=1}^{k} M_i \mid I\right)}{P(I')P\left(\bigcap_{i=1}^{k} M_i \mid I'\right)}. \tag{6}$$

From the fact that $M_1, \ldots, M_k$ are more strongly associated conditionally given I' than given I, we have

$$\frac{P\left(\bigcap_{i=1}^{k} M_i \mid I\right)}{P\left(\bigcap_{i=1}^{k} M_i \mid I'\right)} \leq \frac{\prod_{i=1}^{k} P(M_i \mid I)}{\prod_{i=1}^{k} P(M_i \mid I')}. \tag{7}$$

Thus (5) follows by substituting (7) into (6).

4. An application of the theorem We are concerned in the O.J. trial with evaluating the probability of guilt given the totality of DNA evidence. Let M_1 be the event that a blood drop found near the victims' bodies is consistent with the defendant's blood. Cellmark Diagnostics, the DNA laboratory of record in the trial, said that only one person in 170 million could be expected to match the genetic markers identified in the blood drop. Thus $P(M_1 \mid I) = (1.7 \times 10^8)^{-1} = 5.88 \times 10^{-9}$. Let M_2 be the event that blood on a sock belonging to the defendant and found in his bedroom is consistent with Nicole's blood. Cellmark said the probability was one in 6.8 million that another person would match the genetic markers they found in the victim's blood (see **[3]**). Thus $P(M_2 \mid I) = (6.8 \times 10^9)^{-1} = 1.47 \times 10^{-10}$. Moreover, as mentioned above, $P(M_1 \mid I') \approx 1$ and $P(M_2 \mid I') \approx 1$.

In order to apply the theorem, one must argue that M_1 and M_2 are more strongly associated conditionally given guilt than given innocence. If I is true, then M_1 and M_2 are independent. But assuming guilt, one can imagine many scenarios in which the occurrence of one would increase the probability of the other. The two blood-match events would reasonably be judged to be more strongly associated conditionally given I' than given I. Substituting these numbers into (5) gives

$$P(I \mid M_1 M_2) \leq \frac{P(I)}{P(I')} \times P(M_1 \mid I) \times P(M_2 \mid I) \approx \frac{P(I)}{P(I')} \times 8.65 \times 10^{-19}.$$

Now we give an upper bound for the ratio $P(I)/P(I')$. If we go to the extreme of saying that O.J. Simpson was no more likely to be the killer than anyone else in the world, then $P(I') \approx 10^{-10}$ (realistically, $P(I')$ is much larger than this), and, of course, $P(I) \approx 1$. Hence $P(I)/P(I') \approx 10^{10}$ and $P(I \mid M_1 M_2) \leq 8.65 \times 10^{-9}$. The conditional probability of innocence given both DNA matches is so small as to place the defendant's guilt beyond any reasonable doubt.

Many other events could be used to condition the events in the trial of the defendant. For example, the glove found at the defendant's house, matching the one found at the crime scene, revealed genetic markers not only matching the victims' DNA but that of the defendant as well; label this match M_3. Let M_4 be the match of the defendant's blood to drops leading from the murder scene to the gate. According to the California state DNA analyst, this could have been left by one person in 240,000, including O.J. Let M_5 be the match of the DNA extracted from a drop of blood smeared in the white Bronco, which was consistent with that of Ron Goldman. Earlier we claimed that M_1 and M_2 were independent given I. Now $M_1, \ldots, M_5$ are not independent given I, because some involve the same persons. However, we need not assert such independence in order to apply the theorem; we need only assert that $M_1, \ldots, M_5$ are more strongly associated conditionally given I' than given I.

5. An instructive analogy Comments of some jurors after the trial revealed the opinion that DNA blood-matching is no more certain than fingerprint matching. A simple analogy can illustrate the reliability of DNA evidence. Every human carries DNA in each blood cell containing information which is analogous to the information from one permutation of a deck of cards. There are $52! \approx 8.0658 \times 10^{67}$ possible card permutations. Suppose that, from a drop of blood at the crime scene, there are only a few DNA sites (corresponding to card positions in the deck) from which information can be extracted.

Sometimes, moreover, depending upon the amount of DNA or the degree of contamination, not all the information from a site can be determined. The analogous situation for cards would be that only the number, the suit, or the color (say B or R) of the card can be determined. How incriminating can this be? Assume that partial information is available at only seven card positions, as follows:

$$(\spadesuit, \diamondsuit, 2-\heartsuit, 3-\clubsuit, 10-\heartsuit, 4-\diamondsuit, B).$$

Correspondingly, suppose that at the same DNA site locations in the defendant's blood sample we find

$$(2-\spadesuit, 10-\diamondsuit, 2-\heartsuit, 3-\clubsuit, 10-\heartsuit, 4-\diamondsuit, K-\clubsuit).$$

This corresponds to a DNA match comparable to the suspect's blood being consistent with that found at the crime scene. Assuming that all permutations are equally likely, the probability of obtaining this match is

$$\frac{13}{52} \times \frac{13}{51} \times \frac{1}{50} \times \frac{1}{49} \times \frac{1}{48} \times \frac{1}{47} \times \frac{24}{46} \approx 6.0154 \times 10^{-9}.$$

Only about six persons in a billion would yield a match.

Acknowledgment. We thank the referees for their constructive suggestions.

REFERENCES

1. R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*, To Begin With, Silver Spring, MD, 1981, 31.
2. Dave Barry, Happy Newt Year, column in the December 31, 1995, Knight-Ridder newspapers.
3. Robin Clark, Will DNA turn the tide?, *Philadelphia Inquirer*, May 21, 1995.
4. R. Gomulkiewicz and N. A. Slade, Legal standards and the significance of DNA evidence, *Human Biology* 69, 1997.
5. Edward J. Larsen, *Summer for the Gods: The Scopes Trial and America's Continuing Debate over Science and Religion*, Basic Books, New York, NY, 1997.
6. S. C. Saunders, An application of conditional probability, or, O.J. + DNA = QED, Mathematics Notes from Washington State University, Vol. 38, No. 2, 1995.
7. David Pringle, Who's the DNA fingerprinting pointing at?, *New Scientist*, January 29, 1994.

Math Bite: Further Generalizations of a Curiosity that Feynman Remembered All His Life

Richard Feynman remembered that, when a boy, he was told by Morrie Jacobs [**3**, page 47] that $\cos 20^\circ \times \cos 40^\circ \times \cos 80^\circ = 1/8$. Beyer, Louck, and Zeilberger pointed out [**1**] that "Morrie Jacobs' identity is the special case $k = 3$, $a = 20^\circ$ of the following identity that follows by induction on k using $\sin 2b = 2\sin b \cos b$, with $b = 2^{k-1}a$":

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

Morrie Jacobs' identity can be generalized in various other ways, several of which are given in the splendid old textbook [**2**] by Durell & Robson. In particular, Jacobs' identity is the special case $m = 4$ of the first of the 4 identities in positive integers m:

$$\prod_{r=1}^{m} \cos\left(\frac{r\pi}{2m+1}\right) = 2^{-m}, \quad \prod_{r=1}^{m} \cos\left(\frac{(2r-1)\pi}{4m}\right) = 2^{1-m}\sqrt{m},$$

[**2**, page 225, Ex. 24], which transform simply to give:

$$\prod_{r=1}^{m} \sin\left(\frac{r\pi}{2m+1}\right) = 2^{-m}, \quad \prod_{r=1}^{m} \sin\left(\frac{(2r-1)\pi}{4m}\right) = 2^{1-m}\sqrt{m},$$

[**2**, page 225, Ex. 25].

Dividing the products of sines by the corresponding products of cosines, we get trivial identities for tangents: $\prod_{r=1}^{m} \tan\left(\frac{r\pi}{2m+1}\right) = 1$, and [**2**, page 225, Ex.30] $\prod_{r=1}^{m} \tan\left(\frac{(2r-1)\pi}{4m}\right) = 1$.

There are various other generalizations, including the following identity [**2**, page 225, Ex. 29] in positive integers n:

$$\prod_{r=1}^{n} \cos\left(\frac{(2r-1)\pi}{2n}\right) = 2^{1-n}\cos\left(\frac{n\pi}{2}\right).$$

For odd n the product equals 0, but for even $n = 2m$, this becomes the identity

$$\prod_{r=1}^{2m} \cos\left(\frac{(2r-1)\pi}{4m}\right) = \frac{2}{(-4)^m}.$$

Also [**2**, page 226, Ex. 31]:

$$\prod_{r=1}^{2m+1} \cos\left(\frac{(4r-1)\pi}{8m+4}\right) = \frac{-1}{\sqrt{2}(-4)^m}, \quad \prod_{r=1}^{2m} \cos\left(\frac{(4r-1)\pi}{8m}\right) = \frac{\sqrt{2}}{(-4)^m}.$$

REFERENCES

1. W.A. Beyer, L.D. Louck, and D. Zeilberger, Math Bite: A generalization of a curiosity that Feynman remembered all his life, this MAGAZINE 69 (1996), 43–44.
2. C.V. Durell and A. Robson, *Advanced Trigonometry*, G. Bell & Sons Ltd; London, 1930.
3. James Gleick, *Genius: The Life and Science of Richard Feynman*, Pantheon Books, New York, NY, 1992.

—GARRY J. TEE
UNIVERSITY OF AUCKLAND
AUCKLAND
NEW ZEALAND

Finding Exact Values For Infinite Sums

COSTAS J. EFTHIMIOU
Department of Physics
Tel Aviv University
Tel Aviv 69978
Israel
and
Newman Laboratory
Cornell University
Ithaca, NY 14850

Introduction In the February 1995 issue of *Math Horizons* [**1**], I. Fisher posed the following problem:

From the well-known results

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \quad \textit{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^2+n} = 1,$$

it follows that

$$1 < \sum_{n=1}^{\infty} \frac{1}{n^2+n/2} < \frac{\pi^2}{6}.$$

Find the exact value of the convergent sum.

This paper offers a solution method that allows one to find exact values for a large class of convergent series of rational terms.

We first illustrate the method for a special case. Then we describe the general result, pointing out further generalizations of the method.

A special case Consider the series

$$S(a,b) = \sum_{n=1}^{\infty} \frac{1}{(n+a)(n+b)}, \tag{1}$$

where $a \neq b$ and neither a nor b is a negative integer. Sums of this form arise often in problems in Quantum Field Theory (see, e.g., [**2**, p. 89ff]).

Decomposing each term of (1) in partial fractions gives

$$S(a,b) = \frac{1}{a-b} \sum_{n=1}^{\infty} \left(\frac{1}{n+b} - \frac{1}{n+a} \right). \tag{2}$$

We will use the identity

$$\frac{1}{A} = \int_0^{\infty} e^{-Ax}\, dx, \tag{3}$$

which holds if $A > 0$. Therefore, for $a, b > -1$ we have

$$S(a,b) = \frac{1}{a-b} \lim_{N\to\infty} \sum_{n=1}^{N} \int_0^\infty e^{-nx}(e^{-bx} - e^{-ax})\, dx$$
$$= \lim_{N\to\infty} \int_0^\infty \frac{e^{-bx} - e^{-ax}}{a-b} \frac{e^{-x}(1-e^{-Nx})}{1-e^{-x}}\, dx$$
$$= \frac{1}{a-b} \int_0^\infty \frac{e^{-bx} - e^{-ax}}{a-b} \frac{e^{-x}}{1-e^{-x}}\, dx. \qquad (4)$$

In deriving the last result, we used the monotone convergence theorem (see, e.g., **[3]**, p. 318]). It applies because the integrand in the second line of equation 4 consists of a nondecreasing sequence of nonnegative functions; hence we can swap the limit and integral operations.

Making the change of variable $t = e^{-x}$ in the integral of equation (4) gives a more symmetric result:

$$S(a,b) = \frac{1}{a-b} \int_0^1 \frac{t^b - t^a}{1-t}\, dt. \qquad (5)$$

Some comments are in order:

1. Although the integral

$$\int_0^1 \frac{t^y}{1-t}\, dt$$

diverges, the integral of equation (5) converges for $a \neq b$ and $a,b > -1$.

2. The problem that appeared in *Math Horizons* corresponds to $a = 1/2$, $b = 0$. In this case, we have

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1/2)} = 2\int_0^1 \frac{dt}{1+\sqrt{t}}.$$

Following the change of variables $t = u^2$, the integral is easy to calculate:

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1/2)} = 4\int_0^1 \left(1 - \frac{1}{1+u}\right) du = 4(1 - \ln 2) \simeq 1.227.$$

3. If a and b differ by an integer k, so $a = b + k$, then the sum (1) "telescopes" and it can be easily calculated from (2):

$$S(a, a-k) = \frac{1}{k} \sum_{j=1}^{k} \frac{1}{j+a}.$$

This can be used as a check on formula (5). Indeed,

$$S(a, a-k) = \frac{1}{k} \int_0^1 \frac{t^k - 1}{t-1} t^a\, dt = \frac{1}{k} \int_0^1 \sum_{i=0}^{k-1} t^{i+a}\, dt = \frac{1}{k} \sum_{i=0}^{k-1} \frac{t^{i+a+1}}{i+a+1}\Bigg|_0^1$$
$$= \frac{1}{k} \sum_{j=1}^{k} \frac{1}{j+a},$$

which agrees with the preceding result.

We can express the result (5) in another equivalent form, using a well-known representation of the *digamma function* $\psi(z)$:

$$\psi(z) \equiv \frac{d}{dz}\ln\Gamma(z) = -\gamma - \int_0^1 \frac{t^{z-1}-1}{1-t}\,dt, \tag{6}$$

where γ is the Euler–Mascheroni constant and $\Gamma(z)$ is the gamma function. (See, e.g., [**4**, p. 258] for information on the digamma and polygamma functions.) Noting the similarity between the integrals in (5) and (6), we find that

$$S(a,b) = \frac{\psi(b+1)-\psi(a+1)}{b-a}. \tag{7}$$

The digamma function satisfies many useful identities. For example,

$$\psi(1+z) = \psi(z) + \frac{1}{z}. \tag{8}$$

The exact value of $\psi(z)$ is known for several values of z:

$$\psi(1) = -\gamma, \quad \psi(1/2) = -\gamma - 2\ln 2. \tag{9}$$

Equations (8) and (9) can be used to evaluate $S(a,b)$ exactly for many values of a and b. For example,

$$\begin{aligned} S(0,1/2) &= 2[\psi(3/2)-\psi(1)] = 2[\psi(1/2)+2-\psi(1)] \\ &= 2[-\gamma - 2\ln 2 + 2 + \gamma] = 4(1-\ln 2), \end{aligned}$$

which agrees with our previous result.

When $a = b$ in (1), the sum can still be calculated. We consider two approaches. The first is to repeat the calculations presented above, but modifying the basic equation (3):

$$\frac{1}{A^2} = \int_0^\infty x e^{-Ax}\,dx, \quad \text{for } A > 0.$$

Following the same reasoning as above, we find

$$\sum_{n=1}^{\infty} \frac{1}{(n+a)^2} = -\int_0^1 \frac{t^a \ln t}{1-t}\,dt. \tag{10}$$

Alternatively, we can obtain the same result by taking the limit in (5) as $b \to a$:

$$\sum_{n=1}^{\infty} \frac{1}{(n+a)^2} = \lim_{b\to a} \int_0^1 \frac{t^b - t^a}{b-a}\,\frac{-1}{1-t}\,dt. \tag{11}$$

Without loss of generality, we can assume that $-1 < a < b$. We notice that for $0 \le t \le 1/2$,

$$\left|\frac{t^b - t^a}{1-t}\right| \le 2t^b,$$

while for $1/2 < t \leq 1$ the integrand of (5) is bounded; let M be its supremum in this subdomain. The function

$$g(t) = \begin{cases} \dfrac{2t^b}{b-a} & \text{if } 0 \leq t \leq 1/2, \\ M & \text{if } 1/2 < t \leq 1 \end{cases}$$

is integrable in [0, 1]. Therefore, the dominated convergence theorem [**3**, p. 167, 321] can be used to interchange the operations of the integral and the limit in (11):

$$\sum_{n=1}^{\infty} \frac{1}{(n+a)^2} = \int_0^1 \lim_{b \to a} \frac{t^b - t^a}{b-a} \frac{-1}{1-t}\, dt = -\int_0^1 \frac{t^a \ln t}{1-t}\, dt.$$

From equation (7) we find

$$\sum_{n=1}^{\infty} \frac{1}{(n+a)^2} = \left. \frac{d\psi(z)}{dz} \right|_{z=a+1}.$$

The functions

$$\psi^{(n)}(z) \equiv \frac{d^{n+1}}{dz^{n+1}} \ln \Gamma(z) = -\int_0^1 \frac{t^{z-1}(\ln t)^n}{1-t}\, dt, \quad n = 1, 2, \ldots \qquad (12)$$

are known as *polygamma functions*. Several identities for the polygamma functions are known [**4**, p. 258ff]. For example,

$$\psi^{(n)}(1) = (-1)^{n+1} n! \zeta(n+1);$$

$$\psi^{(n)}(1/2) = (-1)^{n+1} n! (2^{n+1} - 1) \zeta(n+1);$$

$$\psi^{(n)}(z+1) = \psi^{(n)}(z) + \frac{(-1)^n n!}{z^{n+1}}.$$

Here $\zeta(z)$ is the Riemann zeta function, defined by

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}, \quad \text{if } \mathrm{Re}(z) > 1.$$

As an application of (10), we obtain the well-known result

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = -\int_0^1 \frac{\ln t}{1-t}\, dt = \frac{\pi^2}{6} \simeq 1.645.$$

Equation (10) also implies the less well-known result

$$\sum_{n=1}^{\infty} \frac{1}{(n+1/2)^2} = -\int_0^1 \frac{\sqrt{t} \ln t}{1-t}\, dt = 3\zeta(2) - 4 = \frac{\pi^2}{2} - 4 = 0.935\,.$$

The general case We can now establish a more general result. Let

$$S = \sum_{n=1}^{\infty} \frac{Q_{N-2}(n)}{P_N(n)},$$

where $Q_{N-2}(n)$, and $P_N(n)$ are two polynomials in n, of degree $N-2$ and N, respectively. We shall assume that $P_N(n)$ is expressible in the form

$$P_N(n) = (n+a_1)^{m_1}(n+a_2)^{m_2} \cdots (n+a_k)^{m_k},$$

where the a_i, $i = 1, 2, \ldots, k$ are distinct real numbers, none of them a negative integer. (This ensures the convergence of S.) Then for any polynomial

$$Q_{N-2}(n) = c_0 + \cdots + c_{N-3} n^{N-3} + c_{N-2} n^{N-2}$$

the sum S is written in terms of partial fractions:

$$S(a_1, \ldots a_k; c_0, \ldots, c_{N-2}) = \sum_{n=1}^{\infty} \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{A_{ij}}{(n+a_i)^j},$$

where the constants A_{ij} are uniquely determined by the partial fraction decomposition of each summand. In particular, we notice that, since there is no term of degree $N-1$ in $Q_{N-2}(n)$,

$$\sum_{i=1}^{k} A_{i1} = 0. \tag{13}$$

Using the identity

$$\frac{1}{A^L} = \frac{1}{(L-1)!} \int_0^{\infty} x^{L-1} e^{-Ax}\, dx, \tag{14}$$

we write the series in integral form (which is valid only if $a_i > -1$, for all i):

$$S(a_1, \ldots a_k; c_0, \ldots, c_{N-2}) = \sum_{n=1}^{\infty} \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{A_{ij}}{(j-1)!} \int_0^{\infty} x^{j-1} e^{-(n+a_i)x}\, dx. \tag{15}$$

Working in a similar fashion as in the derivation of equation (4), we find

$$S(a_1, \ldots a_k; c_0, \ldots, c_{N-2}) = \sum_{i=1}^{k} \sum_{j=2}^{m_i} \frac{A_{ij}}{(j-1)!} \int_0^{\infty} x^{j-1} \frac{e^{-(a_i+1)x}}{1-e^{-x}}\, dx + \sum_{i=1}^{k} A_{i1} \int_0^{\infty} \frac{e^{-(a_i+1)x} - 1}{1-e^{-x}}\, dx.$$

(We took extra care for the $j = 1$ term, using the condition (13), in order to guarantee the convergence of the corresponding integral.) We can also express our result in terms of the polygamma functions (12):

$$S(a_1, \ldots a_k; c_0, \ldots, c_{N-2}) = \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{(-1)^j}{(j-1)!} A_{ij} \psi^{(j-1)}(a_i+1), \tag{16}$$

where we define $\psi^{(0)}(z) \equiv \psi(z)$.

We apply our method to some straightforward examples.

EXAMPLE 1. Find

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

where a is not a positive integer, and $a>-1$. Using (16), we find

$$S(a,-a)=\frac{\psi(a+1)-\psi(-a+1)}{2a},\quad a\neq 0.$$

This result can be further simplified if we use the functional relation

$$\psi(-z+1)=\psi(z)+\pi\cot(\pi z),$$

together with (8). The result is

$$S(a,-a)=\frac{1}{2a}\left[\frac{1}{a}-\pi\cot(\pi a)\right].$$

EXAMPLE 2.

$$S(\underbrace{a,\ldots,a}_{N})=\sum_{n=1}^{\infty}\frac{1}{(n+a)^N}=\frac{(-1)^N}{(N-1)!}\psi^{(N-1)}(a+1),$$

where $N\geq 2$ and $a>-1$.

EXAMPLE 3. Find

$$S(0,0,3/2)=\sum_{n=1}^{\infty}\frac{1}{n^2(n+1/2)}.$$

Formula (16) gives

$$S(0,0,3/2)=4\psi(1)+2\psi^{(1)}(1)-4\psi(3/2)=\frac{\pi^2}{3}-8(1-\ln 2)\approx 0.835.$$

EXAMPLE 4.

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

Using formula (16) we find

$$S(1,1,1/2)=\left[4\psi(2)-2\psi^{(1)}(2)-4\psi(3/2)\right]=2(4\ln 2-1)-\frac{\pi^2}{3}\approx 0.255.$$

The reader is invited to write down values for other infinite sums of the same form.

The Laplace transform The identities (3) and (14) used above express the quantities $1/A$ and $1/A^L$ as the Laplace transforms of 1 and x^{L-1} defined respectively. More general, if $f(s)$ is the Laplace transform of $g(x)$,

$$f(s)=\int_0^{\infty}e^{-sx}g(x)\,dx,$$

then the sum $S_I = \sum_{n \in I} f(n)$, where $I \subset \mathbb{Z}$, can be written in the form

$$S_I = \int_0^\infty g(x)\left(\sum_{n \in I} e^{-nx}\right) dx,$$

assuming that the operations of summation and integration are interchangeable. If, in addition, the sum inside parentheses can be found explicitly, then we obtain an integral representation of S_I.

The Laplace transform is a valuable tool in the solution of differential equations. Unfortunately, the Laplace transform does not enjoy the same popularity in the summation of series. We have tried in this note to present some of the limitless possibilities that the method offers. As a final illustration, we propose the following problem:

Show that

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n+a} = \int_0^1 \frac{t^a}{1+t}\, dt.$$

Using this formula, derive the well-known result

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln 2 \approx 0.693,$$

and the less well-known result

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n+1/2} = 2 - \frac{\pi}{2} \approx 0.429.$$

We hope that this will motivate readers to explore more aspects of the method presented here, and to establish additional results in the same spirit.

Acknowledgment. The author thanks the referees for their valuable comments and help in revising the paper.

REFERENCES

1. I. Fischer, Problem 23 in Problem Section, *Math Horizons*, February 1995.
2. P. Ramond, *Field Theory: A Modern Primer*, Addison-Wesley, Reading, MA, 1994.
3. W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, New York, NY, 1976.
4. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover, New York, NY, 1972.

A Polynomial Approach to a Diophantine Problem

LENNY JONES
Shippensburg University
Shippensburg, PA 17257

1. Introduction About the third century A.D. Diophantus posed the question of finding four rational numbers such that the product of any two, increased by one, is the square of a rational number. He solved this problem by finding

$$\frac{1}{16}, \quad \frac{33}{16}, \quad \frac{17}{4}, \quad \frac{105}{16}. \tag{1}$$

In the seventeenth century, Fermat observed that the set $\{1, 3, 8, 120\}$ satisfies the conditions of Diophantus; Fermat asked if there was a fifth integer that could be added to this set such that the product of any two elements of the set is always one less than a perfect square. Not until 1969 was a negative answer given, when Davenport and Baker **[1]** proved that 120 is the only fourth number that can be added to the set $\{1, 3, 8\}$ and have the new set satisfy these conditions. Similarly, Veluppillai **[6]** showed that 420 is the only fourth number that can be added to the set $\{2, 4, 12\}$ and have the new set satisfy these conditions.

Many other four-element sets have been found that satisfy the conditions of Diophantus. For example, in the eighteenth century, Euler **[2]** observed that if $ab + 1$ is a perfect square, then the set

$$\left\{a, \quad b, \quad a+b+2\sqrt{ab+1}, \quad 4(a+\sqrt{ab+1})(b+\sqrt{ab+1})\sqrt{ab+1}\right\} \tag{2}$$

has the desired property. This solution implies that there are infinitely many four-element sets of positive integers that satisfy the conditions of Diophantus. Independently of Euler's solution, Hoggatt and Bergum **[3]** proved that the four-element set

$$\{F_{2n}, \quad F_{2n+2}, \quad F_{2n+4}, \quad 4F_{2n+1}F_{2n+2}F_{2n+3}\}$$

where F_n is the n^{th} Fibonacci number ($F_1 = F_2 = 1$), satisfies the conditions of Diophantus. This result is a special case of (2), with $a = F_{2n}$ and $b = F_{2n+2}$. Other infinite families of four-element sets not produced by Euler's solution have also been found. So far, however, no one has been able to prove that a method exists to generate *all* such four-element sets. For further interesting and extensive discussions, visit **[7]**, **[8]** or **[9]** on the World Wide Web.

The search for five-element sets satisfying the conditions of Diophantus has not been as successful. Euler **[2]** showed that the five-element set of rational numbers

$$\left\{1, \quad 3, \quad 8, \quad 120, \quad \frac{777480}{8288641}\right\}$$

satisfies the conditions of Diophantus, but it is still unknown whether any five-element set of positive integers has this property.

The numbers in (1) can be rewritten as

$$x, \quad x+2, \quad 4x+4, \quad 9x+6 \tag{3}$$

where $x = \frac{1}{16}$. If we think of the elements in (3) as polynomials in x, then the first three are such that the product of any two increased by one is the square of a polynomial with nonnegative integer coefficients. Since x can be replaced by any number, this also produces infinitely many four-element sets of numbers satisfying the conditions of Diophantus. Using ideas from recursion theory and classical methods of solving Pell equations, B.W. Jones **[4]** found all of the infinitely many polynomials $c(x)$ and $d(x)$ such that the set

$$\{x,\quad x+2,\quad c(x),\quad d(x)\}$$

of polynomials has the property that the product of any two of them increased by one is the square of a polynomial with integer coefficients. Jones also showed that no set of five polynomials, containing x and $x+2$, exists with this property.

Various generalizations of these problems have been addressed. For example, Dickson **[2]** gives a method for finding sets of four rational numbers such that the product of any *three* is one less than a perfect square of a rational number. Zheng **[10]** observed that the set $\{1, 2, 5\}$ has the property that the product of any two elements is one *more* than a perfect square. Zheng also proved that no fourth positive integer can be added to this set and have the new set retain these properties. Similarly, the set $\{1, 2, 7\}$ has the property that the product of any two elements is two less than a perfect square. It is easy to prove, using congruence arguments modulo 4, that no four-element set of integers has this particular property.

In this paper we show, using elementary techniques, that for every positive integer k there exists a set of four polynomials with nonnegative rational coefficients such that the product of any two of them increased by k^2 is the square of a polynomial with nonnegative rational coefficients. For $k = 1$, 2, and 4, the polynomials can actually be taken to have nonnegative *integer* coefficients. This enables us to deduce that for each positive integer k there are infinitely many four-element sets of positive integers with the property that the product of any two of them increased by k^2 is a perfect square.

2. The polynomials and the Diophantine problem Since the first three polynomials in (3) work for $k = 1$, we try to generalize this set to arbitrary k. The product of the first two from (3) increased by one is $x(x+2)+1$, which is $(x+1)^2$. If we choose x and $x+2k$ as our first two polynomials, we achieve $(x+k)^2$ in the generalization. To generalize $4x+4$ from (3), we first observe that $x(4x+4)+1 = (2x+1)^2$. So, if we choose our third polynomial to be $4x+4k$, we achieve $(2x+k)^2$ in our generalization since $(2x+k)^2 = x(4x+4k)+k^2$. The polynomial $4x+4k$ also works with $x+2k$ since $(x+2k)(4x+4k)+k^2 = (2x+3k)^2$. Hence, for any positive integer k, the three polynomials $f_1 = x$, $f_2 = x+2k$, and $f_3 = 4x+4k$ are such that the product of any two increased by k^2 is a square.

Let f_4 be the desired fourth polynomial. The conditions that $f_1 f_4 + k^2$ and $f_2 f_4 + k^2$ be squares give the equations

$$xf_4 + k^2 = (s(x))^2 \tag{4}$$

$$(x+2k)f_4 + k^2 = (t(x))^2 \tag{5}$$

for some polynomials $s(x)$ and $t(x)$; note that $s(x)$ and $t(x)$ have the same degree. Equating leading coefficients and constant terms in (4) and (5) shows that $s(x)$ and $t(x)$ have the same leading coefficients and that the constant term of $s(x)$ is k. Suppose first that both $s(x)$ and $t(x)$ have degree two. (The reason for this is that nothing appears to be gained from choosing $s(x)$ and $t(x)$ of higher degree. This will be illustrated in the next section.) Write

$$s(x) = ax^2 + bx + k \quad \text{and} \quad t(x) = ax^2 + cx + d.$$

Substituting this into (4) and solving for f_4 gives

$$f_4 = a^2x^3 + 2abx^2 + (2ak + b^2)x + 2bk. \tag{6}$$

Substituting this into (5) now yields

$$(2a^2k + 2ab - 2ac)x^3 + (4abk - 2ad + 2ak + b^2 - c^2)x^2 + (4ak^2 + 2b^2k + 2bk - 2cd)x + (4bk^2 - d^2 + k^2) = 0. \tag{7}$$

Equating each coefficient and the constant term in (7) to zero gives a nonlinear system. Using MAPLE to solve this system produces the following solutions:

$$(a,b,c,d,k) = (a,b,b,0,0),\ (0,b,b,0,0),\ (0,b,-b,0,0),\ (0,0,0,k,k),\ (0,0,0,-k,k),\ (0,2,2,3k,k),\ (0,2,-2,-3k,k),\ \left(\tfrac{4}{k},6,10,5k,k\right).$$

The first five solutions can be discarded: in the first three, k is zero; in the next two, $f_4 = 0$. The sixth and seventh solutions produce $f_4 = f_3$ which is what we would have gotten had we started with first degree polynomials $s(x)$ and $t(x)$. The last solution above, substituted into (6), gives a third degree polynomial with rational coefficients:

$$f_4 = \frac{16}{k^2}x^3 + \frac{48}{k}x^2 + 44x + 12k.$$

Computing $f_if_j + k^2$ for $1 < i < j < 4$ gives these results:

i,j	2	3	4
1	$(x+k)^2$	$(2x+k)^2$	$(\frac{4}{k}x^2 + 6x + k)^2$
2		$(2x+3k)^2$	$(\frac{4}{k}x^2 + 10x + 5k)^2$
3			$(\frac{8}{k}x^2 + 16x + 7k)^2$

We have proven the following theorem:

THEOREM 1. *For each positive integer k the set of four polynomials*

$$\left\{x,\quad x+2k,\quad 4x+4k,\quad \frac{16}{k^2}x^3 + \frac{48}{k}x^2 + 44x + 12k\right\}$$

has the property that the product of any two of them increased by k^2 is the square of a polynomial with nonnegative rational coefficients. If $k = 1$, 2, or 4, the polynomials actually have nonnegative integer coefficients.

From Theorem 1, we can easily prove the following Corollary.

COROLLARY 2. *For each positive integer k there exist infinitely many sets of four positive integers such that the product of any two elements from a particular set increased by k^2 is a perfect square.*

Proof. For $k = 1$, 2 or 4 this is immediate from Theorem 1, since we may take x to be any positive integer. For all other values of k we can generate infinitely many sets by simply choosing values of x such that k divides x. This guarantees that the fourth polynomial in Theorem 1 has integer values.

3. Conclusions and open questions Following are some interesting related questions that arise naturally:

(a) Is there a positive integer k and a set of five or more polynomials such that the product of any two of them increased by k^2 is a square?
(b) Is the fourth polynomial in Theorem 1 unique?
(c) Is there a positive integer k and a set of five or more positive integers such that the product of any two of them increased by k^2 is a perfect square?

Our investigations suggest negative answers to questions (a) and (c) but a positive answer to (b). For example, if we take $s(x)$ and $t(x)$ in (4) and (5) to be cubic rather than quadratic polynomials, and search as before for a fourth polynomial, we do indeed find a polynomial

$$p(x) = \frac{64}{k^4}x^5 + \frac{320}{k^3}x^4 + \frac{592}{k^2}x^3 + \frac{496}{k}x + 184x + 24k.$$

But $p(x)$ no longer "works" with $4x + 4k$, so $p(x)$ can neither be a fifth polynomial for the set in Theorem 1 nor a replacement for f_4. However, $p(x)$ does "work" with f_4; this gives us a new four-element set

$$\{x, \quad x+2k, \quad f_4, \quad p(x)\}$$

such that the product of any two increased by k^2 is the square of a polynomial with nonnegative rational coefficients. Because this process of generating a new fourth polynomial seems to force the replacement of the third polynomial, one conjectures that this procedure can be continued to find higher and higher degree polynomials that complete a four-element set, but that a five-element set is impossible. In addition, as mentioned before, it is shown in [4] that there is no five-element set corresponding to $k = 1$. Concerning question (c), it is not necessary to first find polynomials that have this property. As we have seen, non-polynomial techniques have been used to produce infinitely many different four-element sets with the desired property. Upon examination of these techniques, however, they also suggest a negative answer to question (c).

Acknowledgment. The author thanks the referees for their valuable suggestions.

REFERENCES

1. H. Davenport and A. Baker, The equations $3x^2 - 2 = y^2$ and $8x^2 - 7 = z^2$, *Quarterly J. of Math.* 20 (1969), 129–137.
2. L. E. Dickson, *History of the Theory of Numbers*, vol. 2, Chelsea, New York, NY, 1952.
3. V. E. Hoggatt and G. E. Bergum, A problem of Fermat and the Fibonacci sequence, *Fib. Quarterly* 15 (1977), 323–330.
4. B. W. Jones, A variation on a problem of Davenport and Diophantus, *Quarterly J. of Math.* 27 (1976), 349–353.
5. J. Roberts, *Lure of the Integers*, Math. Assoc. of America, Washington, DC, 1992.
6. M. Veluppillai, The equations $z^2 - 3y^2 = -2$ and $z^2 - 6x^2 = -5$, in *A Collection of Manuscripts Related to the Fibonacci Sequence*, V. E. Hoggatt, Jr. and M. Bicknell-Johnson (eds.), Fibonacci Association, 1980, pp. 71–75.
7. www.weburbia.demon.co.uk/pg/diophant.htm
8. www.seanet.com/~ksbrown/number.htm/kmath289.htm
9. www.seanet.com/~ksbrown/number.htm/kmath290.htm
10. D. X. Zheng, On the systems of Diophantine equations $y^2 - 2x^2 = 1, z^2 - 5x^2 = 4$ and $y^2 - 5x^2 = 4, z^2 - 10x^2 = 9$, *Sichuan Daxue Xuebao* 24 (1987), 25–29.

On Solid Angles and the Volumes of Regular Polyhedra

JEFFREY NUNEMACHER
Ohio Wesleyan University
Delaware, OH 43015

In the interesting article [2], Paul Shutler presents three different approaches to the problem of computing the volumes of the regular (Platonic) polyhedra. These five polyhedra (the cube and the regular tetrahedron, octahedron, dodecahedron, and icosahedron) display perfect symmetry, i.e., all faces are congruent regular polygons and all polyhedral angles are also congruent to each other. The most mysterious of his methods is based on the Biot–Savart Law of magnetostatics and the concept of solid angle. In this note we show how the physics used in this approach can be replaced by a bit of vector calculus. The result is a method that is geometric and conceptually quite simple.

We first discuss the geometric concept of solid angle, a natural generalization of the idea of planar angle to dimension three. Let K denote a surface in $\mathbb{R}^3$ that intersects each ray from the origin in at most one other point (see FIGURE 1). The solid cone of

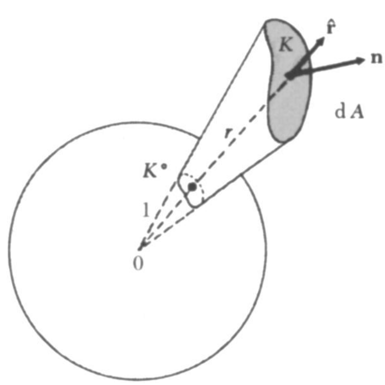


FIGURE 1
(from [1], p. 521)

lines that meet K intersects the unit sphere centered at the origin in a surface K^*. The *solid angle* Ω subtended by K at the origin is defined to be the area of K^*. Notice that the two-dimensional version of this idea results in the ordinary angle subtended by a curve in the plane at the origin. It is possible to express Ω as an integral provided that K is sufficiently (piecewise) smooth. Let $\mathbf{r}$ denote the vector from the origin to a general point in K, $r = |\mathbf{r}|$, $\hat{\mathbf{r}} = \mathbf{r}/r$, and $\mathbf{n}$ the outward-pointing unit normal to K at $\mathbf{r}$. Then, since the inverse square field $\mathbf{r}/r^3$ is divergence-free, a direct application of the divergence theorem shows (see [1], p. 521) that Ω is equal to the surface integral

$$\Omega = \int_K\int \frac{\mathbf{r}\cdot\mathbf{n}}{r^3}\, dA. \tag{1}$$

Consider a regular convex polyhedron P inscribed in a sphere of radius R. Assume that P has N congruent faces, each with n sides. Following [**2**], we introduce three auxiliary quantities: w, the distance from the center of each face to the center of the sphere; l, the length of each side of a face; and h, the perpendicular distance from the center of a face to a side (classically known as the *apothem*). By decomposing P into a union of N congruent pyramids, each with one vertex at the origin, and then using the usual formula for the volume of a pyramid, we find a simple expression for the volume of P: $V = Nnlhw/6$. The problem now is finding constraints to eliminate the quantities l, h, and w, so that V is given in terms of only N, n, and R. Such a formula would then allow us to determine, for instance, the volume of a dodecahedron inscribed in the unit sphere ($R = 1$, $N = 12$, and $n = 5$).

The Pythagorean theorem in three dimensions shows that $R^2 = w^2 + h^2 + l^2/4$, and simple trigonometry on one face yields $\tan(\pi/n) = l/(2h)$. To find a third constraint, we consider the solid angle Ω subtended by one face of the polyhedron, where the origin has been placed at the center of the sphere. By the symmetry of P, we have $\Omega = 4\pi/N$, since the area of the unit sphere is 4π. But Ω can also be calculated using (1), and the resulting surface integral can be evaluated (by hand!). We sketch the calculation.

Divide a face into n triangles with a common vertex placed in the center of the face. Let K denote the right triangle obtained by taking half of one of these n isosceles triangles, which we situate in the xy-plane as shown in FIGURE 2. Then the

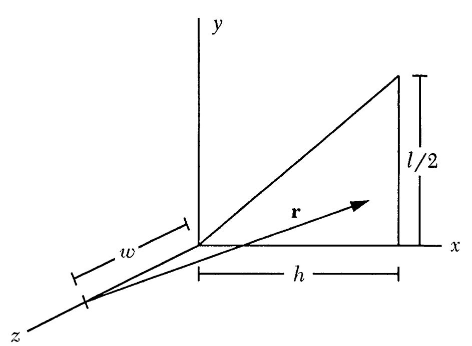


FIGURE 2

center of the sphere lies at distance w directly above the origin, and $\mathbf{r}$ points from the center of the sphere to a general point in the triangle. We have $\mathbf{r}\cdot\mathbf{n} = r\cos\theta$, where θ is the angle that $\mathbf{r}$ makes with the z-axis. Thus $r = \sqrt{x^2 + y^2 + w^2}$ and $\cos\theta = w/r$; hence Ω is given by the iterated integral

$$\Omega = 2nw \int_0^h \int_0^{\frac{lx}{2h}} \frac{1}{\left(x^2 + y^2 + w^2\right)^{3/2}}\, dy\, dx.$$

This integral can be evaluated by using, in the outer integral, the formula

$$\int \frac{x}{(x^2 + a^2)\sqrt{x^2 + b^2}}\, dx = -\frac{1}{\sqrt{a^2 - b^2}} \arcsin \frac{\sqrt{a^2 - b^2}}{\sqrt{a^2 + x^2}} + C,$$

valid when $a > b$. The formula can be derived using standard calculus techniques. We write the result in the form $w = \lambda h$, where

$$\lambda = \frac{\sin\left\{\frac{\pi}{n}\left(1 - \frac{2}{N}\right)\right\}}{\sqrt{\sin^2\frac{\pi}{n} - \sin^2\left\{\frac{\pi}{n}\left(1 - \frac{2}{N}\right)\right\}}}.$$

This is the third of our needed constraints. Using elementary algebra, it is now possible to express V (as well as l, a, h, w, and the surface area A) in terms of R, N, and n. In particular, we obtain the formula for V which was derived in **[2]** using physics:

$$V = \frac{\frac{1}{3}NnR^3\lambda\tan\left(\frac{\pi}{n}\right)}{\left\{1 + \lambda^2 + \tan^2\frac{\pi}{n}\right\}^{3/2}} \qquad (2)$$

It follows from (2) that the volume (to four decimal places) of a dodecahedron inscribed in the unit sphere is 2.7852; that of an icosahedron is 2.5363. Some readers may be surprised that the dodecahedron, which has fewer faces, has the larger volume.

Note Readers may be interested in perusing the *International Journal of Mathematical Education in Science and Technology*, in which **[2]** was published. This journal contains many articles of mathematical interest on a broad range of topics, at about the same level as *Mathematics Magazine*.

REFERENCES

1. Peter Baxandall and Hans Liebeck, *Vector Calculus*, Oxford University Press, New York, NY, 1986.
2. Paul Shutler, Magnetostatics and the volumes of the platonic solids, *Int. J. Math. Educ. Sci. Technol.* 27 (1996), 413–419.

A Web-Searchable Database for *Mathematics Magazine*

We are happy to inform readers of the *Magazine*'s Web-searchable database of Articles, Notes, Proofs Without Words, and more. The database can be found through MAA Online, at

`http://www.maa.org`

(click on Journals and then on Mathematics Magazine), or directly from its "home" at Harvey Mudd College, at

`http://www.math.hmc.edu/MathMag/`

The database now contains information going back to 1927, the *Magazine*'s inception.

The πs Go Full Circle

RUSSELL EULER
JAWAD SADEK
Northwest Missouri State University
Maryville, MO 64468

Introduction To a great extent, mathematics is a study of patterns and generalizations. This paper illustrates the spirit of generalization; it can be viewed as exploration in undergraduate geometry and analysis. We hope it encourages further related studies.

The numerical value of π has been mentioned in a variety of historical settings. For example, the Biblical value of π is 3, as can be deduced from either I Kings 7:23* or II Chronicles 4:2. In 1897, the Indiana State Legislature introduced Bill No. 246† which attempted to determine the value of π by legislation (see [**1**]). Had this bill passed, π would have been 4 in Indiana. Our definition of π will have different values depending on which metric is used. Also, at least one of the new πs (namely, π_1) that we will introduce occurs naturally in the context of taxicab geometry which has a wide range of applications in urban geography (see [**2**], chapters 2 and 6).

The usual definition of π is the ratio of the circumference (C) of a circle to the diameter (d) of the circle. The generalized π we discuss here will be related to an alternating series. An error analysis for approximating this generalized π will also be given.

Basic definitions and preliminary results Let $1 \le p < \infty$ and let $r > 0$. The set $\{(x, y) \in \mathbb{R}^2 \mid |x|^p + |y|^p = r^p\}$ will be called a *generalized circle* with *center* $O = (0, 0)$ and *radius* r; it will be denoted by $C_p(r)$.

For $1 \le p < \infty$, we define the distance between points $X = (x_1, x_2)$ and $Y = (y_1, y_2)$ by

$$D_p(X,Y) = \left[|x_1 - y_1|^p + |x_2 - y_2|^p\right]^{1/p}. \tag{1}$$

Let the function given by $y = f(x)$ for $a \le x \le b$ represent a smooth curve in the xy-plane. We define the *p-arc length* of f between a and b as

$$\int_a^b \left[1 + |f'(x)|^p\right]^{1/p} dx. \tag{2}$$

If $p = 2$, then (1) and (2) are the usual (2-dimensional) Euclidean distance and arc length formulas. The metric $D_1(X,Y)$ is called the *taxicab metric* since the distance from X to Y is the sum of the lengths of a vertical segment ("street") and a horizontal segment connecting X to Y, see [**2**].

*"Now he made the sea of cast metal ten cubits from brim to brim, circular in form, and its height was five cubits, and thirty cubits in circumference."

†"Be it enacted by the General Assembly of the State of Indiana: It has been found that a circular area is to the square on a line equal to the quadrant of the circumference, as the area of an equilateral rectangle is to the square on one side. . . ."

If $p=\infty$, the distance between $X=(x_1, x_2)$ and $Y=(y_1, y_2)$ is

$$D_\infty(X,Y) = \max(|x_1-y_1|, |x_2-y_2|). \tag{3}$$

For all p, $1 \le p \le \infty$, $C_p(r)$ is the set $\{X \in \mathbb{R}^2 \mid D_p(X,O)=r\}$. FIGURE 1 shows the set $C_p(1)$ for $p=1$, 2, 3, 4, and ∞. When $p=2$, $C_2(1)$ is the standard unit circle. The figure illustrates that as $p \to \infty$, the sets $C_p(1)$ approach the set $C_\infty(1)$. See [4, p. 119, Exercise 1] for related ideas.

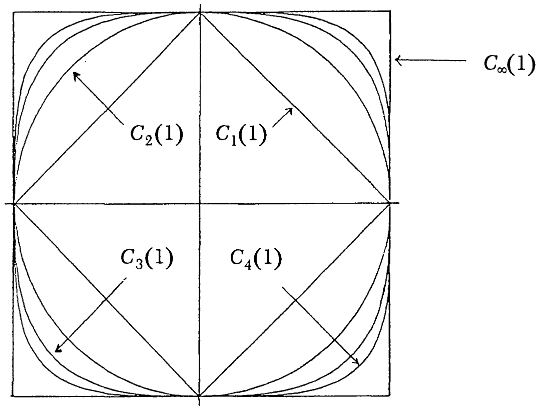


FIGURE 1
$C_p(1)$ for $p=1,2,3,4,\infty$

Let C and d represent, respectively, the p-arc length of the circumference and diameter of $C_p(r)$ where $d=2r$. We define

$$\pi_p = \frac{C}{d}. \tag{4}$$

(We will show later that, for fixed p, π_p is independent of r.)

Examples. Consider the circle $C_1(1)$. By symmetry, $C=4D_1((0,1),(1,0))=8$. Here, $d=2$, so $\pi_1=4$.

For $p=\infty$, the circumference of $C_\infty(1)$ is $C=4D_\infty((1,1),(1,-1))=4\cdot 2=8$. Again, $d=2$ so $\pi_\infty=4$.

Approximating π_p Consider the generalized circle $C_p(r)$, with $d=2r$. Since the graph of the equation $|x|^p+|y|^p=r^p$ is symmetric with respect to the x and y axes, we will restrict our attention to the first quadrant. So, for $p<\infty$, we write $y=f(x)=(r^p-x^p)^{1/p}$, and (2) gives

$$C = 8\int_0^{(\frac{r^p}{2})^{1/p}} \left[1+x^{(p-1)p}(r^p-x^p)^{1-p}\right]^{1/p} dx. \tag{5}$$

For arbitrary p, these integrals cannot be evaluated in closed form. Notice, however, that for $p=1$, equation (5) becomes $C=8\int_0^{r/2} 2\,dx = 8r$, which agrees with our earlier result. Moreover, the transformation $x=rt$ shows that C/d is independent of r since (5) becomes

$$C = 8r\int_0^{\frac{1}{2^{1/p}}} \left[1+t^{(p-1)p}(1-t^p)^{1-p}\right]^{1/p} dt.$$

Thus, without loss of generality, we will let $r=1$ in (5).

By the Binomial Theorem [**3**, pp. 627–628],

$$(1+y)^{1/p} = 1 + \sum_{n=1}^{\infty} \binom{1/p}{n} y^n \text{ for } -1 < y < 1, \text{ where}$$

$$\binom{1/p}{n} = \frac{\frac{1}{p}\left(\frac{1}{p}-1\right)\cdots\left(\frac{1}{p}-n+1\right)}{n!}.$$

Therefore,

$$\left[1 + x^{(p-1)p}(1-x^p)^{1-p}\right]^{1/p} = 1 + \sum_{n=1}^{\infty} \binom{1/p}{n} x^{(p-1)pn}(1-x^p)^{(1-p)n}.$$

Then, using (4) and (5),

$$\pi_p = 4\int_0^{(\frac{1}{2})^{1/p}} \left[1 + \sum_{n=1}^{\infty} \binom{1/p}{n} x^{(p-1)pn}(1-x^p)^{(1-p)n}\right] dx.$$

The Weierstrass M-test implies that the series in the integral converges uniformly on $[0,(1/2)^{1/p}]$, so we can write

$$\pi_p = 4\left(\frac{1}{2}\right)^{1/p} + 4\sum_{n=1}^{\infty} \binom{1/p}{n} \int_0^{(\frac{1}{2})^{1/p}} x^{(p-1)pn}(1-x^p)^{(1-p)n}\,dx.$$

Now we let $x^p = t$; then $x = t^{1/p}$ and $dx = (1/p)t^{1/p-1}\,dt$. Making the change of variables gives

$$\pi_p = 4\left(\frac{1}{2}\right)^{1/p} + \frac{4}{p}\sum_{n=1}^{\infty} \binom{1/p}{n} \int_0^{1/2} t^{(p-1)n+1/p-1}(1-t)^{(1-p)n}\,dt. \qquad (6)$$

For example, if $p = 1$, then (6) becomes $\sum_{n=1}^{\infty}\binom{1}{n} = 1$. If $p = 2$, then (6) becomes

$$\sum_{n=1}^{\infty} \binom{1/2}{n} \int_0^{1/2} t^{n-1/2}(1-t)^{-n}\,dt$$

$$= \frac{1}{2}\int_0^{1/2} t^{1/2}(1-t)^{-1}\,dt - \frac{1}{8}\int_0^{1/2} t^{3/2}(1-t)^{-2}\,dt + \cdots = \frac{\pi}{2} - \sqrt{2}.$$

It is easy to see that the series in (6) is alternating and the terms of the series are nonincreasing in magnitude. The error resulting from approximating π_p by the first N terms satisfies

$$|\text{error}| \le \frac{4}{p}\left|\binom{1/p}{N}\right| \left|\int_0^{1/2} t^{(p-1)N+1/p-1}(1-t)^{(1-p)N}\,dt\right|$$

$$\le \frac{2}{p}\left|\binom{1/p}{N}\right| < \frac{2}{p}\left(\frac{1}{pN}\right) = \frac{2}{p^2 N}.$$

For instance, if $p > M$ and $N = 10$, then this error will be less than $1/(5M^2)$.

To approximate π_p, we will give upper and lower bounds for the integrals in (6), as follows:

$$4\left(\frac{1}{2}\right)^{1/p} + \frac{4}{p}\left(\frac{1}{2}\right)^{1/p} \sum_{n=1}^{\infty} \binom{\frac{1}{p}}{n} \frac{\left(\frac{1}{2}\right)^{(p-1)n}}{(p-1)n+\frac{1}{p}}$$

$$\le \pi_p \le 4\left(\frac{1}{2}\right)^{1/p} + \frac{4}{p}\left(\frac{1}{2}\right)^{1/p} \sum_{n=1}^{\infty} \frac{\binom{1/p}{n}}{(p-1)n+\frac{1}{p}}. \tag{7}$$

To prove (7), first notice that $1 \le (1-t)^{(1-p)n} \le 2^{(p-1)n}$ since $0 \le t \le 1/2$. We let a_n be the nth term of the infinite series in (6) and let

$$b_n = 2^{(p-1)n}\binom{1/p}{n}\int_0^{1/2} t^{(p-1)n+1/p-1}\,dt.$$

Since the series in (6) is alternating and absolutely convergent, the terms can be rearranged to assure that $a_{n+1}+a_n$ and $b_{n+1}+b_n$ are positive. Combining integrals and elementary estimates, straightforward but tedious calculations show that

$$a_{n+1}+a_n \le b_{n+1}+b_n.$$

The upper bound in (7) follows from these observations.

The lower bound in (7) is obtained similarly. Letting

$$c_n = \binom{1/p}{n}\int_0^{1/2} t^{(p-1)n+1/p-1}\,dt,$$

one shows that $a_{n+1}+a_n \ge c_{n+1}+c_n$, and the lower bound follows.

Notice that when $p=1$, (7) becomes $4 \le \pi_1 \le 4$, which agrees with our earlier result.

The entries in Table 1 were computed using (7). For the upper bound in the table, we used only the first term of the series in the upper bound in (7) since additional

TABLE 1. Estimates for π_p

Lower Bound	π_p	Upper Bound
4	π_1	4
3.0	π_2	3.2
3.4	π_5	3.5
3.732	π_{10}	3.736
3.8906	π_{25}	3.8908
3.94493	π_{50}	3.94496
3.96320	π_{75}	3.96321
3.97236	π_{100}	3.97237
3.997228371	π_{1000}	3.997228375
3.99972275073	π_{10000}	3.99972275074
3.99997227420886	π_{100000}	3.99997227420887
3.99999722741223866	$\pi_{1000000}$	3.99999722741223867
⋮	⋮	⋮
4	π_∞	4

terms did not contribute significantly. Similarly, for the lower bound in the table, we used only the first two terms of the series in the lower bound in (7). For each p, we truncated the decimal expansion when the last decimal digit for the upper bound of π_p starts to differ from the last decimal digit for the lower bound of π_p. Table 1 motivated the title of this note, since the π_p's "started" and "ended" at 4.

Questions We end with several questions for possible undergraduate research.

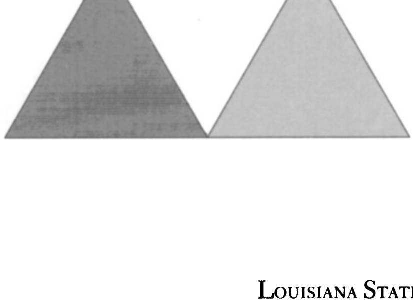
1. For what value of $p \geq 1$ is π_p a minimum?
2. Does there exist p such that $\pi_p = 3$? If so, find that p.
3. Does there exist $p \neq 1$ or $p \neq \infty$ for which π_p is rational?

Acknowledgment. The authors would like to thank the referees for their valuable suggestions.

REFERENCES

1. W. E. Edington, "House Bill No. 246, Indiana State Legislature, 1897," Proceedings of the Indiana Academy of Science, 45 (1935), 206–210.
2. E. F. Krause, *Taxicab Geometry: An Adventure in Non-Euclidean Geometry*, Dover Publications, Inc., New York, NY, 1986.
3. R. E. Larson, R. P. Hostetler, B. H. Edwards, *Calculus with Analytic Geometry*, sixth edition, D. C. Heath and Company, New York, NY, 1998.
4. W. Rudin, *Real and Complex Analysis*, second edition, McGraw-Hill Book Company, New York, NY, 1974.

Proof Without Words: $\frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \cdots = \frac{1}{3}$.



—RICK MABRY
LOUISIANA STATE UNIVERSITY IN SHREVEPORT
SHREVEPORT, LA 71115-2399

PROBLEMS

GEORGE T. GILBERT, *Editor*
Texas Christian University

ZE-LI DOU, KEN RICHARDSON, and SUSAN G. STAPLES, *Assistant Editors*
Texas Christian University

Proposals

To be considered for publication, solutions should be received by July 1, 1999.

1564. *Proposed by Wu Wei Chao, Guang Zhou Normal University, Guang Zhou City, Guang Dong Province, China.*

Let P be the intersection of the diagonals of convex quadrilateral $ABCD$ with $\angle BAC + \angle BDC = 180°$. Suppose that the distance from A to the line BC is less than the distance from D to BC. Show that

$$\left(\frac{AC}{BD}\right)^2 > \frac{AP\cdot CD}{DP\cdot AB}.$$

1565. *Proposed by Joaquín Gómez Rey, I. B. "Luis Buñuel," Alcorcón, Madrid, Spain.*

A philatelist has $(n+1)!-1$ stamps and decides to sell a portion of them in n steps. In each step he will sell $1/(k+1)$ of his remaining total plus $1/(k+1)$ of one stamp, for $k = 1, \dots, n$. However, these n steps are ordered randomly. Let p_n denote the probability that he does not sell the same number of stamps in two successive steps. Evaluate $\lim_{n\to\infty} p_n$.

1566. *Proposed by Stephen G. Penrice, Morristown, New Jersey.*

Let circle C circumscribe (nondegenerate) rectangle R. Let α be the ratio of the area of C to the area of R, and let β be the ratio of the circumference of C to the perimeter of R. Show that α and β cannot both be algebraic.

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must, in general, be accompanied by solutions and by any bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution.

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

Solutions and new proposals should be mailed to George T. Gilbert, Problems Editor, Department of Mathematics, Box 298900, Texas Christian University, Fort Worth, TX 76129, or mailed electronically (ideally as a LaTeX file) to `g.gilbert@tcu.edu`. *Readers who use e-mail should also provide an e-mail address.*

1567. *Proposed by Dennis Spellman, Philadelphia, Pennsylvania, and William P. Wardlaw, United States Naval Academy, Annapolis, Maryland.*

Find the Smith normal form over the integers of the $n \times n$ matrix A with entries $a_{ij} = j^i$.

(The Smith normal form of an integral matrix A is a diagonal matrix D that is obtained through applying to A a sequence of elementary row and column operations with integral matrices of determinant ± 1. It is defined to be the unique such matrix whose diagonal entries d_{ii} are nonnegative integers satisfying (i) $d_{ii} \neq 0$ if and only if $i \leq \text{rank}(A)$ and (ii) d_{ii} divides $d_{(i+1)(i+1)}$ for $1 \leq i < \text{rank}(A)$.)

1568. *Proposed by Emre Alkan, student, University of Wisconsin, Madison, Wisconsin.*

Determine which finite, nonsimple groups G satisfy the following: If H and K are subgroups of G, then either (i) $H \subset K$, (ii) $K \subset H$, or (ii) $H \cap K = \{e\}$.

Quickies

Answers to the Quickies are on page 70.

Q887. *Proposed by Michael Golomb, professor emeritus, Purdue University, West Lafayette, Indiana.*

For $-\pi/2 < x < \pi/2$, prove the inequality

$$\sin x \ln\left(\frac{1 + \sin x}{1 - \sin x}\right) \geq 2x^2.$$

Q888. *Proposed by George T. Gilbert, Texas Christian University, Fort Worth, Texas.*

Standard sports folklore is that the longer a series of games, the likelier it is that the better team wins. Consider a series played until one team wins n games. Prove that if one team has probability $p > 1/2$ of winning a single game, then the probability this team wins the series increases with n.

Solutions

An Upper Bound for a Product — February 1998

1539. *Proposed by Donald Knuth, Stanford University, Stanford, California.*

Let p and q be positive numbers with $p + q = 1$, and suppose $0 < \epsilon < q$. Prove that

$$\left(\frac{p}{p+\epsilon}\right)^{p+\epsilon}\left(\frac{q}{q-\epsilon}\right)^{q-\epsilon} < e^{-2\epsilon^2}.$$

(The condition $0 < \epsilon < q$ is a minor correction of the original problem statement.)

Solution by Eugene Lee, Alias | Wavefront, Inc., Seattle, Washington.

Fix any positive p, q with $p + q = 1$. We will show that the positive function

$$f(\epsilon) = e^{2\epsilon^2}\left(\frac{p}{p+\epsilon}\right)^{p+\epsilon}\left(\frac{q}{q-\epsilon}\right)^{q-\epsilon},$$

defined over $0 \le \epsilon < q$, satisfies $f(\epsilon) < 1 = f(0)$ for $\epsilon > 0$. Logarithmic differentiation gives

$$\phi(\epsilon) := \frac{f'(\epsilon)}{f(\epsilon)} = 4\epsilon + \log\frac{p(q-\epsilon)}{q(p+\epsilon)}.$$

Now $\phi(0) = 0$ and

$$\phi'(\epsilon) = 4 - \left(\frac{1}{p+\epsilon} + \frac{1}{q-\epsilon}\right) = 4 - \frac{1}{(p+\epsilon)(q-\epsilon)}.$$

By the arithmetic mean-geometric mean inequality, $(p+\epsilon)(q-\epsilon) \le 1/4$, hence $\phi'(\epsilon) \le 0$. Note that equality can possibly happen at most for a single ϵ, when $p + \epsilon = q - \epsilon$. This implies $\phi(\epsilon) < 0$ and $f'(\epsilon) < 0$ for all $\epsilon > 0$, hence the claim.

Also solved by Jean Bogaert (Belgium), Paul Bracken (Canada), John Christopher, Daniele Donini (Italy), Mordechai Falkowitz (Canada), Kazuo Goto (Japan), Joe Howard, Hans Kappus (Switzerland), Kee-Wai Lau (China), Can A. Minh, Gao Peng, Heinz-Jürgen Seiffert (Germany), and the proposer. There was one incorrect solution.

The Limit of a Sequence of Polynomial Roots — February 1998

1540. *Proposed by Michael Golomb, Purdue University, West Lafayette, Indiana.*

(a) Show that $x^n + (x-1)^n - (x+1)^n$ has a unique non-zero real root r_n.
(b) Show that r_n increases monotonically.
(c) Evaluate $\lim_{n\to\infty} r_n/n$.

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

(a), (b) Set $f_n(x) := x^n + (x-1)^n - (x+1)^n$ and note that $f_n(0) = 0$ if n is even and $f_n(0) = -2$ if n is odd. It is easily checked that $r_1 = 2$ is the unique non-zero root of $f_1(x)$. For $n = 1, \dots, k$, assume that $f_n(x)$ has a unique non-zero root r_n and that $r_1 < r_2 < \cdots < r_k$. Assume further that $f_n(x)$ is positive on (r_n, ∞), negative on $(0, r_n)$, and, on $(-\infty, 0)$, negative for n odd and positive for n even. By the induction hypotheses and the observation that $f'_{n+1}(x) = (n+1)f_n(x)$, it follows that $f_{k+1}(x)$ has a unique nonzero root r_{k+1} with $r_k < r_{k+1}$ and that the additional sign conditions hold. The proofs of (a) and (b) are complete.

(c) We show $\lim_{n\to\infty} r_n/n = 1/\ln[(1+\sqrt{5})/2] \approx 2.078$. For a fixed $u > 0$,

$$\lim_{n\to\infty}\frac{f_n(un)}{(un)^n} = \lim_{n\to\infty}\left[1 + \left(1 - \frac{1}{un}\right)^n - \left(1 + \frac{1}{un}\right)^n\right] = 1 + e^{(-1/u)} - e^{(1/u)},$$

and straightforward calculations show

$$1 + e^{(-1/u)} - e^{(1/u)} \begin{cases} < 0 & \text{if } u < 1/\ln\left[(1+\sqrt{5})/2\right], \\ = 0 & \text{if } u = 1/\ln\left[(1+\sqrt{5})/2\right], \\ > 0 & \text{if } u > 1/\ln\left[(1+\sqrt{5})/2\right]. \end{cases}$$

All this implies that for large n, if $u < 1/\ln[(1+\sqrt{5})/2]$ then $f_n(un) < 0$ and $r_n > un$. On the other hand if $u > 1/\ln[(1+\sqrt{5})/2]$ then $f_n(un) > 0$ and $r_n < un$. Thus $\lim_{n\to\infty} r_n/n = 1/\ln[(1+\sqrt{5})/2]$.

Also solved by Paul Bracken (Canada), Charles Diminnie and Roger Zarnowski, Daniele Donini (Italy), Danrun Huang, Thomas Jager, Kee-Wei Lau (China), Eugene Lee, Can A. Minh, William A. Newcomb, Gao Peng, TAMUK Problem Solvers, and the proposer. There were 21 incorrect solutions. Most errors were in part (c). Although many answers were correct, these solutions were invalid due to either an a priori assumption of the existence of the limit or an incorrect way of computing the limit.

The Limit of the nth Root of a Difference **February 1998**

1541. *Proposed by Wu Wei Chao, He Nan Normal University, Xin Xiang City, He Nan Province, China.*

Assume $a_1 > 1$ and define $a_{n+1} = 1/a_n + a_1 - 1$ for $n = 1, 2, 3, \ldots$. Evaluate

$$\lim_{n\to\infty} |a_{n+1} - a_n|^{1/n}.$$

I. *Solution by Evgenii S. Freidkin, Middlesex County College, Edison, New Jersey.*

We will show that

$$\lim_{n\to\infty} |a_{n+1} - a_n|^{1/n} = \frac{(a_1-1)^2 + 2 - (a_1-1)\sqrt{(a_1-1)^2+4}}{2}.$$

Clearly $1 < a_2 < a_1$ and an easy induction shows that $1 < a_n < a_1$ for all $n > 1$. Set

$$r = \frac{(a_1-1) + \sqrt{(a_1-1)^2+4}}{2} > 1,$$

which satisfies the equation $r = 1/r + a_1 - 1$. Then, from the recursion,

$$|a_{n+1} - r| = \left|\frac{1}{a_n} - \frac{1}{r}\right| < \frac{|a_n - r|}{r} < \cdots < \frac{|a_1 - r|}{r^n}.$$

Therefore, $\lim_{n\to\infty} a_n = r$.

A similar calculation from the recursion yields

$$|a_{n+1} - a_n| = \frac{|a_n - a_{n-1}|}{a_n a_{n-1}} = \cdots = \frac{|a_1 - 1|}{a_n a_{n-1}^2 \cdots a_2^2 a_1^2}.$$

It follows that

$$\lim_{n\to\infty} |a_{n+1} - a_n|^{1/n} = \frac{1}{r^2} = \frac{(a_1-1)^2 + 2 - (a_1-1)\sqrt{(a_1-1)^2+4}}{2}.$$

II. *Solution by J. C. Binz, University of Bern, Bern, Switzerland.*

Put $b_0 = 1$, $b_1 = a_1$, and $b_{n+1} = (a_1 - 1)b_n + b_{n-1}$ for $n \geq 1$. We prove that $b_n / b_{n-1} = a_n$ by induction, the main step being

$$\frac{b_{n+1}}{b_n} = \frac{(a_1 - 1)b_n + b_{n-1}}{b_n} = a_1 - 1 + \frac{1}{a_n} = a_{n+1}.$$

The recursion for b_n has characteristic equation $r^2 - (a_1 - 1)r - 1 = 0$ with roots

$$\frac{(a_1 - 1) \pm \sqrt{(a_1 - 1)^2 + 4}}{2},$$

which we denote by r_+ and r_-, with $r_+ > 1$ and $|r_-| < 1$. Therefore, $b_n = \alpha_+ r_+^n + \alpha_- r_-^n$, where the initial conditions yield $\alpha_+ = (r_+ + 1)/(r_+ - r_-) \neq 0$ and $\alpha_- = (r_- + 1)/(r_- - r_+)$. Note that $\lim_{n \to \infty} |b_n|^{1/n} = r_+$. Now

$$a_{n+1} - a_n = \frac{b_{n+1}b_{n-1}b_n^2}{b_n b_{n-1}} = \frac{(-1)^n (a_1 - 1)}{b_n b_{n-1}}.$$

Finally,

$$\lim_{n \to \infty} |a_{n+1} - a_n|^{1/n} = \lim_{n \to \infty} \frac{(a_1 - 1)^{1/n}}{|b_n|^{1/n} |b_{n-1}|^{1/n}} = \frac{1}{r_+^2} = r_-^2$$

$$= \frac{(a_1 - 1)^2 + 2 - (a_1 - 1)\sqrt{(a_1 - 1)^2 + 4}}{2}.$$

Also solved by Matt Baker (graduate student), Jean Bogaert (Belgium), Paul Bracken (Canada), Daniele Donini (Italy), Frank A. Horrigan, Thomas Jager, Hans Kappus (Switzerland), LC^U, Eugene Lee, Can A. Minh, William A. Newcomb, Gao Peng, Heinz-Jürgen Seiffert (Germany), Nicholas C. Singer, TAMUK Problem Solvers, Michael Vowe (Switzerland), Western Maryland College Problems Group, and the proposer. There was one incorrect solution.

An Infinite Game on the Real Line **February 1998**

1542. *Proposed by Jerrold W. Grossman and Barry Turett, Oakland University, Rochester, Michigan.*

Sam and Joe (names favored by the late Paul Erdős) play an infinite game on the real number line. They start at distinct initial positions and alternate turns. At each turn a player must move to some point strictly between the players' current positions. Being monotonic and bounded, the sequence of positions for each player converges. A player wins the game if his limit is rational and loses if his limit is irrational.

(a) Show that Joe can force Sam to lose.
(b) Find a strategy by which Joe will win with probability 1 if Sam plays randomly (i.e., at each turn, Sam chooses a point in the gap between the players, independent of previous choices, based on the uniform distribution).
(c) Does the result in (a) hold if the winning set is an arbitrary set of measure zero?

(Obviously they can play cooperatively and end up with a win/win situation. Furthermore, either player can unilaterally guarantee that the results for both players are identical by forcing the gap between them to vanish.)

Solution by William A. Newcomb, Walnut Creek, California.

Let s_n and j_n be the nth numbers chosen by Sam and Joe, respectively, and let S and J be their respective limits. Let Sam choose first. Then, in parts (a) and (b), we can assume for definiteness $j_1 > s_1$. (Joe has also, of course, a successful strategy with $j_1 < s_1$, involving merely the reversal of inequality signs in the obvious places.)

(a) Observe first that $s_k < S < j_n$ for all positive integers k and n. Now let $(r_n)_{n\geq 1}$ be an enumeration of all the rationals. Joe chooses $j_n = r_n$ if available and otherwise arbitrarily. Then, in either case, r_n is outside the interval (s_n, j_n) and is therefore not equal to S. As this is true for every n, S is necessarily irrational.

(b) For some specific rational $r > s_1$, Joe will at first aim to produce $J = r$. To this end, he chooses first

$$j_1 = r + \frac{r - s_1}{4}.$$

At his nth turn, if it is still true that $s_n < r$, he will choose

$$j_n = r + \frac{r - s_n}{2^{n+1}} < j_{n-1}.$$

Given that $s_n < r$, the probability that also $s_{n+1} < r$ is greater than $1 - 2^{-n-1}$. Then with probability P, where

$$P > \left(1 - \tfrac{1}{4}\right)\left(1 - \tfrac{1}{8}\right)\left(1 - \tfrac{1}{16}\right)\cdots > 1 - \left(\frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots\right) = \frac{1}{2} > 0,$$

s_n will be less than r for all n, and Joe will have succeeded in forcing $J = r$. If, on the other hand, it happens at any stage that $s_n \geq r$, thus forcing $J > r$, then Joe simply starts over with a new value of r, again with the same nonzero probability P of success. The probability of success in not more than m such attempts is $1 - (1 - P)^m$, which approaches 1 as m approaches ∞.

(c) The answer is no. As will be shown, Sam can always force $S \in C$, where C is the Cantor set, which has measure zero. Recall the definition of the Cantor set as the set of all real numbers that have ternary expansions consisting entirely of zeros and twos, and let C_0 be the subset of C for which the expansion contains both infinitely many zeros and infinitely many twos. Now every point $s \in C_0$ has the property that every open interval with one (left or right) endpoint s contains a member of C_0, so Sam can always choose a member of C_0 at every turn, which of course is also a member of C. But C is a closed set, so the limit point S also is a member of C.

Also solved by Scott Metcalf and the proposers. There was one incorrect solution.

Extrema of Volumes of Truncated Simplexes **February 1998**

1543. *Proposed by Michael Golomb, Purdue University, West Lafayette, Indiana.*

Let S be a given n-dimensional simplex with centroid C. A hyperplane through C divides the simplex into two regions, one or both of which are simplexes. Find the extrema of the volumes of those regions which are simplexes.

Solution by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.

We will show that the maximum volume is $\text{vol}(S)/2$ and the minimum volume is $[n/(n+1)]^n\ \text{vol}(S)$, where $\text{vol}(S)$ denotes the volume of S.

Let $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n$ and $\mathbf{C}$ denote vectors from one vertex to all other vertices of S and C, respectively. Let a hyperplane through C cut these vectors at points given by $x_1\mathbf{V}_1, x_2\mathbf{V}_2, \ldots, x_n\mathbf{V}_n$, respectively, where the x_i's lie in $[0,1]$. Then, since $\mathbf{C} = \sum_{i=1}^n \mathbf{V}_i/(n+1)$ lies the hyperplane, there are nonnegative weights $w_1, w_2, \ldots, w_n$, with sum 1, such that $\sum_{i=1}^n w_i x_i \mathbf{V}_i = \sum_{i=1}^n \mathbf{V}_i/(n+1)$. Since the $\mathbf{V}_i$'s are linearly independent, $w_i x_i = 1/(n+1)$, so that $\sum_{i=1}^n 1/x_i = n+1$. Since the volume cut off the simplex is

$$\operatorname{vol}(S)\prod_{i=1}^n x_i = \operatorname{vol}(S)\Big/\prod_{i=1}^n 1/x_i,$$

we want to determine the extrema of the latter product subject to the constraints $\sum_{i=1}^n 1/x_i = n+1$ and $x_i \in [0,1]$.

The arithmetic mean-geometric mean inequality implies

$$\operatorname{vol}(S)\prod_{i=1}^n x_i \geq \operatorname{vol}(S)\Big/\left(\frac{1}{n}\sum_{i=1}^n 1/x_i\right)^n = \operatorname{vol}(S)\left[n/(n+1)\right]^n,$$

so that the minimum volume occurs when $x_i = n/(n+1)$ for all i.

To obtain the maximum volume, we let $1/x_i = 1 + y_i$, so that we now want to minimize $\prod_{i=1}^n (1+y_i)$ subject to $\sum_{i=1}^n y_i = 1$. Expanding the product, we see that

$$\prod_{i=1}^n (1+y_i) \geq 1 + \sum_{i=1}^n y_i = 2,$$

with equality if and only if one y_i is 1 and the rest are 0. This yields a maximum volume of $\operatorname{vol}(S)/2$.

Also solved by William A. Newcomb and the proposer.

Answers

Solutions to the Quickies on page 65.

A887. Since both sides of the inequality are even functions, we may assume that $0 \leq x < \pi/2$. Observe

$$\sin x = \int_0^x \cos t\,dt, \qquad \frac{1}{2}\ln\left(\frac{1+\sin x}{1-\sin x}\right) = \int_0^x \sec t\,dt.$$

Apply the Cauchy–Schwarz inequality, $\int_I f^2 \cdot \int_I g^2 \geq (\int_I fg)^2$, with $I = [0,x]$, $f(x) = \sqrt{\cos x}$, and $g(x) = \sqrt{\sec x}$ to obtain the asserted inequality.

A888. A team will be the first to win n games if and only if it would win a majority of $2n-1$ games (playing all games). Increasing the series from $2n-1$ to $2n+1$ games can change the winner of the series if and only if one team wins exactly n of the first $2n-1$ games and then loses the last two. Setting $q = 1-p$, the probability of the better team winning the series has changed by

$$\binom{2n-1}{n-1}p^{n-1}q^n \cdot p^2 - \binom{2n-1}{n-1}p^n q^{n-1}\cdot q^2 = \binom{2n-1}{n-1}p^n q^n (p-q) > 0$$

for $p > 1/2$.

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 mathematics literature. Readers are invited to suggest items for review to the editors.

Brown, Malcolm W., Ancient Archimedes text turns up, and it's for sale, *New York Times* (27 October 1998) A5 (City Ed.), D5 (Ntl Ed.). Archimedes text sold for $2 million, (30 October 1998) A27 (City Ed.).

The sole manuscript containing Archimedes' *The Method* has been sold at auction to an undisclosed private American collector (not Bill Gates) for $2 million. One of the most significant discoveries of the twentieth century in the history of mathematics, the manuscript describes how Archimedes was inspired by ideas from mechanics and the use of infinitesimals to discover mathematical theorems (which he then had to prove by geometric means that were considered rigorous). The manuscript was found in Constantinople in 1907 in palimpsest form—that is, the texts of Archimedes' work had been washed off (but not completely) so that the vellum could be reinked with other text. The Archimedes text dates from the 10th century; somehow the volume passed into private hands after World War I (it was stolen, says the Greek Patriarch of Jerusalem, who has sued for its return). The volume also contains the only Greek version of *On Floating Bodies* and a fragment of an otherwise lost work; it has not been available for inspection by scholars since 1909. It's too bad that the American Mathematical Society and/or the MAA didn't enter the auction, and that no American mathematician was rich enough to buy the work and make it available for research.

Nahin, Paul J., *An Imaginary Tale: The Story of* $\sqrt{-1}$, Princeton University Press, 1998; xvi + 257 pp, $24.95. ISBN 0-691-02795-1.

We mathematicians often seem at a loss for words—for apt words for mathematical concepts. We overwork some nondescriptive words ("regular," "normal") and employ common words but with discipline-specific meanings; a result is to make it harder for non-mathematicians to understand what we are talking about. Take, for example, the course title "Complex Analysis"; to a non-mathematician, this must be more involved and even harder than whatever regular "analysis" may be. The longer title, "Analysis of Complex Numbers," would connect with the person's experience of $i = \sqrt{-1}$ but begs the unhappy question of what is complex about complex numbers or imaginary about imaginary ones. Author Nahin relates that prior to Descartes (to whom we owe the terms "real number" and "imaginary number"), square roots of negative numbers were called "sophisticated" or "subtle" (did Descartes improve the situation?), while it was Gauss who coined the term "complex number." Nahin's captivating book sets out in detail all the historical steps in understanding the arithmetic, algebra, geometry, and applications of $\sqrt{-1}$, culminating in the derivation of Kepler's Third Law through use of the Cauchy integral formulas. For full appreciation, the reader needs to know calculus; but college algebra and a tolerance for symbols will carry the reader through the first half of the book.

Baxter, Martin, and Andrew Rennie, *Financial Calculus: An Introduction to Derivative Pricing*, Cambridge University Press, 1996. ISBN 0–521–52289–3. Wilmott, Paul, Sam Howison, and Jeff Dewynne, *The Mathematics of Financial Derivatives*, Cambridge University Press, 1995; xiii + 317 pp, $49.95, $24.95 (P). ISBN 0–521–49699–3, 0–521–49789–2.

Interest in mathematics of finance was greatly heightened by the award of the 1997 Nobel Prize in Economics to Robert C. Merton (Harvard University) and Myron S. Scholes (Stanford University) for a method (the Black-Scholes formula) of valuing financial derivatives. (The later involvement of the winners in a hedge fund debacle probably won't lessen the attraction of students to finance.) Here are two potential textbooks for learning about the mathematics of pricing financial derivatives. The two differ on what the fundamental mathematics is: Wilmott et al. stress partial differential equations (linear parabolic second order, like the heat equation or the diffusion equation), while Baxter and Rennie put central focus on martingales—a topic not even mentioned by Wilmott et al. Both texts, written by professors at British universities, disingenuously soft-pedal prerequisites with phrasing that begs for commentary: "[no] particular prior body of knowledge, except for some (classical) differential calculus and experience with symbolic notation" (Wilmott et al.) [is it a good idea to try to learn applied mathematics without knowing anything about the area of application?]; "the early calculus, probability and algebra courses of an undergraduate degree or equivalent in mathematics, chemistry, engineering or similar subjects" (Baxter and Rennie) [so engineers are qualified to learn how options pricing works but economics students aren't sufficiently trained]. In fact, courses in mathematical probability and in differential equations would be useful background. Both texts offer exercises but Wilmott et al. offer a greater number. The wise policy, of course, would be to adopt *both* texts and also require reading in the less-mathematical ("more practically oriented") books on options pricing that are studied by students in economics (see p. x of Wilmott et al.).

Csicsery, George Paul, *N Is a Number: A Portrait of Paul Erdős*, color film, 57 min, 1993. Available from Csicsery (check/purchase order), P.O. Box 2833, Oakland, CA 94618, (510) 428–9284, `75430.3310@compuserve.com` , or A K Peters (credit card). 16 mm print $2700; VHS (NTSC) cassette with performance rights $200, home use $42. ISBN 0–933621–62–0.

"He had to have this life, to spend 24 hours a day doing what he wants," says Vera Śoś in this fine film about the human side of about Paul Erdős. Shot in a variety of locations over a period of several years, the film's anecdotes show him telling jokes in his lectures, interacting with mathematicians, playing with children, conversing about social problems and international affairs, and playing remarkably good table tennis for an 80-year-old. His story is told in part through engaging interviews ,with Ron Graham, Vera Śoś, Joel Spencer, and others, who describe a life spare in attachment to material things but rich in mathematics and friends. (December 1998 *Discover* lists this among top 10 science videos of the past 20 years.)

Kidwell, Peggy Aldrich, Stalking the elusive computer bug, *IEEE Annals of the History of Computing* 20 (4) (1998) 5–9.

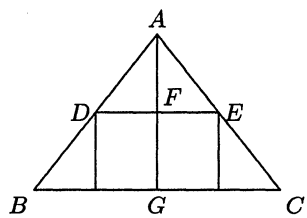
This article summarizes what is known about the history of the term "bug" as used in computer science. It includes the anecdote and a photograph of the 1947 logbook in which Grace Hopper taped a dead moth found in a computer relay and wrote "First actual case of bug being found." What may be surprising, however, is that the term "bug" was already in use by U.S. telegraph operators by 1875, who used it to refer to a false signal. Thomas Edison, a former telegrapher, extended the meaning to refer to "little faults and difficulties" that occur in the course of developing an invention, "implying that some imaginary insect has secreted itself inside and is causing all the trouble." The terms "bug" and "debugging" were used in reports by Hopper several years before she found the famous moth.

NEWS AND LETTERS

59th Annual William Lowell Putnam Mathematical Competition

A-1 A right circular cone has base of radius 1 and height 3. A cube is inscribed in the cone so that one face of the cube is contained in the base of the cone. What is the side-length of the cube?

Solution. Consider a plane cross-section through a vertex of the cube and the axis of the cone as shown.

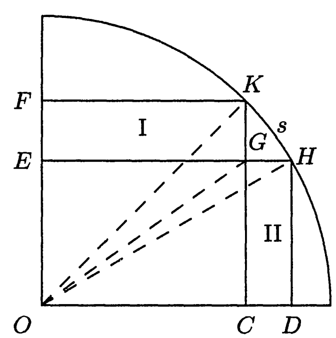


Let s be the side of the cube. Segment DE is a diagonal of the top of the cube and has length $\sqrt{2}\,s$. By similarity of triangles ADE and ABC, we have $\frac{DE}{BC} = \frac{AF}{AG}$, or $\frac{\sqrt{2}\,s}{2} = \frac{3-s}{3}$. Hence

$$s = \frac{6}{3\sqrt{2}+2} = \frac{3\sqrt{2}}{\sqrt{2}+3} = \frac{9\sqrt{2}-6}{7}.$$

A-2 Let s be any arc of the unit circle lying entirely in the first quadrant. Let A be the area of the region lying below s and above the x-axis and let B be the area of the region lying to the right of the y-axis and to the left of s. Prove that $A+B$ depends only on the arc length, and not on the position, of s.

Solution. Let I be the area in rectangle *EGKF* and let II be the area in rectangle *CDHG.*



Clearly, I $= 2 \cdot \text{Area}\Delta OGK$, and II $= 2 \cdot \text{Area}\Delta OGH$. Thus,

$$\begin{aligned} A+B &= \text{I} + \text{II} + 2\,\text{Area}\,GHK \\ &= 2\,(\text{Area}\,\Delta OGK + \text{Area}\,\Delta OGH + \text{Area}\,GHK) \\ &= 2\,(\text{Area Sector}\,OHK) = \theta, \end{aligned}$$

where θ is the length of arc s.

A-3 Let f be a real function on the real line with continuous third derivative. Prove that there exists a point a such that

$$f(a) \cdot f'(a) \cdot f''(a) \cdot f'''(a) \geq 0.$$

Solution. Otherwise, by the Intermediate Value Theorem, $f(x)$ and its first three derivatives cannot change sign. A positive function on $(-\infty, \infty)$ cannot be strictly concave nor a negative function strictly convex, so $f(x)f''(x)$ is a positive function. Applying the same reasoning to $f'(x)$, $f'(x)f'''(x)$ is positive, and the product of two positive functions is positive.

A-4 Let $A_1 = 0$ and $A_2 = 1$. For $n > 2$, the number A_n is defined by concatenating the decimal expansions of A_{n-1} and A_{n-2} from left to right. For example, $A_3 = A_2A_1 = 10$, $A_4 = A_3A_2 = 101$, $A_5 = A_4A_3 = 10110$, and so forth. Determine all n such that 11 divides A_n.

Solution. A_n is divisible by 11 precisely when n is one more than a multiple of 6. We know that a number is divisible by 11 if and only if the alternating sum of its digits is zero. So, let d_n be the alternating sum of the digits in A_n starting from the left. Then $d_n = d_{n-1} + (-1)^{r_{n-1}} d_{n-2}$, where r_n is the number of digits in A_n. Clearly, $r_n = F_n$, the nth Fibonacci number. The Fibonacci numbers have parity pattern $\left((-1)^{F_n}\right) : -1, -1, 1, -1, -1, 1, \ldots$. Thus, the rules for calculating d_n are $d_1 = 0$, $d_2 = 1$, $d_3 = d_2 - d_1 = 1$, $d_4 = d_3 + d_2 = 2$, $d_5 = d_4 - d_3 = 1$, $d_6 = d_5 - d_4 = -1$, $d_7 = d_6 + d_5 = 0$, $d_8 = d_7 - d_6 = 1$, and so forth in this repeating pattern. This completes the proof.

A-5 Let $\mathcal{F}$ be a finite collection of open discs in $\mathbb{R}^2$ whose union contains a set $E \subseteq \mathbb{R}^2$. Show that there is a pairwise disjoint subcollection $D_1, \ldots, D_n$ in $\mathcal{F}$ such that

$$\bigcup_{j=1}^{n} 3D_j \supseteq E.$$

Here, if D is the disc of radius r and center P, then $3D$ is the disc of radius $3r$ and center P.

Solution. It suffices to prove the result when $E = \bigcup \mathcal{F}$. For this, we induct on the number of sets in $\mathcal{F}$. The result obviously holds when $|\mathcal{F}| = 1$, so assume the result is true whenever $|\mathcal{F}| < m$, and consider a family $\mathcal{F}$ with $|\mathcal{F}| = m$. Let F denote a disc in $\mathcal{F}$ with the smallest radius. Then, by induction, there is a pairwise disjoint subcollection $D_1, D_2, \ldots, D_n$ in $\mathcal{F}' = \mathcal{F} \setminus \{F\}$ such that $\bigcup_{j=1}^{n} 3D_j \supseteq \bigcup \mathcal{F}'$. Suppose $F \cap D_i \neq \emptyset$ for some i. Then, using the triangle inequality, it is easy to show that $3D_i \supseteq F$, so $\bigcup_{j=1}^{n} 3D_j \supseteq \bigcup \mathcal{F}$. Otherwise, $D_1, D_2, \ldots, D_n, F$ is a disjoint subcollection in $\mathcal{F}$ such that $\bigcup_{i=1}^{n} D_i \cup F \supseteq \bigcup \mathcal{F}$, and we are done.

A-6 Let A, B, C denote distinct points with integer coordinates in $\mathbb{R}^2$. Prove that if

$$\Big(|AB| + |BC|\Big)^2 < 8 \cdot [ABC] + 1,$$

then A, B, C are three vertices of a square. Here $|XY|$ is the length of segment XY and $[ABC]$ is the area of triangle ABC.

Solution. The set $\mathbb{Z}^2$ is preserved by 90° rotation about any of its points. Therefore, there exists a point C' in $\mathbb{Z}^2$ with $|BC'| = |BA|$, BC' perpendicular to BA such that C and C' belong to the same half-plane with respect to the line through A and B. If $C \neq C'$, then $|CC'| = s \geq 1$. Let $r = |AB|$. By rotating and translating the coordinate system, we may assume that $B = (0,0), A = (r,0), C' = (0,r)$, and $C = (0,r) + s(\cos\theta, \sin\theta)$, where θ is the angle to the line $C'C$ measured from the positive horizontal axis. Thus

$$\begin{aligned}
(|AB| + |BC|)^2 &= \Big(r + \sqrt{s^2\cos^2\theta + (r + s\,\sin\theta)^2}\Big)^2 \\
&= 2r^2 + s^2 + 2rs\,\sin\theta + 2r\sqrt{s^2\cos^2\theta + (r + s\,\sin\theta)^2} \\
&\geq 2r^2 + 1 + 2rs\,\sin\theta + 2r(r + s\,\sin\theta) \\
&= 8\,\frac{r(r + s\,\sin\theta)}{2} + 1 \\
&= 8\,[ABC] + 1.
\end{aligned}$$

B-1 Find the minimum value of

$$\frac{(x + 1/x)^6 - (x^6 + 1/x^6) - 2}{(x + 1/x)^3 + (x^3 + 1/x^3)}$$

for $x > 0$.

Solution. Write

$$\begin{aligned}
\frac{(x + 1/x)^6 - (x^6 + 1/x^6) - 2}{(x + 1/x)^3 + (x^3 + 1/x^3)} &= \frac{\big((x + 1/x)^3 - (x^3 + 1/x^3)\big)\big((x + 1/x)^3 + (x^3 + 1/x^3)\big)}{(x + 1/x)^3 + (x^3 + 1/x^3)} \\
&= (x + 1/x)^3 - (x^3 + 1/x^3) \\
&= 3x^2(1/x) + 3x(1/x)^2 \\
&= 3x + 3/x\,.
\end{aligned}$$

The function $f(x) = 3(x + 1/x)$ tends to $+\infty$ as $x \to 0^+$ or $x \to +\infty$. Also, $f'(x) = 3 - 3/x^2 = 0$ when $x = \pm 1$. Only $x = 1$ is relevant. Since $f''(x) = 6/x^3$ and $f''(1) > 0$, it follows that $x = 1$ is the minimum and $f(1) = 6$ is the minimum value.

B-2 Given a point (a, b) with $0 < b < a$, determine the minimum perimeter of a triangle with one vertex at (a, b), one on the x-axis, and one on the line $y = x$. You may assume that a triangle of minimum perimeter exists.

Solution. Let θ be the angle between the positive x-axis and the line $y = mx$. By two reflections (the first about the line $y = mx$, the next about the reflection of the x-axis across $y = mx$), the shortest perimeter for such a triangle is seen to correspond to the straight line from (a, b) to the rotation of (a, b) by angle 2θ clockwise. The minimum perimeter is therefore, by the Law of Cosines, $\sqrt{2(a^2 + b^2)(1 - \cos 2\theta)}$. For $m = 1$, this is $\sqrt{2(a^2 + b^2)}$.

B-3 Let H be the unit hemisphere $\{(x, y, z) : x^2 + y^2 + z^2 = 1, z \geq 0\}$, C the unit circle $\{(x, y, 0) : x^2 + y^2 = 1\}$, and P the regular pentagon inscribed in C. Determine the surface area of that portion of H lying over the planar region inside P, and write your answer in the form $A \sin \alpha + B \cos \beta$, where A, B, α, and β are real numbers.

Solution. Observe that twice the desired surface area equals the surface area of the whole sphere minus 5 spherical caps subtended by a "chord" with central angle $2\pi/5$. The surface area of a spherical cap of angle α can be computed in spherical coordinates by

$$\int_0^{\alpha/2} \int_0^{2\pi} \sin \phi \, d\theta \, d\phi = 2\pi(1 - \cos(\alpha/2)).$$

Thus, the desired area is

$$\frac{1}{2}\Big(4\pi - 5(2\pi(1 - \cos(\pi/5)))\Big) = 5\pi \cos(\pi/5) - 3\pi.$$

B-4 Find necessary and sufficient conditions on positive integers m and n so that

$$\sum_{i=0}^{mn-1} (-1)^{\lfloor i/m \rfloor + \lfloor i/n \rfloor} = 0.$$

Solution. If m and n are both odd, the sum has an odd number mn of terms all of which are ± 1, and therefore the sum cannot be 0. We shall prove the sum is *not* 0 if and only if $m = 2^k m'$ and $n = 2^k n'$, for an integer $k \geq 1$ and some odd integers m', n'.

LEMMA. $\left\lfloor \dfrac{mn - 1 - i}{m} \right\rfloor = n - 1 - \left\lfloor \dfrac{i}{m} \right\rfloor.$

Proof. If $m = 1$, this is clear. Assume $m \geq 2$. Suppose $k \leq i/m < k + 1$. Then

$$\frac{mn-1}{m} - (k+1) < \frac{mn-1-i}{m} \leq \frac{mn-1}{m} - k$$
$$n - k - 1 - \frac{1}{m} < \frac{mn-1-i}{m} \leq n - k - \frac{1}{m}$$

Since $\dfrac{mn-1-i}{m}$ is a multiple of $1/m < 1$, we conclude:

$$n - k - 1 \leq \frac{mn-1-i}{m} < n - k,$$

which is the statement of the lemma.

Denote the given sum by $S_{m,n}$.

Substitute $mn-1-i$ for i in the sum. Then

$$\begin{aligned} S_{m,n} &= \sum_{i=0}^{mn-1} (-1)^{\lfloor \frac{mn-1-i}{m} \rfloor + \lfloor \frac{mn-1-i}{n} \rfloor} \\ &= \sum_{i=0}^{mn-1} (-1)^{n-1-\lfloor \frac{i}{m} \rfloor + m-1-\lfloor \frac{i}{n} \rfloor} \\ &= (-1)^{m+n} S_{m,n}, \end{aligned}$$

using the lemma.

COROLLARY. $S_{m,n} = 0$ if $m+n$ is odd.

That leaves only the case where m, n are both even to consider. Then

$$\begin{aligned} S_{2m,2n} &= \sum_{i=0}^{4mn-1} (-1)^{\lfloor \frac{i}{2m} \rfloor + \lfloor \frac{i}{2n} \rfloor} \\ &= \sum_{i=0}^{2mn-1} (-1)^{\lfloor \frac{i}{2m} \rfloor + \lfloor \frac{i}{2n} \rfloor} + \sum_{i=0}^{2mn-1} (-1)^{2m-1-\lfloor \frac{i}{2m} \rfloor + 2n-1-\frac{i}{2n}}, \end{aligned}$$

after substituting $2mn-1-i$ for i in the second half. Hence,

$$\begin{aligned} S_{2m,2n} &= 2 \sum_{i=0}^{2mn-1} (-1)^{\lfloor \frac{i}{2m} \rfloor + \lfloor \frac{i}{2n} \rfloor} \\ &= 2 \sum_{i=0}^{mn-1} \left((-1)^{\lfloor \frac{2i}{2m} \rfloor + \lfloor \frac{2i}{2n} \rfloor} + (-1)^{\lfloor \frac{2i+1}{2m} \rfloor + \lfloor \frac{2i+1}{2n} \rfloor} \right) \\ &= 4 S_{m,n}, \end{aligned}$$

since, clearly, $\left\lfloor \dfrac{2i}{2m} \right\rfloor = \left\lfloor \dfrac{2i+1}{2m} \right\rfloor$.

Repeatedly applying this result proves $S_{2^k m', 2^k n'} = 4^k S_{m',n'}$, which establishes our claim.

B-5 Let N be the positive integer with 1998 decimal digits, all of them 1; that is,

$$N = \underbrace{1111\cdots11}_{1998 \text{ digits}}.$$

Find the thousandth digit after the decimal point of $\sqrt{N}$.

Solution. (By a student contestant) Observe that

$$10^{3998} - 12 \cdot 10^{1998} + 36 < 10^{3998} - 10^{2000} < 3^{3998} - 10^{2000} + 25.$$

This leads to the following chain of inequalities:

$$(10^{1999}-6)^2 < 10^{2000}(10^{1998}-1) < (10^{1999}-5)^2$$

$$\left(\frac{10^{1999}-6}{3}\right)^2 < 10^{2000}\left(\frac{10^{1998}-1}{9}\right) = 10^{2000}N < \left(\frac{10^{1999}-5}{3}\right)^2$$

$$\frac{10^{1999}-1}{3}-2 < \frac{10^{1999}-6}{3} < 10^{1000}\sqrt{N} < \frac{10^{1999}-5}{3} < \frac{10^{1999}-1}{3}-1$$

$$\underbrace{333\cdots33}_{1998}1 < 10^{1000}\sqrt{N} < \underbrace{333\cdots33}_{1998}2.$$

Hence the thousandth digit after the decimal point of $\sqrt{N}$, which is equal to the units digit of $10^{1000}\sqrt{N}$, is 1.

B-6 Prove that, for any integers a, b, c, there exists a positive integer n such that $\sqrt{n^3+an^2+bn+c}$ is not an integer.

Solution. (By Byron Walden, University of Santa Clara) Let $s_n = n^3+an^2+bn+c$. Then

$$\begin{aligned} s_1 &\equiv 1+a+b+c \pmod 4\\ s_2 &\equiv 2b+c \pmod 4\\ s_3 &\equiv -1+a-b+c \pmod 4\\ s_4 &\equiv c \pmod 4\end{aligned}$$

So

$$s_2-s_4 \equiv 2b \pmod 4 \quad\text{and}\quad s_1-s_3 \equiv 2b+2 \pmod 4.$$

Thus, either

$$s_2-s_4 \equiv 2 \pmod 4 \quad\text{or}\quad s_1-s_3 \equiv 2 \pmod 4.$$

Because squares are congruent either to 0 or to 1 (mod 4), not both s_2-s_4 and s_1-s_3 can be differences of squares (mod 4). Thus at least one of s_1, s_2, s_3, s_4 is not a square.

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

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