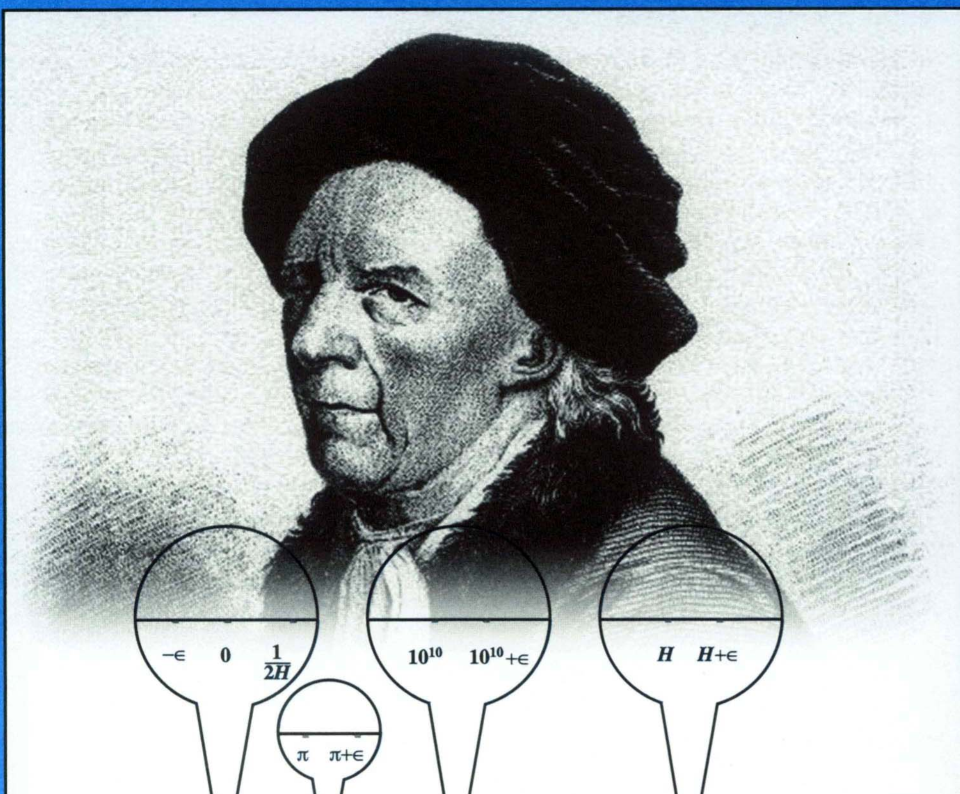
Vol. 74, No. 5, December 2001

# MATHEMATICS MAGAZINE




- Higher Trigonometry, Hyperreal Numbers, and Euler's Analysis of Infinities
- Smullyan's Vizier Problem

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

## EDITORIAL POLICY

*Mathematics Magazine* aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate 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 among various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 74, pp. 75–76, and is available from the Editor or at www.maa.org/pubs/mathmag.html. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Submit new manuscripts to Frank A. Farris, Editor, Mathematics Magazine, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053-0373. Manuscripts should be laser printed, with wide line spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should mail three copies and keep one copy. In addition, authors should supply the full five-symbol 2000 Mathematics Subject Classification number, as described in *Mathematical Reviews*.

Cover image: Euler Surveys the Hyperreal Line, by Jason Challas, who lectures on computer art and other hyperreal topics in the Department of Art and Art History at Santa Clara University.

## AUTHORS

**Mark McKinzie** teaches in the Mathematics Department at Monroe Community College in Rochester, New York. He has a Ph.D. in mathematics from the University of Wisconsin; his dissertation, on the early history of power series, was completed under the supervision of M. Bleicher. Work on this paper started while he was an intern in the Information Sciences Division of Bell Laboratories, and sparked his interest in the history of mathematics in general, and of series in particular.

**Curtis Tuckey** is director of the Oracle Voice Laboratory. Before joining Oracle Corporation he held various research and development positions at Motorola, Lucent Technologies, AT&T, and General Motors. He has occasionally taught at Loyola, DePaul, and Northwestern. This paper was written while he was a research member of the Information Sciences Division of Bell Laboratories. He has a Ph.D. in mathematics from the University of Wisconsin, where he wrote a dissertation in nonstandard analysis under H. J. Keisler. He lives in Chicago.

**Michael Khoury, Jr.** is currently a mathematics and education major at Denison University. He has spent two summers doing research there, and topics have included algebraic number theory and infinite series. His other pursuits include foreign languages and creative writing. The content of this article is the fruit of his high school days, which were largely spent with books of puzzles and problems in hand; he thanks his professors for encouraging him to write this paper. Asked about his own knighthood or knavery, he responded, "Two is prime."

Vol. 74, No. 5, December 2001

# MATHEMATICS MAGAZINE

*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 *MATHEMATICS MAGAZINE* to an individual member of the Association is $25 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $149. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount.) The nonmember/library subscription price is $300 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 Dave Riska (*driska@maa.org*), 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

# Higher Trigonometry, Hyperreal Numbers, and Euler's Analysis of Infinities

MARK MCKINZIE
Monroe Community College
Rochester, NY 14623

CURTIS TUCKEY
1217 W. Arthur Avenue
Chicago, IL 60626

In a textbook published in 1748, without the barest mention of the derivative, Euler derived the fundamental equations of a subject that was later to become known as *higher trigonometry*: he explained the series for the exponential and logarithmic functions,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$
$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots,$$

proved the Euler identity,

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

computed the series for the sine and cosine,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots,$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots,$$

proved the factorization formula for the sine,

$$\sin x = x\left(1 - \frac{x^2}{(1\pi)^2}\right)\left(1 - \frac{x^2}{(2\pi)^2}\right)\left(1 - \frac{x^2}{(3\pi)^2}\right)\cdots,$$

and deduced his celebrated formula,

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots = \frac{\pi^2}{6},$$

among many other facts. The textbook is Euler's *Introductio in Analysin Infinitorum* (Introduction to the Analysis of Infinities). "All this follows from ordinary algebra," he claimed, and all this in a textbook for beginners!

> Often I have considered the fact that most of the difficulties which block the progress of students trying to learn analysis stem from this: that although they understand little of ordinary algebra, still they attempt this more subtle art. From this it follows not only that they remain on the fringes, but in addition they entertain strange ideas about the concept of the infinite, which they must try to use. Although analysis does not require an exhaustive knowledge of algebra,

> even of all the algebraic techniques so far discovered, still there are topics whose consideration prepares a student for a deeper understanding. However, in the ordinary treatise on the elements of algebra, these topics are either completely omitted or are treated carelessly. For this reason, I am certain that the material I have gathered in this book is quite sufficient to remedy that defect.... There are many questions which are answered in this work by means of ordinary algebra, although they are usually discussed with the aid of analysis. In this way the interrelationship between the two methods becomes clear. [**12**, p. *v*]

What is this "ordinary algebra" that Euler spoke of, and how did it allow him to deduce results that we now classify as requiring differential calculus? The answer lies here: although Euler did not use the notion of the derivative to deduce these results (and certainly not theorems like Taylor's Theorem, which depend on the derivative) his notion of ordinary algebra went beyond what most of our contemporaries would include. In particular, Euler explicitly included the arithmetic of infinite and infinitesimal quantities, and implicitly used a general principle for simplifying calculations involving infinitely many infinitesimals. Because of this, Euler is often portrayed in popular accounts and classroom lectures as a reckless symbol-manipulator, who worked in a number system fraught with nonsense and contradiction, but who through sheer intuitive brilliance somehow came to correct conclusions. The following passages, taken from popular books on the history of mathematics, are typical.

> It is perhaps only fair to point out that some of Euler's works represent outstanding examples of eighteenth-century formalism, or the manipulation, without proper attention to matters of convergence and mathematical existence, of formulas involving infinite processes. He was incautious in his use of infinite series, often applying to them laws valid only for finite sums. Regarding power series as polynomials of infinite degree, he heedlessly extended to them well-known properties of finite polynomials. Frequently, by such careless approaches, he luckily obtained truly profound results .... [**13**, p. 435]

> Today, we recognize that Euler was not so precise in his use of the infinite as he should have been. His belief that finitely generated patterns and formulas automatically extend to the infinite case was more a matter of faith than science, and subsequent mathematicians would provide scores of examples showing the folly of such hasty generalizations. [**7**, p. 222]

In contrast we take Euler's calculations involving infinite and infinitesimal numbers seriously, and find that Euler's *Introductio* is written with grace, wit, and care. There is the occasional misstep, but on the whole, Euler's use of the infinite and infinitesimal is consistent and clear. Furthermore, there is a modern context, replete with infinite and infinitesimal numbers, in which Euler's methods can be made intelligible, rigorous, and useful to modern readers.

What follows is our own version of Euler's mathematical tale, sensitively rehabilitated to contemporary tastes for rigor.

## Exponentials and logarithms in Euler's *Introductio*

Euler began his introductory chapter on exponentials and logarithms [**12**, Chap. VI] by saying,

> Although the concept of a transcendental function depends on integral calculus, there are certain kinds of functions which are more obvious, which can be conveniently developed, and which open the door to further investigations.

He went on to explain the usual laws of exponents and logarithms, and illustrated the usefulness of tables of logarithms, much as one would in a precalculus course today, with examples from business and the life sciences.

> A certain man borrowed 400,000 florins at the usurious rate of five percent annual interest. Suppose that each year he repays 25,000 florins. The question is, how long will it be before the debt is repaid completely?...
>
> Since after the flood all men descended from a population of six, if we suppose that the population after two hundred years was 1,000,000, we would like to find the annual rate of growth.

To demonstrate the usefulness of tables of logarithms, Euler asked,

> If the progression 2, 4, 16, 256, ... is formed by letting each term be the square of the preceding term, find the value of the twenty-fifth term.

In the succeeding chapter, Euler developed the series for the exponential and logarithmic functions, and showed how to use series to compile tables of logarithms. What interests us here is the means by which Euler obtained those series. Euler began his discussion of the series for the exponential function as follows [**12**, Chap. VII]:

> Since $a^0 = 1$, when the exponent on $a$ increases, the power itself increases, provided that $a$ is greater than 1. It follows that if the exponent is infinitely small and positive, then the power also exceeds 1 by an infinitely small amount. Let $\epsilon$ be an infinitely small number, or, a fraction so small that, although not equal to 0, still $a^\epsilon = 1 + \psi$, where $\psi$ is also an infinitely small number. From the preceding chapter we know that unless $\psi$ were infinitely small, then neither would $\epsilon$ be infinitely small. It follows that $\psi = \epsilon$ or $\psi > \epsilon$ or $\psi < \epsilon$. Which of these is true depends on the value of $a$, which is not now known, so we let $\psi = \lambda\epsilon$. Then we have $a^\epsilon = 1 + \lambda\epsilon$ .... [**12**, §114]

(We have changed Euler's $\omega$ to $\epsilon$ and his $j$, in what follows, to $K$.) Euler then reasoned that if $x$ is any finite, positive, noninfinitesimal number, and $K$ is $x/\epsilon$, then by a simple calculation using the Binomial Theorem (discussed in §71 of the *Introductio*), a series for $a^x$ is given by

$$
\begin{aligned}
a^x = a^{K\epsilon} = (a^\epsilon)^K = (1+\lambda\epsilon)^K &= \left(1+\frac{\lambda x}{K}\right)^K \\
&\doteq 1 + \frac{1}{1}\lambda x + \frac{1(K-1)}{1\cdot 2K}\lambda^2x^2 + \frac{1(K-1)(K-2)}{1\cdot 2K\cdot 3K}\lambda^3x^3 + \cdots \\
&= 1 + \lambda x + \frac{K-1}{K}\cdot\frac{1}{1\cdot 2}\lambda^2x^2 + \frac{K-1}{K}\frac{K-2}{K}\cdot\frac{1}{1\cdot 2\cdot 3}\lambda^3x^3 + \cdots.
\end{aligned}
$$

Euler then reasoned that since $x$ is noninfinitesimal and $\epsilon$ is infinitesimal, $K$ will necessarily be infinite, and hence one may substitute 1 for the fractions $\frac{K-1}{K}$, $\frac{K-2}{K}$, $\frac{K-3}{K}$, and so on, to obtain

$$a^x = 1 + \lambda x + \frac{1}{2!}\lambda^2x^2 + \frac{1}{3!}\lambda^3x^3 + \cdots.$$

Finally, Euler examined the case in which the base $a$ is taken to correspond to $\lambda$ being equal to unity—the natural exponential function—and showed that in general $\lambda$ is the natural logarithm of $a$.

This argument, also discussed by Edwards [**9**, pp. 272–274] and Dunham [**8**], among others, is a gem of eighteenth-century mathematical reasoning, but there are several issues that must be dealt with before something like it could honestly be given in a modern context.

- Euler freely uses the arithmetic of infinite and infinitesimal numbers. If such numbers are to be used in a modern context, the rules for dealing with them must be presented as clearly, concisely, and consistently as the rules for ordinary numbers.
- Even granted a sound treatment of infinite and infinitesimal numbers, the reasoning by which one is allowed to make infinitely many substitutions—the numbers $\frac{K-1}{K}$, $\frac{K-2}{K}$, $\frac{K-3}{K}$, and so on, each being replaced by 1—must be explained. In each substitution instance, an error is incurred; for example, the difference between 1 and $\frac{K-1}{K}$ is $\frac{1}{K}$. Individually these differences are infinitesimal, but (as Euler was well aware) it is possible for infinitely many infinitesimals to add up to a noninfinitesimal amount.
- The argument as given employs the Binomial Theorem for nonintegral exponents, a theorem that Euler chose not to prove in the *Introductio*, and something that we would hesitate to assume in a modern precalculus course.

In our rehabilitation of Euler's methods for modern use, we deal with these issues as follows.

- We work in a consistent axiomatic system that clearly specifies the properties of infinite and infinitesimal numbers.
- We provide a criterion, based on the intuitive notion of *determinacy*, for deciding whether neglecting infinitely many infinitesimals leads to a negligible difference in an infinite sum.
- In our construction of the series for the exponential function, we find that the Binomial Theorem for natural exponents, a theorem that is verified by mathematical induction in traditional precalculus courses, suffices. (Later, in connection with the series for the logarithm, we give an elementary proof of the Binomial Theorem for fractional exponents.)

Once these issues are dealt with, we will return to Euler's argument and show how it can be rigorously rehabilitated in this context. We will then go on to obtain the series for the sine, cosine, and logarithm.

## The arithmetic of the infinite and infinitesimal

The first requirement of our rehabilitation of Euler's arguments is that his methods be formulated within a mathematical system in which the properties of infinite and infinitesimal numbers are explained at least as clearly as the properties of the real numbers. For this we turn to the system of *hyperreal* numbers, as described axiomatically in Keisler's textbook, *Calculus: An Infinitesimal Approach* [**23**].

In elementary courses, the real numbers are not defined explicitly; instead they are defined implicitly by their arithmetic properties, an approach that is essentially axiomatic. In more advanced courses one builds a model for the real numbers, typically using equivalence classes of Cauchy sequences of rational numbers. Similarly, the hyperreal numbers can either be introduced axiomatically or by building a model using equivalence classes of sequences of real numbers.

Keisler's textbook is intended for use in an introductory calculus course. He introduces the properties of the hyperreal numbers gradually, with appropriate examples and exercises, over the first forty pages of the book. The real numbers are described informally in the main body of the textbook, but presented more precisely in an appendix by citing the field axioms, the order axioms, the definition of the natural numbers, the root axiom (that principal $n^{\text{th}}$ roots exist for positive numbers), and the completeness axiom. Further axioms describe the hyperreal numbers as a field containing infinite and infinitesimal numbers in addition to all the real numbers. (He discusses a set-theoretic construction of the hyperreals in his guide for teachers [**22**].) Keisler sets the stage for extending the real numbers by reminding students that successive extensions of the notion of number have been the milestones of their mathematical educations.

> In grade school and high school mathematics, the real number system is constructed gradually in several stages. Beginning with the positive integers, the systems of integers, rational numbers, and finally real numbers are built up....
>
> What is needed [for an understanding of the calculus] is a sharp distinction between numbers which are small enough to be neglected and numbers which aren't. Actually, no real number except zero is small enough to be neglected. To get around this difficulty, we take the bold step of introducing a new kind of number, which is infinitely small and yet not equal to zero....
>
> The real line is a subset of the hyperreal line; that is, each real number belongs to the set of hyperreal numbers. Surrounding each real number $r$, we introduce a collection of hyperreal numbers infinitely close to $r$. The hyperreal numbers infinitely close to zero are called infinitesimals. The reciprocals of nonzero infinitesimals are infinite hyperreal numbers. The collection of all hyperreal numbers satisfies the same algebraic laws as the real numbers....
>
> We have no way of knowing what a line in physical space is really like. It might be like the hyperreal line, the real line, or neither. However, in applications of the calculus it is helpful to imagine a line in physical space as a hyperreal line. The hyperreal line is, like the real line, a useful mathematical model for a line in physical space. [**23**, pp. 1, 24, 25, 27]

In the picture of the hyperreal line (FIGURE 1), observe that $-\epsilon$, 0, and $1/2H$ are infinitesimal; $\pi + \epsilon$ is a finite, noninfinitesimal number that is infinitely close to $\pi$; $H$

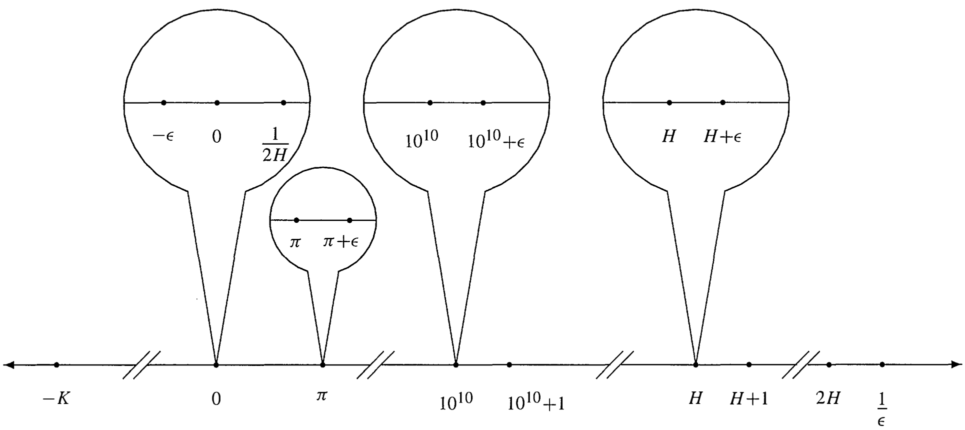



**Figure 1** The hyperreal line

is infinite, but infinitely close to $H + \epsilon$; $H$ is a finite, noninfinitesimal distance from $H + 1$, and infinitely far from $2H$.

Key computational properties of the hyperreal numbers are given in the following table.

RULES FOR INFINITE, FINITE, AND INFINITESIMAL NUMBERS. *Assume that $\epsilon$, $\delta$ are infinitesimals; $b$, $c$ are hyperreal numbers that are finite but not infinitesimal; $H$, $K$ are infinite hyperreal numbers; and $n$ is a finite natural number.*

- Real numbers. *The only infinitesimal real number is* 0. *Every real number is finite.*
- Negatives. $-\epsilon$ *is infinitesimal;* $-b$ *is finite but not infinitesimal;* $-H$ *is infinite.*
- Reciprocals. *If* $\epsilon \neq 0$, *then* $1/\epsilon$ *is infinite;* $1/b$ *is finite but not infinitesimal;* $1/H$ *is infinitesimal. Note that* $1/0$ *remains undefined.*
- Sums. $\epsilon + \delta$ *is infinitesimal;* $b + \epsilon$ *is finite but not infinitesimal;* $b + c$ *is finite (possibly infinitesimal);* $H + \epsilon$ *and* $H + b$ *are infinite.*
- Products. $\delta \cdot \epsilon$ *and* $b \cdot \epsilon$ *are infinitesimal;* $b \cdot c$ *is finite but not infinitesimal;* $H \cdot b$ *and* $H \cdot K$ *are infinite.*
- Quotients. $\epsilon/b$, $\epsilon/H$, *and* $b/H$ *are infinitesimal;* $b/c$ *is finite but not infinitesimal;* $b/\epsilon$, $H/\epsilon$, *and* $H/b$ *are infinite, provided that* $\epsilon \neq 0$.
- Powers. $\epsilon^n$ *is infinitesimal;* $b^n$ *is finite but not infinitesimal;* $H^n$ *is infinite.*
- Roots. *If* $\epsilon > 0$ *then* $\sqrt[n]{\epsilon}$ *is infinitesimal; if* $b > 0$ *then* $\sqrt[n]{b}$ *is finite but not infinitesimal; if* $H > 0$ *then* $\sqrt[n]{H}$ *is infinite.*

*Notice that there are no general rules for deciding whether the combinations* $\epsilon/\delta$, $H/K$, $H\epsilon$, *and* $H + K$, *are infinitesimal, finite, or infinite.*

DEFINITION. We write $x \simeq y$ to mean that $x - y$ is infinitesimal. If $x \simeq y$, we say that $x$ is *infinitely close* to $y$.

Keisler's entire course is based on three fundamental principles relating the real and hyperreal numbers: the Extension Principle, the Transfer Principle, and the Standard Part Principle. The *Extension Principle* posits the existence of nonzero infinitesimals in the hyperreal field, and for each real function $f$, a function $^*f$ extending $f$ to the hyperreal numbers. The function $^*f$ is called the *hyperreal extension* of $f$. (A function is a set of ordered pairs such that no two pairs have the same first element and different second elements. If $f$ and $g$ are functions, then by "$g$ extends $f$" or "$g$ is an extension of $f$" we mean that $f$ is a subset of $g$. A *real function of one variable* is a function in which the domain and range are sets of real numbers. A *real function of n variables* is a function in which the domain is a set of $n$-tuples of real numbers and the range is a set of real numbers.) The *Transfer Principle* says that every *real statement* that holds for a particular real function holds for its hyperreal extension as well. Equations and inequalities are examples of real statements.

> Here are seven examples that illustrate what we mean by a *real statement*.... (1) Closure law for addition: for any $x$ and $y$, the sum $x + y$ is defined. (2) Commutative law for addition: $x + y = y + x$. (3) A rule for order: If $0 < x < y$ then $0 < 1/y < 1/x$. (4) Division by zero is never allowed: $x/0$ is undefined. (5) An algebraic identity: $(x - y)^2 = x^2 - 2xy + y^2$. (6) A trigonometric iden-

tity: $\sin^2 x + \cos^2 x = 1$. (7) A rule for logarithms: If $x > 0$ and $y > 0$ then $\log_{10}(xy) = \log_{10} x + \log_{10} y$. [**23**, pp. 28–29]

(Keisler later gives a precise characterization of the real statements [**23**, p. 907].) A consequence of the Transfer Principle is that one does not ordinarily need to distinguish between $*f$ and $f$, since any real statement true of one of these functions will be true of the other: for simplicity we use the same function symbol $f$ for both $*f$ and $f$. Finally, the *Standard Part Principle* says that every finite hyperreal number is infinitely close to exactly one real number; this principle is useful for translating results about finite hyperreal quantities into equivalent statements about real quantities.

In our development, which emphasizes discrete mathematics, the natural numbers play a larger role than they do in most presentations of the calculus. Key properties of the natural numbers are that they contain 0 and 1, are closed under $+$ and $\cdot$, and that they satisfy the Natural Induction Principle (also known as the Principle of Mathematical Induction). For example, the binomial formula,

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2}a^{n-2}b^2 + \cdots + b^n,$$

for all real $a$, $b$ and all natural $n$, and the geometric sum formula,

$$\frac{1-a^{n+1}}{1-a} = 1 + a + a^2 + \cdots + a^n,$$

for all $a$ except 1 and all natural $n$, are often proved by induction. We will only require the Natural Induction Principle for equations and inequalities. In the following, a *real sequence* is a real function in which the domain is the set of natural numbers.

NATURAL INDUCTION PRINCIPLE. *Let $\phi(n)$ be an equation or inequality of real sequences; that is, let $\phi(n)$ be of the form $a_n = b_n$, $a_n \neq b_n$, $a_n < b_n$, or $a_n \leq b_n$, where a and b are real sequences. If $\phi(0)$ holds and if for all natural m, we have that $\phi(m+1)$ holds whenever $\phi(m)$ holds, then $\phi(n)$ holds for all natural n.*

Another important tool is the Principle of Definition by Recursion, which says that one may define a real sequence by specifying its value at 0, and specifying for each natural $n$ its value at $n+1$ as determined by its value at $n$. (See [**2**] for an elementary discussion of recursion schemes and their solutions.) For example, the *factorial-power* function,

$$x^{\underline{n}} = \underbrace{x(x-1)(x-2)\cdots(x-n+1)}_{n \text{ factors}},$$

is defined for all natural $n$ by the equations,

$$x^{\underline{0}} = 1, \quad x^{\underline{n+1}} = x^{\underline{n}} \cdot (x-n).$$

A *real series* is a real sequence of partial sums $a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots,$ where $a$ is a real sequence; we use the notation $a_0 + a_1 + a_2 + \cdots$ to denote this series. Real series are defined more formally by recursion. For example, the sum of the first $n$ square numbers is defined for all natural $n$ by the equations

$$s_0 = 0, \quad s_{n+1} = s_n + (n+1)^2.$$

The *integers* are defined to be the natural numbers together with their negatives. An important function from the real numbers to the integers is the *greatest-integer* function, where $\lfloor x \rfloor$ is the greatest integer less than or equal to $x$. Following Keisler's presentation, we define the *hyperintegers* to be the range of $^*\lfloor\ \rfloor$. The *hypernatural* numbers are defined to be the nonnegative hyperintegers. They extend the natural numbers to include infinite elements that satisfy the same real statements (such as the recursive definitions of addition and multiplication) as the ordinary, finite natural numbers.

There are several ways of defining *hypersequences* having hypernatural indices and hyperreal values. The simplest way is to start with any real sequence $s$ and take its hyperreal extension, $^*s$. By the Transfer Principle, $^*s_n = s_n$ for all finite $n$, but $^*s_N$ has new, hyperreal values for infinite $N$.

*Example*: $0, 1, 4, 9, 16, \ldots, N^2, \ldots$ ($N$ hypernatural). If $t$ is the sequence $0, 1, 4, 9, 16, \ldots, n^2, \ldots$ ($n$ natural), then $^*t$ is $0, 1, 4, 9, 16, \ldots, N^2, \ldots$ ($N$ hypernatural). In terms of sets, $t$ is $\{(n, n^2) : n \text{ natural}\}$ and $^*t$ is $\{(N, N^2) : N \text{ hypernatural}\}$.

*Example*: $0, 1, 5, 14, 30, \ldots, (0^2 + 1^2 + 2^2 + 3^2 + \ldots + N^2), \ldots$ ($N$ hypernatural). If $s$ is the real series (sequence of partial sums) defined on the natural numbers by $s_0 = 0$, $s_{n+1} = s_n + (n+1)^2$, then by the Extension Principle $^*s$ is defined on the hypernatural numbers, and by the Transfer Principle $^*s$ satisfies the same real statements as $s$—in particular, the same recursion equations. Thus $s_N$, also written $\sum_{n=0}^{N} n^2$ or even $0^2 + 1^2 + 2^2 + 3^2 \cdots + N^2$, makes sense for infinite as well as finite $N$.

More generally, one may start with a real function of one or more variables, take its hyperreal extension, and then substitute hyperreal values for some of its arguments.

*Example*:

$$1,\ \lambda,\ \frac{K-1}{K}\cdot\frac{\lambda^2}{1\cdot 2},\ \frac{K-1}{K}\cdot\frac{K-2}{K}\cdot\frac{\lambda^3}{1\cdot 2\cdot 3},\ \ldots.$$

These are the coefficients of the binomial expansion of $(1 + (\lambda x/K))^K$, which Euler used in his construction of the series for $a^x$. These terms are given by

$$\beta_n = \frac{K^{\underline{n}}}{K^n}\frac{\lambda^n}{n!},$$

where $\lambda$ and $K$ are fixed hyperreal numbers and $n$ ranges over the hypernatural numbers. The hyperreal function $\beta$ arises from the three-argument real function $b$ defined by

$$b(k, l, n) = \frac{k^{\underline{n}}}{k^n}\frac{l^n}{n!}$$

by fixing $K$ and $\lambda$ and setting $\beta_n = {}^*b(K, \lambda, n)$ for all hypernatural $n$.

DEFINITION. A *hypersequence* is any function defined on the hypernatural numbers by composing the hyperreal extensions of real functions of one or more variables and allowing hyperreal arguments. A *hyperseries* is the hypersequence of partial sums of a hypersequence. We often use the term *series* to refer to either a real series or a hyperseries.

By the Transfer Principle one can extend the binomial formula and the geometric sum formula to hyperreal terms and hypernatural exponents.

BINOMIAL THEOREM. *For all hyperreal a, b, and all hypernatural n,*

$$(a+b)^n = \sum_{k=0}^{n} \frac{n^{\underline{k}}}{k!} a^{n-k} b^k = a^n + \frac{n^{\underline{1}}}{1!} a^{n-1} b^1 + \frac{n^{\underline{2}}}{2!} a^{n-2} b^2 + \frac{n^{\underline{3}}}{3!} a^{n-3} b^3 + \cdots + b^n.$$

GEOMETRIC SUM THEOREM. *For all hyperreal a except unity and all hypernatural n,*

$$\frac{1-a^{n+1}}{1-a} = \sum_{k=0}^{n} a^k = 1 + a + a^2 + \cdots + a^n.$$

Hypersequences are examples of the *internal* sequences of Robinson's theory; see [**21**, pp. 94ff]. Because of the special role of hypersequences in our exposition, we will find it convenient to assume one further principle.

HYPERNATURAL INDUCTION PRINCIPLE. *Let $\phi(n)$ be an equation or inequality of hypersequences; that is, let $\phi(n)$ be of the form $a_n = b_n$, $a_n \neq b_n$, $a_n < b_n$, or $a_n \leq b_n$, where a and b are hypersequences. If $\phi(0)$ holds and if for all hypernatural m, we have that $\phi(m+1)$ holds whenever $\phi(m)$ holds, then $\phi(n)$ holds for all hypernatural n.*

In more advanced treatments, the Hypernatural Induction Principle can be seen to follow from the Natural Induction Principle and a version of the Transfer Principle that takes into account statements involving quantifiers, in addition to the (quantifier-free) real statements.

Sullivan, in her article in the *American Mathematical Monthly* [**44**], provides evidence that elementary calculus can be effectively taught to high school and college students using Keisler's system of hyperreal numbers. A recent reform-calculus book that uses infinitesimal methods is Stroyan's *Calculus using Mathematica* [**42**]. Interested readers might also consult Luxemburg's article in the *Monthly* [**32**], Lightstone's articles in the *Monthly* [**30**] and this MAGAZINE [**31**], Davis and Hersh's "Nonstandard analysis" in *Scientific American* [**6**], Simpson's article from the *Mathematical Intelligencer* [**41**], and Henle and Kleinberg's slender volume, *Infinitesimal Calculus* [**18**]. For more advanced treatments, see [**40**], [**43**], or [**21**]. Keisler's article [**24**] contains a brief history of infinitesimals. For a nonstandard connection between Euler's mathematics and modern functional analysis, see [**45**].

## Determinate series

Much of the *Introductio* concerns the expansion of well-known functions into series:

> Since both rational functions and irrational functions of $x$ are not of the form of polynomials $A + Bx + Cx^2 + Dx^3 + \cdots$, where the number of terms is finite, we are accustomed to seek expressions of this type with an infinite number of terms which give the value of the rational or irrational function. Even the nature of a transcendental function seems to be better understood when it is expressed in this form, even though it is an infinite expression. Since the nature of polynomial functions is very well understood, if other functions can be expressed by different powers of $x$ in such a way that they are put in the form $A + Bx + Cx^2 + Dx^3 + \cdots$, then they seem to be in the best form for the mind to grasp their nature, even though the number of terms is infinite. [**12**, §59]

Implicit in this statement is the assumption that "infinite polynomials" share well-known properties of finite polynomials. In our rehabilitation of Euler's methods, the *polynomials with an infinite number of terms* become *polynomials of infinite hypernatural degree*: $a_0 + a_1x + a_2x^2 + \cdots + a_Nx^N$, where $N$ is an infinite hypernatural number. By the Transfer Principle, such hyperreal polynomials satisfy the same real statements as real polynomials of finite degree, and in particular can be algebraically manipulated according to the usual rules. Hyperreal polynomials cannot provide *exact* expressions for nonpolynomial real functions, but the extension to the hyperreals does present the opportunity for approximating a real function to within *infinitesimal* error—and for most practical purposes this is close enough. In our rehabilitation of Euler's methods, the goal of expressing a real function as a polynomial with an infinite number of terms becomes: for a real function $f$, to find a hypersequence $a$ and an infinite hypernatural $N$ such that for all real $x$ (or for all real $x$ in some range),

$$f(x) \simeq a_0 + a_1x + a_2x^2 + \cdots + a_Nx^N.$$

It would be computationally inconvenient if the value of $a_0 + a_1x + a_2x^2 + \cdots + a_Nx^N$ were to depend perceptibly on the *particular* infinite value of $N$. Therefore we give special consideration to series that are *determinate* in the sense that once one has taken the summation to an infinite number of terms, the contribution made by adding still more terms is infinitesimal.

Euler did not discuss the notion of determinacy in the *Introductio* or anywhere else—with one exception. In a paper on harmonic series presented in 1734, Euler stated a principle which may be read as follows: "A series that has a finite sum when continued infinitely will receive insignificant growth even if it is continued further; in fact, that which is added after infinitely many terms will be infinitely small." (*Series, quae infinitum continuata summum habet finitam, etiamsi ea duplo longius continuetur, nullum accipiet augmentum, sed id, quod post infinitum adiicitur cogitatione, re vera erit infinite parvum.* [**11**, §2]) He used this principle to show that a harmonic series

$$\frac{c}{a} + \frac{c}{a+b} + \frac{c}{a+2b} + \frac{c}{a+3b} + \cdots$$

does not have a finite sum, but that series such as

$$\frac{c}{a} + \frac{c}{a+b} + \frac{c}{a+4b} + \frac{c}{a+9b} + \cdots$$

and in general series whose $n^{\text{th}}$ term is $c/(a + n^\alpha b)$, $\alpha > 1$, do have finite sums. (In all of these cases, the assumption that $a$, $b$, and $c$ are positive is implicit.) Because of its essential use of infinite and infinitesimal numbers, we find it worthwhile to recount Euler's argument that the harmonic series is not determinate [**11**, §3]:

> Let the series $c/a$, $c/(a+b)$, $c/(a+2b)$, etc., be continued infinitely to the infinitesimal term $c/(a+(i-1)b)$, where $i$ denotes an infinite number, the index of this term. Now if this series is continued from the next term $c/(a+ib)$ through the $ni^{\text{th}}$ term $c/(a+(ni-1)b)$, the number of these added terms is $(n-1)i$. The sum of these terms is less than
>
> $$\frac{(n-1)ic}{a+ib},$$
>
> and greater than
>
> $$\frac{(n-1)ic}{a+(ni-1)b}.$$
>
> Since $i$ is infinitely large, $a$ is negligible in each denominator; thus the sum is greater than
>
> $$\frac{(n-1)c}{nb},$$

and less than

$$\frac{(n-1)c}{b}.$$

Note the salient features of this argument. The number $i$ is explicitly taken to be *infinite*, and a sum of $i$ terms is taken, *terminating* with $c/(a+(i-1)b)$. After summing these infinitely many terms there is a *next term*, $c/(a+ib)$. A tail sum is taken of the next $(n-1)i$ terms, and a lower bound is obtained for this tail sum using *ordinary algebra*, which can then be simplified because *$a/i$ is infinitesimal*:

$$\frac{(n-1)ic}{a+(ni-1)b} = \frac{(n-1)c}{a/i+(n-1/i)b} \simeq \frac{(n-1)c}{nb}.$$

For example, if we take $n=2$ (continuing the sum twice as far), we have a tail sum that is greater than or infinitely close to $c/2b$, and hence not infinitesimal. Therefore the series is not determinate.

What does the notion of determinacy have to do with the *Introductio*? Euler's techniques for expanding functions into series depend at various points on the negligibility of infinitely many infinitesimals in an infinite sum. There are easy examples to the contrary (arising, for example, in the computation of areas as infinite sums of infinitesimal rectangles) so to be rigorous, one must have a criterion for deciding when one can neglect infinitesimals in an infinite sum. The notion of determinacy provides such a criterion.

DEFINITION OF DETERMINACY. *A hypersequence $s_0, s_1, s_2, \ldots$ is said to be* determinate *iff $s_M \simeq s_N$ for all infinite $M$ and $N$. If $a_0, a_1, a_2, \ldots$ is a hypersequence, then a series $a_0+a_1+a_2+\cdots$ is said to be determinate iff the hypersequence of partial sums defined by $s_n = a_0+a_1+a_2+\cdots+a_n$ is determinate.*

The following theorem says that one can neglect infinitely many infinitesimals in an infinite sum provided the relevant series are both determinate.

SUMMATION COMPARISON THEOREM. *If the series $a_0+a_1+a_2+\cdots$ and $b_0+b_1+b_2+\cdots$ are determinate, and if for each natural $n$, $a_n \simeq b_n$, then for all hypernatural $n$, $a_0+a_1+\cdots+a_n \simeq b_0+b_1+\cdots+b_n$.*

We will postpone the proof of this theorem to a later section. As an example of a determinate series we verify that a geometric series $1+x+x^2+\cdots$ is determinate for certain values of $x$. Let $x$ be a hyperreal number such that $|x|<1$ and $|x| \not\simeq 1$, and let $J$ and $K$ be infinite hypernatural numbers with $K>J$. Then by the Geometric Sum Theorem, $x^J+\cdots+x^K = (x^J - x^{K+1})/(1-x)$. This is infinitesimal because both $x^J$ and $x^{K+1}$ are infinitesimal and $1/(1-x)$ is finite.

The following theorem contains two general tests for determinacy. The proof is left to the reader.

COMPARISON TEST FOR DETERMINACY. ***(i)*** *Let $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$ be sequences of positive terms. If $b_0+b_1+b_2+\cdots$ is determinate and if there is a finite $k$ such that $a_n \le b_n$ for $n \ge k$, then $a_0+a_1+a_2+\cdots$ is determinate as well.* ***(ii)*** *If $|c_0|+|c_1|+|c_2|+\cdots$ is determinate, then $c_0+c_1+c_2+\cdots$ is determinate as well.*

The requirement that *once one has added an infinite number of terms, the contribution made by adding still more terms must be infinitesimal* bears a striking resemblance to the Cauchy condition for convergence of real series, which says that *once one has added a sufficiently large finite number of terms, the contribution made by adding still* *more terms must always be less than some previously specified amount.* This resemblance has been discussed by Eneström [**10**] and Pringsheim [**39**], and more recently by Laugwitz [**29**, 205–208]. Our presentation was inspired by Laugwitz's discussion. Did Euler anticipate the "Cauchy" criterion for convergence? The answer is far from being free of controversy (see McKinzie [**35**]) and moreover, even if he did, the discovery seems to have made no difference to the historical development of the calculus. Eneström, compiler of the definitive catalog of Euler's published works, lamented:

> I have looked in vain for a reference to the Eulerian convergence condition in the accessible mathematical writings of the 18$^{\text{th}}$ century. The discovery appears therefore to remain completely unheeded, and the mathematicians who attack the convergence question at the start of the 19$^{\text{th}}$ century were surely not influenced by Euler. [**10**]

That much said, we still have no qualms about using our own definition of determinacy, simply and clearly stated in Eulerian language, in our rehabilitation of Euler's methods: as we shall see it is precisely what we need to make Euler's arguments rigorous.

## The exponential series

Having outlined an axiomatic system that specifies the properties of infinite and infinitesimal numbers, and having provided a criterion for the negligibility of infinitesimals in an infinite sum, we are now ready to present our rehabilitation of Euler's derivation of the series for the exponential function.

Exponentiation is defined for 0 and other natural $n$ by

$$a^0 = 1$$
$$a^n = \underbrace{a \cdots a}_{n \text{ factors}},$$

or more formally using recursion, then extended to positive rational exponents $\frac{m}{n}$ using the root axiom:

$$a^{\frac{m}{n}} = \left(\sqrt[n]{a}\right)^m \quad (a > 0);$$

then extended to negative rational numbers by taking reciprocals,

$$a^{-\frac{m}{n}} = 1/a^{\frac{m}{n}} \quad (a > 0).$$

From these definitions and the basic field and order axioms for the real numbers, one can show that for all positive $a$ greater than 1 and all rational $p$ and $q$, the following familiar rules hold for rational exponentiation: $a^p a^q = a^{p+q}$, $(a^p)^q = a^{pq}$, and $a^p < a^q$ if and only if $p < q$.

Extending the definition of $a^x$ further, from rational to real $x$, and verifying that the rules just given also hold for the extension, is more involved. Instead, we *assume* that $a^x$ is a real function (defined for all real $x$) and, using the properties mentioned in the previous paragraph as motivation, assume the following axioms for the exponential function.

AXIOMS. *For all positive real a greater than* 1 *and all real x and y, the following rules hold:* $a^0 = 1$; $a^{-x} = 1/a^x$; $a^x a^y = a^{x+y}$; $(a^x)^y = a^{xy}$; *and* $a^x < a^y$ *iff* $x < y$.

By the Transfer Principle, the axioms stated above hold for hyperreal numbers as well. In addition to these rules, we also require the following proposition.

PROPOSITION. *Assuming that a is finite and greater than* 1, *we have the following results. If* $\epsilon > 0$ *and* $\epsilon \simeq 0$ *then* $a^\epsilon > 1$ *and* $a^\epsilon \simeq 1$. *If x and y are finite, then* $a^x \simeq a^y$ *iff* $x \simeq y$.

*Proof.* We prove this in three steps. (i) Let $N$ be an infinite hypernatural number. We want to conclude that $a^{1/N}$ exceeds 1 by an infinitesimal amount. By the axioms for exponentiation, $a^{1/N} > a^0 = 1$, so we may write $a^{1/N} = 1 + u$, where $u$ is positive. By the Binomial Theorem, $a = (a^{1/N})^N = (1+u)^N = 1 + Nu +$ (positive terms), from which we conclude that $0 < u < (a-1)/N \simeq 0$, and hence $a^{1/N} > 1$ and $a^{1/N} \simeq 1$. (ii) Now let $\epsilon$ be positive, and take $N = \lfloor 1/\epsilon \rfloor$, the greatest hypernatural number less than or equal to $1/\epsilon$, so that $1/(N+1) < \epsilon \leq 1/N$. Then by the axioms for exponentiation, $a^{1/(N+1)} < a^\epsilon \leq a^{1/N}$, from which it follows that $\epsilon \simeq 0$ iff $N$ is infinite iff $a^\epsilon \simeq 1$. Furthermore, $a^{-\epsilon}$, which is $1/a^\epsilon$, is infinitely close if 1 iff $a^\epsilon$ is as well. (iii) Assuming that both $x$ and $y$ are finite, we conclude that

$$a^x \simeq a^y \quad \text{iff} \quad \frac{a^x}{a^y} \simeq 1 \quad \text{iff} \quad a^{x-y} \simeq 1 \quad \text{iff} \quad x - y \simeq 0 \quad \text{iff} \quad x \simeq y.$$

It is important in the first step that $a^x$ and $a^y$ are neither infinite nor infinitesimal. ∎

Our goal for this section is to show that there is a finite $\lambda$ such that for all finite $x$ and infinite $N$,

$$a^x \simeq a + (\lambda x) + \frac{1}{2!}(\lambda x)^2 + \frac{1}{3!}(\lambda x)^3 + \cdots + \frac{1}{N!}(\lambda x)^N.$$

Let $x$ be finite, and for the moment, positive. We choose an infinite hypernatural number $K$, which we hold fixed for the rest of this section, and choose a fraction $J/K$ that is infinitely close to $x$. This can be done by taking $J = \lfloor Kx \rfloor$, so that $0 < x - J/K < 1/K$, and hence that $x \simeq J/K$.

By the proposition, we know that $a^x \simeq a^{J/K}$, so let us now work with $a^{J/K}$. We write $a^{J/K}$ as $(a^{1/K})^J$, and consider $a^{1/K}$. By the proposition, $a^{1/K}$ exceeds 1 by an infinitesimal amount. We do not know whether that amount is greater or less than $1/K$, so (following Euler) we introduce a positive scaling factor, $\lambda$, depending on $K$:

$$a^{1/K} = 1 + \lambda\frac{1}{K}.$$

It is easy to see that $\lambda$ must be finite: by the Binomial Theorem, $a = (a^{1/K})^K = (1 + \lambda/K)^K = 1 + \lambda +$ (positive terms), so that $0 < \lambda < a$.

We may now expand $a^{J/K}$, written as $(1 + \lambda\frac{1}{K})^J$, as follows:

$$\begin{aligned} a^x \simeq a^{J/K} = (a^{1/K})^J &= \left(1 + \lambda\frac{1}{K}\right)^J \\ &= 1 + J\left(\lambda\frac{1}{K}\right) + \frac{J^{\underline{2}}}{2!}\left(\lambda\frac{1}{K}\right)^2 + \cdots + \frac{J^{\underline{J}}}{J!}\left(\lambda\frac{1}{K}\right)^J && (1) \\ &= 1 + \left(\lambda\frac{J}{K}\right) + \frac{1}{2!}\frac{J^{\underline{2}}}{J^2}\left(\lambda\frac{J}{K}\right)^2 + \cdots + \frac{1}{J!}\frac{J^{\underline{J}}}{J^J}\left(\lambda\frac{J}{K}\right)^J \\ &\simeq 1 + (\lambda x) + \frac{1}{2!}(\lambda x)^2 + \cdots + \frac{1}{J!}(\lambda x)^J && (2) \end{aligned}$$

$$\simeq 1 + (\lambda x) + \frac{1}{2!}(\lambda x)^2 + \cdots + \frac{1}{N!}(\lambda x)^N \tag{3}$$

Line (1) follows from the Binomial Theorem, and lines (2) and (3) will be justified by the Summation Comparison Theorem once we have shown that the series in question are determinate and that their respective terms are infinitely close.

LEMMA. *The series* $1 + y + \frac{1}{2!}y^2 + \frac{1}{3!}y^3 + \cdots$ *is determinate for all finite* $y$.

*Proof.* Fix $y > 0$ and let $n_0 = \lfloor y \rfloor$. Then for $n > n_0$,

$$\frac{y^n}{n!} = \frac{y^{n_0}}{n_0!}\frac{y^{n-n_0}}{(n_0+1)(n_0+2)\cdots n} \le \frac{y^{n_0}}{n_0!}\left(\frac{y}{n_0+1}\right)^{n-n_0} = b\,c^{n-n_0},$$

where $b = y^{n_0}/n_0!$ and $c = y/(n_0+1)$, so that $|c| < 1$, $c \not\simeq 1$, and $b$ is finite. As we saw earlier, the series $1 + c + c^2 + c^3 + \cdots$ is determinate for $0 < c < 1$, so the result follows from the Comparison Test for Determinacy. ■

Setting $y$ to $\lambda x$ in the lemma shows that the series in (2) is determinate, and since for positive $x$,

$$\frac{1}{k!}\frac{J^{\underline{k}}}{J^k}\left(\lambda\frac{J}{K}\right)^k \le \frac{1}{k!}(\lambda x)^k,$$

the Comparison Test for Determinacy implies that the series in (3) is determinate. We next note that since $J$ is infinite, we have $J^{\underline{k}}/J^k \simeq 1$ and $(J/K)^k \simeq x^k$ for all finite $k$, and hence

$$\frac{J^{\underline{k}}}{J^k}\left(\frac{J}{K}\right)^k\frac{\lambda^k}{k!} \simeq \frac{\lambda^k}{k!}x^k$$

for all finite $k$ as well. Using the Summation Comparison Theorem, and similar reasoning for negative exponents, we obtain the desired theorem.

THEOREM. *If* $a$ *is finite and greater than* 1, *then there is a finite* $\lambda$ *such that*

$$a^x \simeq 1 + (\lambda x) + \frac{1}{2!}(\lambda x)^2 + \frac{1}{3!}(\lambda x)^3 + \cdots + \frac{1}{N!}(\lambda x)^N$$

*for all finite* $x$ *and infinite* $N$.

## The natural exponential series

Suppose one wanted to compute $10^x$ using the previous theorem. What value of $\lambda$ would one use? From our original equation, $a^{1/K} = 1 + \lambda\frac{1}{K}$, one can deduce that $\lambda = K(a^{1/K} - 1)$, but this formula is difficult to evaluate, in that it requires the extraction of a large-order root of $a$. Later in this article, we will use a series to compute $\lambda$, but in the mean time one can ask, why not take $\lambda$ to be some value convenient for computation, and use the value of $a$ corresponding to that value of $\lambda$? Euler noted that "[s]ince we are free to choose the base $a \ldots$, we now choose $a$ in such a way that $\lambda = 1$." [**12**, §§122–123] That is, we choose $a = (1 + 1/K)^K$, so that the corresponding series for $a^x$ is

$$1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots.$$

But does this help? Now we need to know this special value of $a$. Noting this difficulty, Euler wrote,

> [If] we now choose $a$ such that $\lambda = 1 \ldots$ then the series
>
> $$1 + \frac{1}{1} + \frac{1}{1\cdot 2} + \frac{1}{1\cdot 2\cdot 3} + \frac{1}{1\cdot 2\cdot 3\cdot 4} + \cdots$$
>
> is equal to $a$. If the terms are represented as decimal fractions and summed, we obtain the value for $a = 2.71828182845904523536028\ldots$. For the sake of brevity, for this number ... we will use the symbol $e$ .... [**12**, §122]

The function $e^x$ is called the *natural exponential function.* According to Cajori [**3**, §400], Euler first used the letter $e$ to represent the natural exponential base in a manuscript of 1727 or 1728, published posthumously in 1862. The notation first found its way into print in Euler's *Mechanica sive motus scientia analytice exposita* of 1736. Its use in such influential works as the *Mechanica* and the *Introductio* established the notation as standard. See also Coolidge [**4**] and Maor [**34**].

Let us verify that $(1 + 1/K)^K$ is determinate in the sense that for different infinite values of $K$ the corresponding values are all infinitely close.

PROPOSITION. *For all infinite $M$, $N$, and $P$,*

$$\left(1 + \frac{1}{M}\right)^M \simeq 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{P!} \simeq \left(1 + \frac{1}{N}\right)^N.$$

*Proof.* Expanding the binomial power and repeatedly applying the Summation Comparison Theorem, we find that for all infinite $M$ and $P$,

$$\begin{aligned}\left(1 + \frac{1}{M}\right)^M &= 1 + M\left(\frac{1}{M}\right) + \frac{M^{\underline{2}}}{2!}\left(\frac{1}{M}\right)^2 + \cdots + \frac{M^{\underline{M}}}{M!}\left(\frac{1}{M}\right)^M \\ &= 1 + 1 + \frac{M^{\underline{2}}}{M^2}\frac{1}{2!} + \cdots + \frac{M^{\underline{M}}}{M^M}\frac{1}{M!} \\ &\simeq 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{M!} \simeq 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{P!}.\end{aligned}$$

Since this computation holds for infinite $N$ as well as $M$, the result follows. ∎

At this stage one might be tempted to define $e$ to be any one of the values $(1 + 1/K)^K$, for $K$ infinite, and, by now-familiar computations obtain the relation,

$$e^x \simeq 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{N!}x^N \tag{4}$$

for all finite $x$ and infinite $N$. But this does not seem to adequately pin down the value of $e$; one would prefer to have $e$ stand for some specific real number rather than having it be an arbitrary choice from an anonymous class of hyperreal numbers, albeit all infinitely close. Our remedy is to use the Standard Part Principle, which says that every finite hyperreal number has exactly one real number that is infinitely close to it:

DEFINITION. *$e$ is the unique real number that is infinitely close to $(1+1/K)^K$, where $K$ is infinite.*

Finally, we must verify that this value of $e$, which actually differs (infinitesimally) from $(1+1/K)^K$ for each $K$, still satisfies (4). This is a consequence of the following proposition, about different exponential functions whose bases are infinitely close.

PROPOSITION. *If $a$ and $b$ are finite and greater than 1, and $a \simeq b$, then $a^x \simeq b^x$ for all finite $x$.*

*Proof.* If $a \simeq b$ and $a, b > 1$, then we may write $b = a(1+\epsilon)$ where $\epsilon \simeq 0$. Then $b^x = (a(1+\epsilon))^x = a^x(1+\epsilon)^x$. We need only verify that $(1+\epsilon)^x \simeq 1$. Let $n = \lfloor x \rfloor$. Then $n \le x < n+1$, and by the order axiom for exponentiation, $(1+\epsilon)^n \le (1+\epsilon)^x < (1+\epsilon)^{n+1}$. Since $\epsilon \simeq 0$ and $n$ is finite, by the Binomial Theorem we have $1 \le (1+\epsilon)^n = 1 + \epsilon \cdot$ (a sum of $n$ finite terms) $\simeq 1$, and $(1+\epsilon)^{n+1} = (1+\epsilon)^n(1+\epsilon) \simeq 1$, and hence $(1+\epsilon)^x \simeq 1$, and finally $b^x \simeq a^x$. ■

Since $e \simeq (1+1/K)^K$, by the previous proposition we conclude that

THEOREM. *For all finite $x$ and infinite $N$,*

$$e^x \simeq 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{N!}x^N. \tag{5}$$

By the way, this theorem helps explain the relationship between $\lambda$ and $a$. For if $\lambda$ happens to satisfy the equation $e^\lambda = a$, then

$$a^x = (e^\lambda)^x = e^{\lambda x} \simeq 1 + (\lambda x) + \frac{1}{2!}(\lambda x)^2 + \frac{1}{3!}(\lambda x)^3 + \cdots + \frac{1}{N!}(\lambda x)^N,$$

for all finite $x$ and infinite $N$. This shows that the $\lambda$ we chose earlier can be taken to be the exponent to which $e$ must be raised in order to yield $a$: in other words, the natural logarithm of $a$. We will return to the natural logarithm later in this article.

## The Euler identities and the series for sine and cosine

Although series for the exponential function, logarithm, and trigonometric functions were known to Newton and others prior to 1670 (see [**25**, p. 436ff ]), Euler's *Introductio in Analysin Infinitorum* of 1748 provided a systematic account of these formulas as deduced from basic principles. According to Boyer,

> [The *Introductio*] contains the earliest algorithmic treatment of logarithms as exponents and of the trigonometric functions as numerical ratios. It was the first textbook to list systematically the multiple-angle formulas, calling attention to the periodicities of the functions; and it included the first general analytic treatment of these as infinite products, as well as their expansion into infinite series. The well-known "Euler identities," relating the trigonometric functions to imaginary exponentials, are also found here. [**1**, pp. 224–225]

In this section, we will use the multiple-angle formulas to deduce a form of the Euler identities, and use these identities to derive the series for sine and cosine.

The Euler identities are well known to us in the form

$$\cos x = \tfrac{1}{2}[e^{ix} + e^{-ix}], \qquad \sin x = \tfrac{1}{2i}[e^{ix} - e^{-ix}],$$

or equivalently, $e^{ix} = \cos x + i \sin x$, but to appreciate these formulas one must understand what is meant by $e^{ix}$. This is difficult because the axioms for exponentiation discussed so far are silent on the subject of imaginary exponents. It is tempting to take our relation $e^x \simeq (1 + x/N)^N$ for finite $x$ and infinite $N$, postulate that it holds for imaginary exponents as well,

$$e^{ix} \simeq \left(1 + \frac{ix}{N}\right)^N, \tag{6}$$

and then perform algebraic operations on the right-hand side. We will give a definition of $e^{ix}$ very close to this one toward the end of the article, but in the mean time we avoid the difficulty of having to pin down the meaning of $e^{ix}$ by providing a form of the Euler identities that does not require the relation in (6), nor even mention of $e^{ix}$, but instead uses algebraic terms of the form $(1 + ix/N)^N$. We show that for all finite $x$ and infinite $N$,

$$\cos x \simeq \frac{1}{2}\left[\left(1 + \frac{ix}{N}\right)^N + \left(1 - \frac{ix}{N}\right)^N\right], \tag{7}$$

$$\sin x \simeq \frac{1}{2i}\left[\left(1 + \frac{ix}{N}\right)^N - \left(1 - \frac{ix}{N}\right)^N\right], \tag{8}$$

and then use these relations directly.

This maneuver frees us from having to define $e^{ix}$, but what about $(1 + ix/N)^N$? We still have to explain how complex numbers fit into the hyperreal framework. In elementary courses, the complex numbers are defined by starting with the real numbers and adjoining a new number $i$, together with the axiom $i^2 = -1$ and a "transfer principle" that says that the usual rules of algebra apply to the extended system of *complex* numbers. This method suits our purposes exactly, except that now we adjoin $i$ to the hyperreals rather than the reals, and call the resulting numbers the *hypercomplex* numbers. Every hypercomplex number can be written as $a + bi$ where $a$ and $b$ are hyperreal. For two hypercomplex numbers $c$ and $d$, we write $c \simeq d$ to mean that the modulus of their difference is infinitesimal (or equivalently, that their respective real and imaginary parts are infinitely close). We say that a hypercomplex number $c$ is infinitesimal if its modulus is infinitesimal, and finite if its modulus is finite. We note that the Binomial Theorem holds for hypercomplex binomials by the Transfer Principle, and that the Summation Comparison Theorem holds for series of hypercomplex terms by the same argument (given in a later section) as for series of hyperreal terms.

We begin by proving two standard formulas.

PROPOSITION. *For all real x and natural n,*

$$\cos x = \frac{1}{2}\left[\left(\cos\frac{x}{n} + i\sin\frac{x}{n}\right)^n + \left(\cos\frac{x}{n} - i\sin\frac{x}{n}\right)^n\right], \tag{9}$$

$$\sin x = \frac{1}{2i}\left[\left(\cos\frac{x}{n} + i\sin\frac{x}{n}\right)^n - \left(\cos\frac{x}{n} - i\sin\frac{x}{n}\right)^n\right]. \tag{10}$$

*Proof.* Using the familiar formulas for the sine and cosine of a sum of angles, one can show by induction that for all $n$ and $\theta$, $\cos n\theta = \frac{1}{2}[(\cos\theta + i\sin\theta)^n + (\cos\theta - i\sin\theta)^n]$ and $\sin n\theta = \frac{1}{2i}[(\cos\theta + i\sin\theta)^n - (\cos\theta - i\sin\theta)^n]$. Substituting $x/n$ for $\theta$ yields the result. ∎

We will obtain (7) and (8) from this proposition using small-angle approximations for the sine and cosine. If $\theta$ is an infinitesimal angle (that is, an angle subtending an

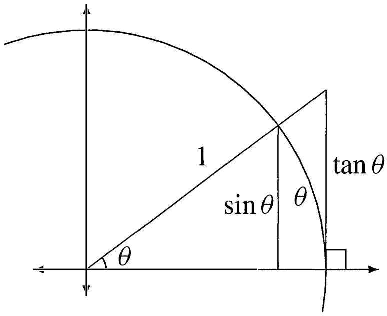



**Figure 2** $0 < \sin\theta < \theta < \tan\theta$ in the first quadrant

infinitesimal arc) then it is obvious from the inequality in FIGURE 2 that $\sin\theta \simeq \theta$ and hence that $\cos\theta = (1 - \sin^2\theta)^{1/2} \simeq 1$. But the presence of the exponents in (9) and (10) prevents us from using these results to take $n$ infinite and substitute 1 for $\cos(x/n)$ and $x/n$ for $\sin(x/n)$ to get (7) and (8). For these substitutions to be valid we need a sharper result that involves the notion of *relative infinitesimal*. For $\epsilon \neq 0$, we will say that *$a$ is infinitesimal with respect to $\epsilon$*, and write $a \simeq 0 \pmod{\epsilon}$, to mean that $a/\epsilon \simeq 0$. Similarly, $a \simeq b \pmod{\epsilon}$ means that $a - b \simeq 0 \pmod{\epsilon}$.

PROPOSITION. *If $0 < \theta < \pi/2$, then*

$$\theta - \frac{1}{2}\theta^3 < \sin\theta < \theta \quad \textit{and} \quad 1 - \frac{1}{2}\theta^2 < \cos\theta < 1.$$

*If $\theta$ is a nonzero infinitesimal, then*

$$\sin\theta \simeq \theta \pmod{\theta} \quad \textit{and} \quad \cos\theta \simeq 1 \pmod{\theta}.$$

*Proof.* Assume that $0 < \theta < \pi/2$; the case of negative $\theta$ will be an easy consequence. From geometry (see FIGURE 2) we have $\sin\theta < \theta < \tan\theta$. From $\theta < \tan\theta$ we deduce $\theta\cos\theta < \sin\theta$. By the double-angle formula, $\cos\theta = 1 - 2\sin^2(\frac{\theta}{2})$, we get $\theta(1 - 2\sin^2(\frac{\theta}{2})) < \sin\theta$. Since $\sin(\frac{\theta}{2}) < \frac{\theta}{2}$ in this range, we conclude (remarkably) that $\theta - \frac{1}{2}\theta^3 = \theta(1 - 2(\frac{\theta}{2})^2) < \theta(1 - 2\sin^2(\frac{\theta}{2})) < \sin\theta < \theta$, which implies $-\frac{1}{2}\theta^2 < (\sin\theta - \theta)/\theta < 0$. Then $\theta \simeq 0$ implies $(\sin\theta - \theta)/\theta \simeq 0$, and hence $\sin\theta \simeq \theta \pmod{\theta}$. For the cosine approximation, the formula $\cos\theta = 1 - 2\sin^2(\frac{\theta}{2})$ implies that $1 - \frac{1}{2}\theta^2 < 1 - 2\sin^2(\frac{\theta}{2}) = \cos\theta < 1$, and thus for $\theta \simeq 0$, we have $\cos\theta \simeq 1 \pmod{\theta}$. ∎

With this proposition we can now prove the theorem.

THEOREM. *For all finite $x$ and infinite $N$,*

$$\cos x \simeq \frac{1}{2}\left[\left(1 + \frac{ix}{N}\right)^N + \left(1 - \frac{ix}{N}\right)^N\right],$$

$$\sin x \simeq \frac{1}{2i}\left[\left(1 + \frac{ix}{N}\right)^N - \left(1 - \frac{ix}{N}\right)^N\right].$$

*Proof.* Let $x$ be finite and $N$ infinite. By the Transfer Principle applied to (9) and (10) we get

$$\cos x = \frac{1}{2}\left[\left(\cos\frac{x}{N} + i\sin\frac{x}{N}\right)^N + \left(\cos\frac{x}{N} - i\sin\frac{x}{N}\right)^N\right], \tag{11}$$

$$\sin x = \frac{1}{2i}\left[\left(\cos\frac{x}{N} + i\sin\frac{x}{N}\right)^N - \left(\cos\frac{x}{N} - i\sin\frac{x}{N}\right)^N\right], \tag{12}$$

where $x/N \simeq 0$. Since $\cos\frac{x}{N} \simeq 1 \pmod{\frac{x}{N}}$ and $\sin\frac{x}{N} \simeq \frac{x}{N} \pmod{\frac{x}{N}}$, we may write $\cos\frac{x}{N} = 1 + \frac{x}{N}\epsilon$ and $\sin\frac{x}{N} = \frac{x}{N} + \frac{x}{N}\delta$, where $\epsilon$ and $\delta$ are infinitesimals depending on $x$ and $N$ (take $\epsilon = (\cos\frac{x}{N} - 1)/\frac{x}{N}$ and $\delta = (\sin\frac{x}{N} - \frac{x}{N})/\frac{x}{N}$; these are infinitesimal by the previous proposition). Then

$$\cos\frac{x}{N} \pm i\sin\frac{x}{N} = 1 \pm \frac{ix + (\epsilon + i\delta)x}{N}. \tag{13}$$

By the Binomial Theorem and the Summation Comparison Theorem, one can easily show that $(1 \pm c/N)^N \simeq (1 \pm d/N)^N$ whenever $c$ and $d$ are finite, hypercomplex numbers that are infinitely close. Thus (13) implies that

$$\left(\cos\frac{x}{N} \pm i\sin\frac{x}{N}\right)^N = \left(1 \pm \frac{ix + (\epsilon + i\delta)x}{N}\right)^N \simeq \left(1 \pm \frac{ix}{N}\right)^N,$$

which by (11) and (12) yields the result. ∎

The familiar series for sine and cosine can now be obtained by applying the Binomial Theorem to "multiply out" the $N^{\text{th}}$ powers in (7) and (8), and then applying the Summation Comparison Theorem. This proves the following theorem.

THEOREM. *For all finite $x$ and infinite $H$,*

$$\cos x \simeq 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \pm \frac{x^{2H}}{(2H)!},$$

$$\sin x \simeq x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \pm \frac{x^{2H+1}}{(2H+1)!}.$$

## The binomial series

Many times already we have used the formula,

$$(1+x)^m = 1 + mx + \frac{m^{\underline{2}}}{2!}x^2 + \frac{m^{\underline{3}}}{3!}x^3 + \cdots + \frac{m^{\underline{m}}}{m!}x^m, \quad \text{natural } m,$$

a result that was known centuries prior to Euler (though not in this notation), and which can be verified using induction. In 1665 Newton discovered a generalization of the coefficients of the binomial expansion using a complicated interpolation between the rows and columns of a tabular form of Pascal's triangle, and conjectured that these generalized coefficients could be used to obtain a binomial series for negative and fractional exponents. Newton tested the conjecture on many examples, including squaring the series for $(1+x^2)^{1/2}$,

$$\begin{aligned}&(1+x^2)^{1/2}(1+x^2)^{1/2}\\ &\quad = \left(1 + \frac{1}{2}x^2 + \frac{\frac{1}{2}(\frac{1}{2}-1)}{2!}x^4 + \cdots\right)\left(1 + \frac{1}{2}x^2 + \frac{\frac{1}{2}(\frac{1}{2}-1)}{2!}x^4 + \cdots\right)\end{aligned}$$

to obtain

$$1 + x^2 + 0x^4 + 0x^6 + 0x^8 + 0x^{10} + \cdots,$$

but he never published a deductive proof of the general formula. (Edwards [**9**, pp. 178–187] gives a detailed account of the discovery of the Binomial Theorem; see also [**25**, p. 438].) Euler states Newton's "universal theorem" [**12**, §71],

$$(P+Q)^{\frac{m}{n}} = P^{\frac{m}{n}} + \frac{m}{n}P^{\frac{m-n}{n}}Q + \frac{m(m-n)}{n\cdot 2n}P^{\frac{m-2n}{n}}Q^2 + \frac{m(m-n)(m-2n)}{n\cdot 2n\cdot 3n}P^{\frac{m-3n}{n}}Q^3 + \cdots$$

but omits the proof "since it can be done so much more easily with the aid of some principles of differential calculus" [**12**, §76]. Surprisingly, we will show that the proof of the Binomial Theorem for fractional exponents—which we write as

$$(1+x)^{\frac{m}{n}} \simeq 1 + \frac{m}{n}x + \left(\frac{m}{n}\right)^{\underline{2}}\frac{x^2}{2!} + \left(\frac{m}{n}\right)^{\underline{3}}\frac{x^3}{3!} + \cdots + \left(\frac{m}{n}\right)^{\underline{H}}\frac{x^H}{H!},$$

where $H$ is infinite, $|x| < 1$, and $x \not\simeq 1$—is well within the scope of this article, and forms a natural part of our rehabilitation of Euler's methods. The proof uses the Binomial Theorem for Factorial Powers, which can be verified by induction. (See [**16**] for other uses of factorial powers.)

BINOMIAL THEOREM FOR FACTORIAL POWERS. *For all real $a$, $b$ and all natural $n$,*

$$(a+b)^{\underline{n}} = \sum_{k=0}^{n} \frac{n^{\underline{k}}}{k!} a^{\underline{n-k}} b^{\underline{k}}.$$

BINOMIAL THEOREM (FRACTIONAL EXPONENTS). *If $|x| < 1$, and $m$ and $n$ are finite and positive, and $H$ is infinite, then*

$$(1+x)^{m/n} \simeq 1 + \frac{m}{n}x + \left(\frac{m}{n}\right)^{\underline{2}}\frac{x^2}{2!} + \cdots + \left(\frac{m}{n}\right)^{\underline{H}}\frac{x^H}{H!}.$$

*Proof.* Fix an infinite hypernatural $H$ and a hyperreal $x$ such that $|x| < 1$ and $x \not\simeq 1$. We introduce the notation $(1+x)^{\boxed{a}}$ for the sum,

$$(1+x)^{\boxed{a}} \stackrel{\text{def}}{=} 1 + ax + a^{\underline{2}}\frac{x^2}{2!} + \cdots + a^{\underline{H}}\frac{x^H}{H!}.$$

(The dependence on $H$ is not explicit in our notation.) Generalizing Newton's calculation for $(1+x^2)^{1/2}$, we will show that

$$\left((1+x)^{\boxed{m/n}}\right)^n \simeq (1+x)^m$$

and hence that

$$(1+x)^{\boxed{m/n}} \simeq (1+x)^{m/n}, \tag{14}$$

which is the statement of our theorem in our new notation.

The key formula in the proof of (14) is that for finite positive $a$ and $b$,

$$(1+x)^{\boxed{a+b}} \simeq (1+x)^{\boxed{a}}(1+x)^{\boxed{b}}. \tag{15}$$

This is easy to see for integral $m$, $n$:

$$(1+x)^{\boxed{m+n}} = (1+x)^{m+n} = (1+x)^m(1+x)^n = (1+x)^{\boxed{m}}(1+x)^{\boxed{n}}.$$

For the general case we make use of the Binomial Theorem for Factorial Powers, after first multiplying out, gathering like terms, and neglecting the tail. Using the lemma (following this theorem), we write

$$\begin{aligned}
&(1+x)^{\boxed{a}}(1+x)^{\boxed{b}} \\
&= \left(1+ax+a^{\underline{2}}\frac{x^2}{2!}+\cdots+a^{\underline{H}}\frac{x^H}{H!}\right)\left(1+bx+b^{\underline{2}}\frac{x^2}{2!}+\cdots+b^{\underline{H}}\frac{x^H}{H!}\right) \\
&= \sum_{k=0}^{2H} c_k x^k, \quad \text{where } c_k = \sum_{i+j=k} \frac{a^{\underline{i}}}{i!}\frac{b^{\underline{j}}}{j!}.
\end{aligned}$$

Observe that for all *finite* $k$,

$$c_k = \sum_{i+j=k} \frac{a^{\underline{i}}}{i!}\frac{b^{\underline{j}}}{j!} = \sum_{j=0}^{k} \frac{a^{\underline{k-j}}}{(k-j)!}\frac{b^{\underline{j}}}{j!} = \frac{1}{k!}\sum_{j=0}^{k} \frac{k^{\underline{j}}}{j!} a^{\underline{k-j}} b^{\underline{j}} = \frac{1}{k!}(a+b)^{\underline{k}}.$$

For the first step in this chain of equalities it is essential that $k$ be finite; the last step follows from the Binomial Theorem for Factorial Powers. In the lemma (following) we will verify that both $\sum_{k=0}^{2H} c_k x^k$ and $\sum_{k=0}^{H}(a+b)^{\underline{k}}x^k/k!$ are determinate; if for the moment we assume this as fact, then by the Summation Comparison Theorem, we may conclude that

$$(1+x)^{\boxed{a}}(1+x)^{\boxed{b}} = \sum_{k=0}^{2H} c_k x^k \simeq \sum_{k=0}^{H} \frac{(a+b)^{\underline{k}}}{k!}x^k = (1+x)^{\boxed{a+b}},$$

which shows (15). If $m$ and $n$ are finite, then by applying (15) a total of $n$ times, we get

$$(1+x)^m = (1+x)^{\boxed{m}} = (1+x)^{\boxed{\underbrace{\frac{m}{n}+\cdots+\frac{m}{n}}_{n \text{ terms}}}} \simeq \left((1+x)^{\boxed{m/n}}\right)^n,$$

and hence that $(1+x)^{m/n} \simeq (1+x)^{\boxed{m/n}}$, as required. ■

The Binomial Theorem can be extended to negative rational exponents by a similar argument, and, by the Sequential Theorem (see the next section) to the case where $m$ and $n$ are infinite, so long as $m/n$ is finite. From there it is but a very small step to the theorem for real exponents. This is left as an exercise for the reader.

The previous theorem requires the following lemma.

LEMMA. ***(i)*** *The series* $1+|a|+|a^{\underline{2}}x^2/2!|+|a^{\underline{3}}x^3/3!|+\cdots$ *is determinate for finite positive* $a$, $|x|<1$, $x \not\simeq 1$. ***(ii)*** *If* $a_0, a_1, a_2, \ldots$ *and* $b_0, b_1, b_2, \ldots$ *are hypersequences then for all hypernatural* $H$,

$$\left(\sum_{i=0}^{H} a_i\right)\left(\sum_{j=0}^{H} b_i\right) = \sum_{k=0}^{2H} c_k \quad \textit{where} \quad c_k = \sum_{i+j=k} a_i b_j.$$

*Moreover, if* $|a_0| + |a_1| + |a_2| + \cdots$ *and* $|b_0| + |b_1| + |b_2| + \cdots$ *are determinate and have finite partial sums, then* $|c_0| + |c_1| + |c_2| + \cdots$ *is determinate and has finite partial sums.*

*Proof.* **(i)** We ask the reader to verify that for integral $k > a > 0$, we have $|a^{\underline{k}}| < k!$, and hence that $|a^{\underline{k}} x^k / k!| \le |x^k|$. This shows that for $k > a > 0$ the absolute values of the terms in $(1 + x)^{\boxed{a}}$ are bounded by a determinate geometric series. **(ii)** Note that the product $(\sum_{i=0}^{H} a_i)(\sum_{j=0}^{H} b_j)$ when multiplied out, is the sum of all terms $a_i b_j$ for $i$ and $j$ between 0 and $H$. These terms can be arranged in a table.

| | 0 | 1 | 2 | $\cdots$ | $H$ |
|---|---|---|---|---|---|
| 0 | $a_0b_0$ | $a_0b_1$ | $\mathbf{a_0b_2}$ | $\cdots$ | $a_0b_H$ |
| 1 | $a_1b_0$ | $\mathbf{a_1b_1}$ | $a_1b_2$ | $\cdots$ | $a_1b_H$ |
| 2 | $\mathbf{a_2b_0}$ | $a_2b_1$ | $a_2b_2$ | $\cdots$ | $a_2b_H$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $H$ | $a_Hb_0$ | $a_Hb_1$ | $a_Hb_2$ | $\cdots$ | $a_Hb_H$ |

$$\mathbf{c_2 = a_2b_0 + a_1b_1 + a_0b_2}$$

For each $k$, $c_k = \sum_{i+j=k} a_i b_j$ is the sum of all terms on the northeasterly diagonal at $k$, making $\sum_{k=0}^{2H} c_k$ the sum of all of the diagonals, and hence the sum of all terms in the table. To see that $|c_0| + |c_1| + |c_2| + \cdots$ is determinate, observe that for $N$ infinite,

$$\sum_{k=N}^{N+M} |c_k| \le \sum_{k=N}^{N+M} \sum_{i+j=k} |a_i b_j| \le \sum_{i=\frac{N}{2}}^{H} \sum_{j=0}^{H} |a_i b_j| + \sum_{j=\frac{N}{2}}^{H} \sum_{i=0}^{H} |a_i b_j|$$

$$= \left(\sum_{i=\frac{N}{2}}^{H} |a_i|\right)\left(\sum_{j=0}^{H} |b_j|\right) + \left(\sum_{i=0}^{H} |a_i|\right)\left(\sum_{j=\frac{N}{2}}^{H} |b_j|\right) \simeq 0.$$

The second inequality is obtained by noting that if $i + j \ge N$, then either $i \ge N/2$ or $j \ge N/2$. The final step ($\simeq 0$) follows because the series are determinate and have finite partial sums. ■

## Proof of the Summation Comparison Theorem

In contrast with the other theorems in this article, which concern concrete functions and equations, the Summation Comparison Theorem is a result about all functions and equations of a general class. It should not be surprising then that the proof is more abstract and relies on more basic definitions and principles than the proofs of the other theorems. We will show how the Summation Comparison Theorem follows from the Least Counterexample Principle, an equivalent of the Hypernatural Induction Principle, by way of the Sequential Theorem. The Sequential Theorem is an important result due to Robinson [**40**, Theorem 3.3.20]. In a course of study, the proof could be delayed.

LEAST COUNTEREXAMPLE PRINCIPLE. *Let* $\phi(n)$ *be an equation or inequality of hypersequences; that is, let* $\phi(n)$ *be of the form* $a_n = b_n$, $a_n \ne b_n$, $a_n < b_n$, *or* $a_n \le b_n$, *where* $a$ *and* $b$ *are hypersequences. Then either* $\phi(n)$ *holds for all hypernatural* $n$, *or else there is an* $m$ *such that* $\phi(m)$ *fails but such that* $\phi(n)$ *holds for all hypernatural* $n$ *less than* $m$.

For example, consider the inequality $\phi(n)$ given by $(1 - n/H) > 0$, for a fixed hypernatural $H$. This inequality is false for $n$ equal to $H$, but it is true for all $n$ less than $H$. Thus $n = H$ is a least counterexample for $\phi(n)$. On the other hand, the equation $a_n = 0$, where $a_n$ is defined to be 0 for $n$ finite and 1 for $n$ infinite, does not have a least counterexample, even though there are counterexamples. It does not however disobey the Least Counterexample Principle, because the function $a$, so defined, is not a hypersequence (that is, $a$ cannot be obtained by composition of natural extensions of real functions with hyperreal parameters).

SEQUENTIAL THEOREM. *Let $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$ be hypersequences. If $a_n \simeq b_n$ for all natural n, then there is an infinite N such that $a_n \simeq b_n$ for all hypernatural n smaller than N.*

*Proof.* Since the relation $a_n \simeq b_n$ is not an equation or inequality, one cannot apply the Least Counterexample Principle directly. Instead, we note that if $a_n \simeq b_n$ for all finite $n$, then it is also true that $-\frac{1}{n} < a_n - b_n$ and $a_n - b_n < \frac{1}{n}$ for all finite $n > 1$. By applying the Least Counterexample Principle to these inequalities, we conclude that there is an infinite $N$ such that $-\frac{1}{n} < a_n - b_n$ and $a_n - b_n < \frac{1}{n}$ for all $n$ between 1 and $N$. By the original assumption that $a_n \simeq b_n$ for all finite $n$, and using the fact that $1/n$ is infinitesimal for infinite values of $n$, we conclude that $a_n \simeq b_n$ for all $n < N$. ■

SUMMATION COMPARISON THEOREM. *If the series $a_0 + a_1 + a_2 + \cdots$ and $b_0 + b_1 + b_2 + \cdots$ are determinate, and if for each natural n, $a_n \simeq b_n$, then for all hypernatural n, $a_0 + a_1 + \cdots + a_n \simeq b_0 + b_1 + \cdots + b_n$.*

*Proof.* If $a_n \simeq b_n$ for all finite $n$ then $a_0 + \cdots + a_n \simeq b_0 + \cdots + b_n$ for all finite $n$ as well. By the Sequential Theorem, there is an infinite $J$ such that for all $n$ less than $J$, $a_0 + \cdots + a_n \simeq b_0 + \cdots + b_n$. Let $N$ be greater than $J$. If the sums are determinate, then by definition, $a_J + \cdots + a_N$ and $b_J + \cdots + b_N$ are both infinitesimal, and hence *for all n*, $a_0 + \cdots + a_n \simeq b_0 + \cdots + b_n$. ■

## The logarithm and beyond

Immediately after defining the exponential function $a^x$ and discussing the basic rules for exponentiation, Euler defined the logarithm for bases $a$ greater than 1.

> Just as, given a number $a$, for any value of $x$, we can find a value of $y$ [$= a^x$], so, in turn, given a positive value for $y$, we would like to give a value for $x$, such that $a^x = y$. This value of $x$, insofar as it is viewed as a function of $y$, is called the *logarithm* of $y$. ... It is customary to designate the logarithm of $y$ by the symbol, $\log y$. [**12**, §102]

We usually write $\log_a y$, making the dependence on the base $a$ explicit in the notation. From the definition that for $y > 0$, $\log_a y$ is the $x$ such that $a^x = y$, and from the rules for exponentials given earlier, the following rules for logarithms (for $a > 1$ and $x, y > 0$) follow immediately: $\log_a 1 = 0$, $\log_a x^{-1} = -\log_a x$, $\log_a(xy) = \log_a x + \log_a y$, $\log_a x^y = y \log_a x$, and $\log_a x < \log_a y$ iff $x < y$.

Early in the *Introductio*, Euler explained how these properties of the logarithm, insofar as they reduce exponentiation to multiplication and multiplication to addition, make compiled tables of logarithms extremely useful for performing computations. One of his examples, employing a table of logarithms to the base 10, is as follows.

> If the population in a certain region increases annually by one-thirtieth and at one time there were 100,000 inhabitants, we would like to know the population

> after 100 years. For the sake of brevity, we let the initial population be $n$, so that $n = 100{,}000$. After one year the new population will be $(1+\frac{1}{30})n = \frac{31}{30}n$. After two years it will equal $(\frac{31}{30})^2 n$. After three years it will equal $(\frac{31}{30})^3 n$. Finally after one-hundred years the population will be $(\frac{31}{30})^{100} n = (\frac{31}{30})^{100} 100{,}000$. The logarithm of this population is $100 \log \frac{31}{30} + \log 100{,}000$. But $\log \frac{31}{30} = \log 31 - \log 30 = 0.014240439$, so that $100 \log \frac{31}{30} = 1.4240439$, which, when increased by $\log 100{,}000 = 5$, gives 6.424039, the logarithm of the desired population. The corresponding population is 2,654,874. So after one-hundred years the population will be more than twenty-six-and-a-half times as large. [**12**, §110]

The question remains: how might one compile such tables of logarithms? Euler gave an example to show how Briggs computed logarithms for his famous *Arithmetica Logarithmica* of 1624 using a calculation-intensive algorithm that required the manual extraction of many successive square roots, but noted that, "In the mean time, much shorter methods have been found by means of which logarithms can be computed more quickly." [**12**, §106] In the succeeding chapter, Euler explained how to calculate logarithms using series.

Let us find a series for the logarithm with base $e$. This logarithm can be written $\log_e$ but is more often written ln. To get started, we relax the requirement that for a given $y$ we must find *the* $x$ such that $e^x$ is *exactly equal* to $y$, and instead try to find, for a given finite $y > 0$,

$$\text{an } x \text{ such that } e^x \simeq y. \tag{16}$$

Earlier we found that for finite $x$ and infinite $N$, $e^x \simeq (1 + x/N)^N$, so let us solve the equation $y = (1 + x/N)^N$ and see whether that solution is of any use. A solution (taking the principal $N^{\text{th}}$ root of $y$) is given by

$$x = N(y^{1/N} - 1).$$

Observe that this does satisfy (16),

$$e^x = e^{N(y^{1/N}-1)} \simeq \left(1 + \frac{N(y^{1/N}-1)}{N}\right)^N = (1 + y^{1/N} - 1)^N = (y^{1/N})^N = y,$$

so that indeed, $e^x \simeq y$. In our discussion of the logarithm, we also need to know that $N(y^{1/N} - 1) \simeq \ln y$.

THEOREM. *If $y$ is finite and positive, then* $\ln y \simeq N(y^{1/N} - 1)$ *for all infinite $N$.*

*Proof.* Let $y$ be finite and positive. By the preceding computation, $e^{N(y^{1/N}-1)} \simeq y = e^{\ln y}$. Then $N(y^{1/N} - 1) \simeq \ln y$ follows from the proposition saying that $x \simeq y$ if and only if $e^x \simeq e^y$ for finite $x$ and $y$. ∎

In [**12**, §119], Euler used the formula $\ln y = N(y^{1/N} - 1)$, for $N$ infinite, to derive the series for the natural logarithm. He expanded the function $\log(1 + y)$ using the Binomial Theorem for fractional exponents to get

$$\begin{aligned}
&\log(1+y)\\
&= N\left[(1+y)^{1/N} - 1\right]\\
&= N\left[\left[1 + \frac{1}{N}y + \frac{1}{N}\left(\frac{1}{N}-1\right)\frac{1}{2!}y^2 + \frac{1}{N}\left(\frac{1}{N}-1\right)\left(\frac{1}{N}-2\right)\frac{1}{3!}y^3 + \cdots\right] - 1\right]y\\
&= y + \left(\frac{1}{N}-1\right)\frac{1}{2!}y^2 + \left(\frac{1}{N}-1\right)\left(\frac{1}{N}-2\right)\frac{1}{3!}y^3 + \cdots.
\end{aligned}$$

Then, using the fact that $N$ is infinite, Euler substituted 0 everywhere for $1/N$, and obtained the equation,

$$\log(1+y) = y - \frac{1}{2}y^2 + \frac{1}{3}y^3 - \cdots.$$

Such substitutions are somewhat more difficult to justify for this series than for our earlier examples, but it is nonetheless within reach of the methods we have discussed thus far.

THEOREM. *For all $y$ with $|y| < 1$ but $|y| \not\simeq -1$, and all infinite $H$,*

$$\log(1+y) \simeq y - \frac{1}{2}y^2 + \frac{1}{3}y^3 - \cdots + (-1)^{H+1}\frac{1}{H}y^H.$$

*Proof.* Assume that $|y| < 1$, $|y| \not\simeq -1$, and let $H$ be infinite. By the Binomial Theorem for fractional exponents, we conclude that for all *finite* $n$,

$$(1+y)^{1/n} \simeq 1 + \frac{1}{n}y + \left(\frac{1}{n}\right)^{\underline{2}}\frac{1}{2!}y^2 + \cdots + \left(\frac{1}{n}\right)^{\underline{H}}\frac{1}{H!}y^H. \tag{17}$$

For $n$ infinite, both sides are infinitely close to 1, so (17) is actually true for all hypernatural $n$, infinite as well as finite. Thus it is tempting to follow Euler's lead and substitute an infinite $N$ for $n$ in (17), subtract 1 from both sides, then multiply by $N$. We cannot quite do this, because (17) has "$\simeq$" rather than "$=$", and because for $N$ infinite it does not follow from $a \simeq b$ that $Na \simeq Nb$ (for a counterexample take $a = 0$, $b = 1/N$). Instead we can do this: (17) implies that for all *finite* $n$,

$$\begin{aligned} &n\left[(1+y)^{1/n} - 1\right] \\ &\quad\simeq n\left[\left(1 + \frac{1}{n}y + \left(\frac{1}{n}\right)^{\underline{2}}\frac{y^2}{2!} + \left(\frac{1}{n}\right)^{\underline{3}}\frac{y^3}{3!} + \cdots + \left(\frac{1}{n}\right)^{\underline{H}}\frac{y^H}{H!}\right) - 1\right] \\ &\quad= y - \frac{\left(1-\frac{1}{n}\right)}{1}\cdot\frac{y^2}{2} + \frac{\left(1-\frac{1}{n}\right)}{1}\cdot\frac{\left(2-\frac{1}{n}\right)}{2}\cdot\frac{y^3}{3} \\ &\qquad - \cdots + (-1)^{H-1}\frac{\left(1-\frac{1}{n}\right)}{1}\cdot\frac{\left(2-\frac{1}{n}\right)}{2}\cdots\frac{\left(H-1-\frac{1}{n}\right)}{H-1}\cdot\frac{y^H}{H}, \end{aligned}$$

where the alternation in signs follows from the fact that for $k > 0$ and $n > 1$ we have $(1/n)^{\underline{k}} = (-1)^{k-1}\cdot\frac{1}{n}(1-\frac{1}{n})(2-\frac{1}{n})\cdots((k-1)-\frac{1}{n})$. Hence by the Sequential Theorem we conclude—not for all—but that for all *sufficiently small* infinite $N$,

$$\begin{aligned} &N\left[(1+y)^{1/N} - 1\right] \\ &\quad\simeq y - \frac{\left(1-\frac{1}{N}\right)}{1}\cdot\frac{y^2}{2} + \frac{\left(1-\frac{1}{N}\right)}{1}\cdot\frac{\left(2-\frac{1}{N}\right)}{2}\cdot\frac{y^3}{3} \\ &\qquad - \cdots + (-1)^{H-1}\frac{\left(1-\frac{1}{N}\right)}{1}\cdot\frac{\left(2-\frac{1}{N}\right)}{2}\cdots\frac{\left(H-1-\frac{1}{N}\right))}{H-1}\cdot\frac{y^H}{H}. \end{aligned}$$

When this last sum is compared with the sum

$$y - \frac{1}{2}y^2 + \frac{1}{3}y^3 - \cdots + (-1)^{H+1}\frac{1}{H}y^H,$$

it is clear that term by term the sums are infinitely close, so we need only verify that both sums are determinate. For $|y| < 1$, $|y| \not\simeq -1$, determinacy follows from the Comparison Test for Determinacy by comparison with a determinate geometric sum. By the Summation Comparison Theorem we finally conclude that

$$H\left[(1+y)^{1/H} - 1\right] \simeq y - \frac{1}{2}y^2 + \frac{1}{3}y^3 - \cdots + (-1)^{H+1}\frac{1}{H}y^H$$

*for all* infinite $H$. The result now follows from the previous theorem. ■

Finally, Euler observed that though the series for the logarithm just given does not converge rapidly, and hence is not itself so effective for computing logarithms, it leads to other series that are quite effective. For example,

$$\log\left(\frac{1+x}{1-x}\right) = \log(1+x) - \log(1-x) = 2x + \frac{2}{3}x^3 + \frac{2}{5}x^5 + \cdots.$$

Euler remarked,

> This last series is strongly convergent if we substitute an extremely small fraction for $x$. For instance, if $x = \frac{1}{5}$, then $\log\frac{6}{4} = \log\frac{3}{2} = \frac{2}{1\cdot 5} + \frac{2}{3\cdot 5^2} + \frac{2}{5\cdot 5^5} + \frac{2}{7\cdot 5^7} + \cdots$. If $x = \frac{1}{7}$, then $\log\frac{4}{3} = \frac{2}{1\cdot 7} + \frac{2}{3\cdot 7^2} + \frac{2}{5\cdot 7^5} + \frac{2}{7\cdot 7^7} + \cdots$, and if $x = \frac{1}{9}$, then $\log\frac{5}{4} = \frac{2}{1\cdot 9} + \frac{2}{3\cdot 9^2} + \frac{2}{5\cdot 9^5} + \frac{2}{7\cdot 9^7} + \cdots$. From the logarithms of these fractions, we can find the logarithms of integers. From the nature of logarithms we have $\log\frac{3}{2} + \log\frac{4}{3} = \log 2$, and $\log\frac{3}{2} + \log 2 = \log 3$, and $2\log 2 = \log 4$. Furthermore we have $\log\frac{5}{4} + \log 4 = \log 5$, $\log 2 + \log 3 = \log 6$, $3\log 2 = \log 8$, $2\log 3 = \log 9$, $\log 2 + \log 5 = \log 10$. [**12**, §123]

Using these series and relationships, Euler was able to show how to begin constructing a table of logarithms.

In the *Introductio*, Euler also exhibited series for other transcendental functions, including the tangent, cotangent, and arctangent, and went on to show how to use infinite products to compute the values of infinite sums. Using infinitesimal methods similar to those described here, Euler factored the sine function into an infinite product and used that factorization to deduce the celebrated formula $1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \cdots = \frac{\pi^2}{6}$. Both of these theorems can be rehabilitated, but the algebra turns out to be more taxing.

THEOREM. *For all finite $x$ and infinite $H$,*

$$\sin x \simeq x\prod_{k=1}^{H}\left(1 - \frac{x^2}{(k\pi)^2}\right).$$

THEOREM. *For all infinite $H$,*

$$\sum_{k=1}^{H}\frac{1}{k^2} \simeq \frac{\pi^2}{6}.$$

A careful analysis of Euler's arguments for these two results is given in [**37**].

## The connection between standard and nonstandard notions

Our theorem saying that $e^x \simeq \sum_{n=0}^{N} \frac{x^n}{n!}$, for all finite $x$ and infinite $N$, is conceptually similar to the standard theorem that says

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

for all real $x$, but these theorems are not the same. The former refers to a hypernatural summation within a proper extension of the real numbers while the latter refers to something we have not yet discussed: the limit of a real sequence of partial sums. The notion of *limit* is usually taken to be the dividing line between algebra and analysis. In this section we give a brief sketch of how to cross that line. Let us first recall the standard definition of convergence for infinite series.

STANDARD DEFINITION OF CONVERGENCE OF SERIES. Let $s$ be a real sequence and let $r$ be real. We say that $s$ *converges* to $r$ if and only if for each positive real $\epsilon$ there is a natural $n$ such that for all natural $m$ greater than $n$, $|s_m - r| < \epsilon$. We write $\sum_{n=0}^{\infty} a_n = r$ to mean that $a$ is a real sequence such that its sequence of partial sums $a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots$ converges to the real number $r$.

Except for our definition of the real number $e$ given above, thus far we have not had to distinguish between the closely related notions of *real number* and *finite hyperreal number*. But we need this distinction if we are to convert our results about hyperreal numbers to results solely about real numbers. The real numbers are distinguished from other ordered fields by the Completeness Axiom. We will not prove this here but the Standard Part Principle is actually a consequence of the Completeness Axiom for the real numbers. (See Keisler [**23**, pp. 36–40, 908–909].)

STANDARD PART PRINCIPLE. *For every finite hyperreal $b$ there is a unique real $r$ such that $r \simeq b$. This real $r$ is called the* standard part *of $b$, denoted $^{\circ}b$.*

Earlier we used the Standard Part Principle to define the real number $e$ to be $^{\circ}(1 + 1/N)^N$, where $N$ is infinite. This definition and (5) together with the assumption that $e^x$ is a real function, imply that

$$e^x = {}^{\circ}\sum_{n=0}^{N} \frac{x^n}{N!} \tag{18}$$

for infinite $N$. Rather than assuming $e^x$ to be defined for all real numbers (as we did above) one could instead take (18) to be the definition of the function $e^x$ and derive the usual properties of exponentiation from this definition. We chose not to do that, but it is a reasonable alternative. However, for *complex* exponentiation the synthetic approach is all we have at our disposal, so we simply define $e^{ix}$ by the identity

$$e^{ix} = {}^{\circ}\left(1 + \frac{ix}{N}\right)^N,$$

for real (and hyperreal) $x$, where $^{\circ}(a + bi) = (^{\circ}a) + (^{\circ}b)i$. From this definition one can deduce the Euler identities in their familiar form.

COROLLARY. *For all real $x$,*

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \textit{and} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

The connection between the infinite sum of a determinate hypersequence and the convergence of a real sequence of partial sums is given by the following theorem, which is a consequence of the Transfer Principle and the Least Counterexample Principle. The proof, although somewhat technical, is within the scope of Keisler's calculus book.

THEOREM. *Let $\beta$ be a hypersequence such that $\beta_0 + \beta_1 + \beta_2 + \cdots$ is determinate, let $b$ be a real sequence, and suppose that $b_n \simeq \beta_n$ for all natural $n$. Then the real sequence of partial sums of $b$ is convergent in the standard sense, and for all infinite hypernatural $N$, $\sum_{n=0}^{\infty} b_n = {}^{\circ}\sum_{n=0}^{N} \beta_n$.*

(Note that the convergence of $b$ is a consequence, not a hypothesis.) This theorem implies standard analogs of all the summations in this article.

COROLLARIES. *For all real $x$,*

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \qquad \log(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} \quad \textit{for } |x| < 1,$$

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \qquad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!},$$

$$(1+x)^{m/n} = \sum_{k=0}^{\infty} \left(\frac{m}{n}\right)^{\underline{k}} \frac{x^k}{k!} \quad \textit{for } |x| < 1.$$

*Proof.* For the first equation, let $N$ be infinite. Then $e^x = {}^{\circ}\sum_{n=0}^{N} \frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. The others are similar. ∎

## Lessons from Euler

The *Introductio* was expressly intended as a precalculus textbook, that is, a book for a course of study prior to differential and integral calculus. The point was not to give short and slick derivations from an extensive body of knowledge, but rather to educate beginners. Euler said,

> Although all of these nowadays are accomplished by means of differential calculus, nevertheless, I have here presented them using only ordinary algebra, in order that the transition from finite analysis to analysis of the infinite might be rendered easier.... At the same time I readily admit that these matters can be much more easily worked out by differential calculus. [**12**, pp. *ix–x*]

We might take a lesson from Euler's great textbook for our own courses. In the standard treatments, discrete mathematics is held disjoint from the calculus, and interesting and useful series are studied only after Taylor's Theorem is proved—usually at the end of the lectures on convergence of sequences and series, well after the derivative is thoroughly studied. In Euler's treatment, beginners get their hands on concrete examples of sequences and series even before the derivative is defined. As rehabilitated here, this approach might also give our own students practice with important topics from discrete mathematics—induction, recursion, finite summations, and axiomatics—in the course of proving elementary analogs of theorems of real analysis. But whether or not our rehabilitation of Euler's methods finds its way into the educational main stream, we hope that by focusing our attention on the intellectual beauty of the underlying mathematics, we have convinced the reader that Euler's insights and arguments, far from being reckless or nonsensical, are directly relevant to the understanding, appreciation, and application of elementary mathematics even in our day.

**Acknowledgment.** An outline of this article was presented to the *Colóquio Internacional de Matemática Não-Standard* (Aveiro, Portugal, 1994); see [**36**]. The authors thank Richard Askey, Michael Benedikt, Michael Bleicher, Dan Kalman, and Jerome Keisler, for sharing many insights on this subject.

## REFERENCES

1. C. B. Boyer, The foremost textbook of modern times, *Amer. Math. Monthly* **58** (1951), 223–226.
2. R. C. Buck, Mathematical induction and recursive definitions, *Amer. Math. Monthly* **70** (1963), 128–135.
3. F. Cajori, *A History of Mathematical Notations*, vol. 2, Open Court, Chicago, 1929. Republished by Dover, 1993.
4. J. L. Coolidge, The number *e*, *Amer. Math. Monthly* **57** (1950), 591–602.
5. N. Cutland, V. Neves, A. F. Oliviera, and J. Sousa-Pinto (eds.), *Developments in Nonstandard Mathematics*, Pitman Research Notes in Mathematics 336, Longman House, Harlow, Essex, 1995.
6. M. Davis and R. Hersh, Nonstandard analysis, *Sci. Amer.* **226** (1972), 78–86.
7. W. Dunham, *Journey through Genius*, Penguin Books, New York, 1991.
8. W. Dunham, *Euler: The Master of Us All*, Dolciani Mathematical Expositions **22**, Math. Assoc. of Amer., 1999.
9. C. H. Edwards, Jr., *The Historical Development of the Calculus*, Springer-Verlag, New York, 1979.
10. G. Eneström, Über eine von Euler augestellte allgemeine Konvergenzbedingung, *Bibliotheca Mathematica* ser. 3 vol. **6** (1905), 186–189.
11. L. Euler, De progressionibus harmonicus observationes, *Commentarii academiæ scientarium Petropolitanæ* **9** (1734), 150–161. Reprinted in L. Euler, *Opera Omnia*, ser. 1, vol. **14**, 87–100.
12. L. Euler, *Introductio in Analysin Infinitorum*, Tomus Primus, Lausanne, 1748. Reprinted as L. Euler, *Opera Omnia*, ser. 1, vol. 8. Translated from the Latin by J. D. Blanton, *Introduction to Analysis of the Infinite*, Book I, Springer–Verlag, New York, 1988.
13. H. Eves, *An Introduction to the History of Mathematics*, 6th ed., Saunders College Publishing, Philadelphia, 1990.
14. E. Farkas and M. Szabo, On the plausibility of nonstandard proofs in analysis, *Dialectica* **38** (1984), 297–310.
15. J. Fenstad, Is nonstandard analysis relevant for the philosophy of mathematics?, *Synthese* **62** (1985), 289–302.
16. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison–Wesley, Reading, MA, 1989.
17. G. H. Hardy, Some notes on certain theorems in higher trigonometry, *Math. Gaz.* **56** (1906), 284–288.
18. J. Henle and E. Kleinberg, *Infinitesimal Calculus*, MIT Press, Cambridge, MA, 1979.
19. C. W. Henson, M. Kaufmann, and H. J. Keisler, The strength of nonstandard methods in arithmetic, *J. Symbolic Logic* **49** (1984), 1039–1058.
20. E. W. Hobson, *A Treatise on Plane Trigonometry*, 7th ed., Cambridge University Press, Cambridge, 1939.
21. A. E. Hurd and P. A. Loeb, *An Introduction to Nonstandard Real Analysis*, Academic Press, Orlando, 1985.
22. H. J. Keisler, *Foundations of Infinitesimal Calculus*, Prindle, Weber, and Schmidt Publishers, Boston, 1976.
23. H. J. Keisler, *Elementary Calculus: An Infinitesimal Approach*, 2d ed., Prindle, Weber, and Schmidt Publishers, Boston, 1986.
24. H. J. Keisler, The hyperreal line, *Real Numbers, Generalizations of the Reals, and Theories of Continua*, P. Ehrlich, editor, Kluwer Academic, Dordrecht, 1994.
25. M. Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, New York, 1972.
26. I. Lakatos, Cauchy and the continuum: the significance of nonstandard analysis for the history and philosophy of mathematics, *Math. Intelligencer* **1**:3 (1978), 151–161.
27. D. Laugwitz, *Zahlen und Kontinuum: Eine Einführung in die Infinitesimalmathematik*, Lehrbücher und Monographien zur Didaktik der Mathematik 5, Bibliographisches Institut, Mannheim, 1986.
28. D. Laugwitz, Hidden lemmas in the early history of infinite series, *Æquationes Math.* **34** (1987), 264–276.
29. D. Laugwitz, Definite values of infinite sums: aspects of the foundations of infinitesimal analysis around 1820, *Arch. Hist. Exact Sci.* **39** (1988/9), 195–245.
30. A. H. Lightstone, Infinitesimals, *Amer. Math. Monthly* **79** (1972), 242–251.
31. A. H. Lightstone, Infinitesimals and integration, this MAGAZINE **46** (1973), 20–30.
32. W. A. J. Luxemburg, What is nonstandard analysis?, *Amer. Math. Monthly* **80** (1973), 38–67.
33. M. Machover, The place of nonstandard analysis in mathematics and in mathematics teaching, *Brit. J. Phil. Sci.* **44** (1993), 205–212.
34. E. Maor, *e: The Story of a Number*, Princeton University Press, Princeton, 1994.
35. M. McKinzie, *The Halley-Euler Method*, Doctoral dissertation, University of Wisconsin, Madison, 2000.
36. M. McKinzie and C. Tuckey, Nonstandard methods in the precalculus curriculum, in [**5**].

37. M. McKinzie and C. Tuckey, Hidden lemmas in Euler's summation of the reciprocals of the squares, *Arch. Hist. Exact Sci.* **51** (1997), 29–57.
38. E. Nelson, *Radically Elementary Probability Theory*, Annals of Mathematics Studies **117**, Princeton University Press, Princeton, 1987.
39. A. Pringsheim, Über ein Eulersches Konvergenzkriterium, *Bibliotheca Mathematica* ser. 3 vol. **6** (1905), 252–256.
40. A. Robinson, *Non-standard Analysis*, North-Holland Publishing Co., Amsterdam, 1966. Reprinted by Princeton University Press, Princeton, 1996.
41. A. P. Simpson, The infidel is innocent, *Math. Intelligencer* **12** (1990), 42–51.
42. K. D. Stroyan, *Calculus using Mathematica*, Academic Press, San Diego, CA, 1993.
43. K. D. Stroyan and W. A. J. Luxemburg, *Introduction to the Theory of Infinitesimals*, Academic Press, New York, 1976.
44. K. Sullivan, The teaching of elementary calculus using the nonstandard analysis approach, *Amer. Math. Monthly* **83** (1976), 370–375.
45. C. Tuckey, *Nonstandard Methods in the Calculus of Variations*, Pitman Research Notes in Mathematics 297, Longman House, Harlow, Essex, 1993.


## Letter to the Editor

Dear Editor:

In my paper, "Avoiding your spouse at a bridge party," appearing in the February 2001 issue of this MAGAZINE, I calculated certain probabilities, associated with couples playing bridge, $b_n$, using the inclusion-exclusion principle. In an aside, I commented that the fact that the probabilities could be expressed as a sum with decreasing terms "is a consequence of our having formulated the expression for the $b_n$ using the inclusion-exclusion principle." Professor Lajos Takacs pointed out to me in a letter that this claim is false. It is true that the terms in the sum are decreasing, but this fact is not a consequence of the inclusion-exclusion principle.

Recall the inclusion-exclusion principle, which can be proved by doing the following exercise from *Probability Theory and Examples, 2nd Ed.*, by Richard Durrett (p. 22, ex. 3.11):

Let $A_1, A_2, \ldots, A_n$ be events and $A = \bigcup_{i=1}^n A_i$. Prove that $1_A = 1 - \prod_{i=1}^n (1 - 1_{A_i})$. Expand out the right hand side, then take expected value to conclude

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i<j} P\left(A_i \bigcap A_j\right)$$
$$+ \sum_{i<j<k} P\left(A_i \bigcap A_j \bigcap A_k\right) - \cdots + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right).$$

In the preceding, $1_A$ is the indicator function equal to 1 if $x \in A$, and 0 otherwise.

A trivial example of a case for which my statement is false is where we have $n$ events $A_1, \ldots, A_n$ such that $A_1 = A_2 = \cdots = A_n$, and $P(A_i) = \alpha$; then

$$\sum_{i_1 < i_2 < \cdots < i_k} P\left(A_{i_1} \bigcap \cdots \bigcap A_{i_k}\right) = \binom{n}{k}\alpha.$$

In this case, then, the terms are just $\alpha$ times the binomial numbers $n, \binom{n}{2}, \ldots, \binom{n}{k}, \ldots, n, 1$ and this sequence is not decreasing. When $n = 3$, for example, the sequence is 3, 3, 1.

I regret the error.


Barbara H. Margolius
Cleveland State University
Cleveland, Ohio 44115
b.margolius@csuohio.edu

# Smullyan's Vizier Problem

MICHAEL KHOURY, JR.
Denison University
Granville, OH 43023

In his book *Satan, Cantor, and Infinity*, renowned logician and puzzler Raymond Smullyan poses a problem in the familiar genre of knights (who always tell the truth) and knaves (who never tell the truth). Although a solution of sorts is presented, evading the difficulty without confronting it, the central difficulty is by no means resolved. In this paper, I will analyze the problem and present two solutions, one of which I prove to be the most efficient algorithm to solve Smullyan's problem.

Though it will later be necessary to alter the problem, for the moment let me present it exactly as Smullyan originally gave it with the following excerpt from *Satan, Cantor, and Infinity*, in which the character Alexander meets with King Zorn after deducing the location of his kidnapped love, the princess Annabelle.

> *"Ah!" said the King, "your next task is to find out whether my Grand Vizier is a knight or a knave. If you succeed, I will have Annabelle released. You may ask the Vizier as many questions as you like, but they must all be answerable by yes or no."*
>
> *"But that's ridiculously easy!" cried Alexander. "I have merely to ask one question—one question whose answer I already know, such as whether two plus two equals four. From his answer, I will of course know whether he is a knight or a knave."*
>
> *"You shouldn't have interrupted!" said the King. "Of course you can find out by asking just one question whose answer you already know. But I was about to say that you are not allowed to ask any question whose answer you already know."*
>
> *The suitor stood lost in thought.*
>
> *"Let me be more explicit," said the King. "You don't have to plan the sequence of questions in advance; at any stage, the questions you decide to ask may depend on the answer already given, but at no stage are you allowed to ask a question whose truthful answer could be known to you."*

The analysis of this simply-stated problem gives rise to a host of logical issues.

## Alexander's instructive blunder and Smullyan's trick

Alexander offers what he believes to be a solution, though Smullyan's King Zorn debunks it handily. Still, it will be to our benefit to examine this pseudosolution.

Alexander's solution begins by asking if the Vizier is an unmarried knight. This question is deceptively shrewd. Notice that unmarried knights, being knights who must needs be truthful, would respond in the affirmative. Any knave, who must always lie, would falsely claim to be a unmarried knight, so all knaves would say "yes." This question has the remarkable property that it will be answered "no" *if and only if* the Vizier is a married knight. We can say that this question *singles out* married knights. In particular, if the Vizier happens to be a married knight, this question will reveal that fact. Also, the correct answer is not known to Alexander in advance. Hence Alexander might pass the test with this one question.

To this point Alexander's analysis is entirely correct. However, in what follows he makes a critical mistake. He reasons that, married knights being singled out, it is only necessary to single out unmarried knights. After both questions, he would have singled out all knights, so the Vizier's knighthood or knavery would be evident. It can be checked that "Are you a married knight?" singles out unmarried knights, and that this question individually is legal. The flaw in Alexander's reasoning is that, although the questions do together settle the problem of identifying the Vizier, and although each question alone would be legal, *each question precludes Alexander's asking the other.* After the first question, Alexander knows whether the Vizier is a married knight—but this is precisely the answer to the next question he wants to ask. Similarly, it would be illegal to ask the questions in the opposite order. Thus Alexander's two-question solution is in fact incorrect.

**Smullyan's trick** Smullyan presents a solution to the problem in his book, but as I mentioned, this solution only evades the difficulty. The solution that Alexander uses is to draw a card from a deck, show it to the Vizier without looking at it, and ask him "Is this card red?" Since he has not seen the card, he does not know the answer to the question *in advance*. Having heard the response, the interrogator can simply look at the card and know immediately whether the Vizier was telling the truth. Voila. Case settled... sort of....

Although King Zorn admits that this solution meets his stipulations, Smullyan's King Zorn states that "it seems like cheating." Personally, this solution leaves an unsatisfactory taste in my mouth. Read on....

**Evading the issue, which is what, exactly?** To avoid semantic difficulty, we must define the concept of *answer* and *response*. In this paper, *answer* shall always refer to the truthful, correct answer to a question, while *response* shall always refer to what the Vizier actually claims. For example, if a knave is asked "Does a triangle have three sides?" the answer will be "yes" while the response will be "no." (It should not be supposed that the answer is independent of the identity of the Vizier and that the response is not; indeed, for questions like "Are you a knight?" the situation is exactly the reverse.) Note that there is an interesting mutual dependency of the answer to a question $\mathcal{Q}$, the response to $\mathcal{Q}$, and the knighthood/knavery of the Vizier. Any two determine the third (the reader should verify this).

This interdependency is the basis of diagnosing knights and knaves. The usual method of identifying knights and knaves is asking them questions whose answers are already known, like "Is two prime?" and comparing the response to the answer. By the interdependency principle, this method will always work. An interesting question, clearly the question Smullyan's King Zorn (and of course Smullyan) intended to address, is whether knights and knaves can be distinguished without directly using this dependency.

We can now say exactly why Smullyan's solution is unsatisfying. The downfall of Smullyan's solution is that it does not address the fundamental question behind the Vizier problem. The identification is still done by comparing the answer and the response to a given question. The only difference is that of which is learned first, a detail which had never been of importance. Therefore, Smullyan's solution to his own problem does not settle the question of whether a knight can be distinguished from a knave without using the combined knowledge of a response and an answer to the same question.

## A suitable restatement

We can now infer with confidence that King Zorn's challenge was motivated by the question of whether a knight can be distinguished from a knave without comparing a response to an answer. Restating the problem in such a way as to ensure that a solution must address this fundamental issue turns out to be tricky, even though we have now made precise what the fundamental question is.

Since the intent of the problem is to prevent Alexander from comparing an answer to a response, a natural restriction could be that Alexander may not know or be able to learn the answer to any of his questions, either at the time they are asked or at any later time. This would certainly rule out Smullyan's solution. Unfortunately, this restriction is far too much for King Zorn to demand; no solution would exist. Suppose that Alexander could determine the identity of the Vizier under these rules. Then, by comparing the Vizier's nature to all of his responses, Alexander could learn the true answers to all the questions, rendering his procedure illegal.

To address the issue, we must make a weaker restriction. A key to Alexander's success is his ability to make an observation of the card during the interrogation. To prevent this, King Zorn can require that Alexander ask his questions from an isolation chamber, a selective sensory deprivation chamber such that the only information he can obtain from the outside world is the response by the Vizier.

As it stands, this "isolation chamber" admits a subtle modification of Smullyan's scheme. Smullyan's scheme is based on the principle of asking a question whose answer becomes known after the question is asked. In that spirit, Alexander could ask "Will you respond to this question in one minute or less?" Once again, the question is legal (since the answer is not determined in advance) and effective (since the answer will be evident as soon as the response is given). In some ways this method is an improvement over the "Is it red?" strategy, since it dispenses with the external phenomenon of the card. To escape this, we will assume that the isolation chamber masks all ancillary information about the response as well, including the Vizier's response time, tone of voice, facial expression, etc. The selective sensory deprivation permits Alexander to perceive *only* whether the Vizier responds "yes" or "no." The practical details of how this masking of information would be carried out is irrelevant to our discussion.

Though this restriction has made impossible both Smullyan's evasion and the above evasion, it is not entirely satisfactory, since Alexander can ask a question like "Is the hundred thousands digit of the result of multiplying 4305836 by 23894 odd?" Alexander can easily delay his own computation of the answer until after he has heard the Vizier's response. Although the isolation of Alexander is important, we are still missing one crucial restriction on the question—the answer may not be *determined* by the information available to Alexander when he asks it. This is harsher than merely prohibiting questions whose answer is *known*. In particular, it closes the loophole of the multiplication problem.

It seems that this reformulation of the problem would express King Zorn's true wishes. It has neither a trivial solution (like the original statement) nor a trivial proof of impossibility (like the restatement given in the second paragraph of this section). Also, any solution of this problem would necessarily avoid comparing an answer to a response.

Therefore, for the purposes of the remainder of this paper, the restrictions on Alexander's actions are understood to be as follows:

- Alexander receives no information from the external world other than the Vizier's responses.

- Alexander may ask no question whose answer can be determined from the information that Alexander has at the time he asks it.
- Alexander may ask only questions that the Vizier can answer by "yes" or "no."

The last of these items is not affected by the analysis of this section and is left as Smullyan gave it.

## Eliminating the bounded

Some of the material in this section is based on a putative proof Alexander offers to show that the King's task is impossible. While Alexander's proof is not correct, as Smullyan's Professor Bacterius explains, parts of it can be modified and repaired to prove that certain *types* of solutions to the problem are impossible.

It is straightforward to dismiss the possibility of a single question that would solve the problem. To see this, suppose that there existed such a question $Q$. Since $Q$ solves the problem, either of the two possible responses tells the interrogator the nature of the Vizier. Since the nature of the Vizier is unknown prior to the question, both outcomes are possible. That is, there are two possible outcomes, each of which implies one of two possible natures of the Vizier. Since knighthood and knavery must each be represented, one of the following two cases holds regarding $Q$:

1. *either* the response "yes" is given and the Vizier is a knight *or* the response "no" is given and the Vizier is a knave, or
2. *either* the response "no" is given and the Vizier is a knight *or* the response "yes" is given and the Vizier is a knave,

where the interrogator is aware whether 1 or 2 is in effect. However, in case 1, whichever response is given, the answer is "yes," so the answer is predetermined. Similarly, in case 2, the answer is predetermined to be "no." Thus such a question $Q$ would not be admissible, and a one-question solution is impossible.

A simple extension of this argument is possible to show that no algorithm can guarantee success when a bound on the number of questions known in advance. That is, the interrogator cannot know before beginning the interrogation that he will succeed in 10 questions, or 200, or 3000, or any other specified number. Again the proof is by contradiction. Suppose a solution exists in which the interrogator knows in advance some bound on the number of questions. Let $n$ be the smallest such bound known in advance by the interrogator. Since he does not know $n-1$ to be a bound, his method must anticipate some case in which the Vizier's identity is still unknown after $n-1$ questions. Since he knows his method will succeed in $n$ questions, he knows the $n^{th}$ question will be decisive in this case. The analysis of the preceding paragraph applies here and shows that no such $n^{th}$ question would be admissible. Thus, no solution is possible when a bound on the number of questions is known in advance.

It is, however, possible to construct a solution that will succeed with an unspecified number of questions. I have in fact two methods.

## First solution

Let us revisit an idea presented by Smullyan and detailed previously, that of asking, "Are you an unmarried knight?" Recall that only a married knight could respond "no" to this question, so this question singles out married knights. A "no" response, therefore, would settle the matter, while a "yes" response would not distinguish between unmarried knights and knaves of any marital status. While one would now like to single out unmarried knights so that all types of knights will be accounted for, leaving only knaves and hence identifying the Vizier, this is prohibited because the required question, "Are you a married knight?" is illegal since the result of the first question gives precisely this information. Thus the answer would be known in advance, as Smullyan points out. This should not surprise us now, however, since the algorithm, if successful, would require at most 2 questions, and hence would be a successful algorithm with a bound on the number of questions known in advance; we have already shown this to be impossible.

An alternate approach would be to single out married knaves, so the Vizier would certainly be identified if he is married. The question to do this is "Are you an unmarried knave?" A "yes" response will be forthcoming if and only if the Vizier is a married knave, as the reader can verify. If this question identifies the Vizier, then we have succeeded. Otherwise, we seem to be in a bit of trouble. Singling out unmarried knaves is just as impossible as identifying unmarried knights, for precisely the same reason. While this approach offers no guarantee, it will succeed if the Vizier is married, *regardless of his nature*. The crucial point is that these two questions have a *positive probability of success*.

Note the following general result, which we can apply later. We can single out knights with a certain characteristic $\mathcal{X}$ by asking, "Are you a knight without characteristic $\mathcal{X}$?" and listening for a "no." Similarly, we can single out knaves with a characteristic $\mathcal{X}$ by asking, "Are you a knave without characteristic $\mathcal{X}$?" and listening for a "yes." (The reader should verify this observation and compare with the questions already used, in which "being married" is $\mathcal{X}$.)

I now present the first method of diagnosing the Vizier. First, the two questions about marriage are asked, settling the issue if the Vizier happens to be married. Next, if that fails to determine his nature (which will be the case if and only if he is unmarried), the corresponding questions are asked to test for knights and knaves with some unrelated characteristic, taking $\mathcal{X}$ to be, perhaps "liking iced tea." These two questions will distinguish the Vizier if he likes iced tea. Furthermore, they will be legal, since the prior questions about marriage will give no information about iced tea preferences. Repeat this with a different characteristic, then another, and so on until you happen to choose a characteristic that the Vizier actually possesses. Once you choose a characteristic that the Vizier has, one of the two questions will force him to give himself away. A typical beginning to the interrogation could be as follows:

| | | |
|---|---|---|
| $\mathcal{X}$ = being married | | |
| | "Are you an unmarried knight?" | "Yes." |
| | "Are you an unmarried knave?" | "No." |
| $\mathcal{X}$ = liking iced tea | | |
| | "Are you a knight who does not like iced tea?" | "Yes." |
| | "Are you a knave who does not like iced tea?" | "No." |
| $\mathcal{X}$ = owning a pet cat | | |
| | "Are you a knight who doesn't own a pet cat?" | "Yes." |
| | "Are you a knave who doesn't own a pet cat?" | "No." |

and so on.

The Vizier's responses in the above represent a sort of worst-case scenario. If the Vizier gives the opposite response from the one given here to any particular question, the interrogation will be complete and the Vizier identified. Provided that the $\mathcal{X}$'s chosen are independent, it is easy to see that all the questions are legal. Also, provided that the $\mathcal{X}$'s chosen have probability bounded away from zero of belonging to the Vizier, the probability that the Vizier will lack the first $n$ chosen approaches zero as $n$

becomes large. In other words, the probability is zero that the Vizier will lack every characteristic you will think of and thereby escape classification. Also equivalently, this strategy will succeed with probability 1. Therefore, this method could be considered a solution. It should be pointed out immediately that an event with probability 1 is not necessarily certain to occur, despite the common term "probabilistic certainty." For example, if two people independently select positive integers, they will choose different integers with probability 1. Yet they *might*, say, both choose 13. As a further defect, this method relies on the interrogator's having to invent an unbounded number of independent characteristics. It is not entirely obvious that an unbounded number of independent questions can be generated; if not, this solution would not be a complete algorithm.

Instead, the spirit of the approach can be retained and the procedure made complete by having the Vizier flip a coin before each pair of questions and taking the characteristic $\mathcal{X}$ always to be "having just flipped a head." This is certainly legal by the above line of reasoning, since the individual coin flips are independent (of course Alexander does not see the coin flip, because he is in the sensory deprivation chamber). Since the Vizier will be identified within two questions after the first head is flipped, the method could fail only if the coin came up tails forever, an event with probability zero. The probability of success, then, is 1. This algorithm does give a probabilistic certainty of identifying the Vizier without use of a response-answer combination.

However, the expected number of questions required to diagnose the Vizier is *finite*. Since the expected number of coin flips required to get a head is well known to be 2, the expected number of question pairs here is also 2.

## Second solution

A different solution relies on a single characteristic with a countably infinite number of possibilities; compare this to the preceding solution, which relied on a countably infinite number of characteristics, each with two possibilities. Before the algorithm is started, the Vizier should be made to choose an element of a countable set; for convenience, we will have the Vizier choose a natural number. We can now single out, using our established technique, a knight with number 1, then a knave with number 1. Next, we single out a knight with number 2, then a knave with number 2. This continues until the Vizier's chosen number is reached, at which point one of the two questions will determine the Vizier's nature.

More specifically, for any integer $n$, let $A_n$ be the question "Are you a knight whose number is not $n$?" and let $B_n$ be the question "Are you a knave whose number is not $n$?" The response to $A_n$ is "no" if and only if the Vizier is a knight who chose $n$, and the response to $B_n$ is yes if and only if the Vizier is a knave who chose $n$. It is clear, then, that if the Vizier is a knight (respectively a knave) and he chose $n$, he will be identified by question $A_n$ (respectively $B_n$). Also notice that if a question does not identify the Vizier, it will not distinguish between any of the other possible cases; thus these questions are legal in any order, provided that the Vizier has not been identified. Therefore, a legal and successful series of questions is $A_1, B_1, A_2, B_2, A_3, B_3, \ldots$.

It should be noted that extreme care should be taken in the selection of the Vizier's number. If Alexander can know any bounds on the number at all, even ridiculously high bounds, the method will not work. The number should be chosen randomly, *not* based on some fixed characteristic of the Vizier. Even a seemingly innocent way of choosing the number can lead to a bound. The number of children the Vizier has, for example, is bounded above by even the wildest overestimate for the population of the Earth. In fact, the Vizier must think of his number only and not write it down, since the resulting upper bound on the amount of time it takes to write his number would imply an upper bound on the number. Although this is nit picking by any standard, it is essential. If there is any bound placed on the Vizier's number, his number of choices will be finite (though gargantuan) and therefore the interrogation will need only a bounded number of questions. We have already proven, however, that an interrogation with a known bound on the number of questions is doomed to failure.

To deal with these problems, we must *postulate* that it is possible for the Vizier to choose a number in such a way that his choices are unbounded. No procedure for doing so can be given explicitly, nor can a probability distribution be specified. In fact, a mathematical justification of such a process is impossible, though it is intuitively very reasonable. Any reader who is unwilling to humor this postulate is free to accept only the first solution as legitimate.

Note the contrast to the other solution. Where the success of the other was a mere probabilistic certainty, this is an absolute certainty. Though the number of questions is unbounded, it is finite for each possible outcome of the interrogation. That is, if the Vizier knew the algorithm the interrogator was using, *the Vizier himself* would be able to specify the finite number of steps it will take (at most twice the number he chose). Observe that this is not true of the first solution. Clearly this cannot be improved upon, since solutions with bounds known by the interrogator himself have already been proven impossible. Still, there are many possible meanings for *improvement*. The second strategy is superior in the sense that it is guaranteed to terminate, but it relies on an intuitive assumption. The first strategy is superior in the sense that it has a finite expected number of questions, but it will succeed only almost surely.

## Simple generalizations

There are several ways in which this problem can be altered to admit new analyses. I consider four here, two in which the problem is made more difficult by relaxing the assumptions about the Vizier, and two in which the problem is made easier (though not simpler) by relaxing the requirement on the question.

First, we consider the problem of identifying the Vizier when there is at least one other option for his nature. Along with knights and knaves, the standard other characters in the genre are alternators, who tell the truth and lie alternately on successive questions, and normals, who freely do as they please.

If the problem is to identify whether the Vizier is a knight, a knave, or an alternator, a very simple method is available. First, any question with an undetermined answer is asked twice; "Do you like your coffee black?" for example. Alternators will be picked out immediately by their inconsistency. If the Vizier were discovered not to be an alternator, we would be in the situation of Smullyan's original problem, and my solution could be applied. Since this algorithm is identical to the previous one, except for the addition of two questions, it will also have a finite expected number of questions (if we use the first method) or guaranteed termination (if we use the second method). No algorithm with a bound on the number of questions can succeed, since any algorithm that solves this problem would certainly have solved the original.

The interrogator's situation is very much the worse if the Vizier might be a normal, that is, if he is free to respond to questions in any way he pleases. Distinguishing a normal from a knight or a knave is impossible because, with no restrictions on his responses, he could impersonate a knight, answering each question exactly as a knight would. (Note that this is *not* the same as a normal deciding to always tell the truth! Consider the Vizier's behavior under the question "Are you normal?" A normal impersonating a knight would say "no," while a normal telling the truth would say "yes."

Confusing folks, these normals.) Therefore, the interrogator could never know conclusively that the Vizier was not normal. An interesting question to consider is what types of limitations could be placed on normals to make them detectable by an algorithm of the type we have been discussing. Phrased alternately, how much freedom can be given to "subnormals" in the responses they are allowed to give, without allowing them to successfully impersonate knights or knaves. This, however, will not be addressed here.

Alternatively, we could change the problem by relaxing the constraints on the questions used. In addition to the requirement that the answer not be predetermined, whose importance to the heart of the problem was discussed at length, we have required that a question be answerable and that it be yes-or-no. We now relax each condition in turn, beginning with the latter.

If the interrogator is permitted to ask questions with more than two possible answers, much of our argument breaks down. First, although knowledge of an answer and a response to the same question determine the nature of the Vizier, it is no longer the case that knowledge of the nature of the Vizier and either the answer or the response to a question determines the other. Worse, the proofs that eliminate bounded solutions are no longer valid. In particular, the breakdown into two cases fails. Consider as an example a multiple-choice question with answers/responses A, B, and C. If one single question is to determine the nature of the Vizier, there are now six cases to consider instead of two:

1. *Either* the response A or B is given and the Vizier is a knight *or* the response C is given and the Vizier is a knave.
2. *Either* the response A or C is given and the Vizier is a knight *or* the response B is given and the Vizier is a knave.
3. *Either* the response B or C is given and the Vizier is a knight *or* the response A is given and the Vizier is a knave.
4. *Either* the response A is given and the Vizier is a knight *or* the response B or C is given and the Vizier is a knave.
5. *Either* the response B is given and the Vizier is a knight *or* the response A or C is given and the Vizier is a knave.
6. *Either* the response C is given and the Vizier is a knight *or* the response A or B is given and the Vizier is a knave.

None of these predetermines the answer. Therefore, the proof that success is impossible in a single question does not carry over.

A one-question solution can indeed be found if the available responses are understood to be, for example, all synonyms of "yes" or "no" (responses such as "affirmative," "absolutely," and "negative"). The question, "Will you answer 'yes' to this question?" suffices to identify the Vizier. A knight could answer with "yes" or any word meaning "no," while a knave must answer with a synonym of "yes" other than "yes" itself. Also, this question is legal, since the truthful answer (whether "yes" will be said) is not determinable in advance.

If the interrogator is permitted to ask questions that need not necessarily be answerable, then there are three possible responses: "yes," "no," and silence. (We assume that the Vizier will always answer if he can, or else the problem is trivially unsolvable because the Vizier can refuse to answer any questions.) By the argument of the preceding paragraph, we no longer have a proof that a solution is impossible with a bound on the number of questions. In fact, a one-question solution is available: "Are you going to respond 'yes' to this question?" A knight can respond to this any way he pleases and still be honest, while a knave cannot respond either way without unwittingly being honest, so he must hold his tongue. Thus the question succeeds. Furthermore, it is legal, because the same analysis shows that the correct answer is not determined, since a "yes" response is possible but not guaranteed. Therefore with this modification, the problem has a one-question solution.

The above generalizations, while not as provocative or significant as the original problem, do contribute at least one important result, in particular the treatment of not-necessarily-answerable questions. Looking retrospectively over this analysis, the variety of issues confronted by the problem of King Zorn's Vizier is quite surprising.

**Acknowledgment.** The author wishes to thank the referees for the reality check and helpful suggestions.

## REFERENCE

1. Raymond M. Smullyan, *Satan, Cantor, and Infinity*. Knopf, New York, 1992.

## Letter to the Editor

Dear Editor:

I enjoyed James Tanton's Proof Without Words of a little known property of equilateral triangles in a recent issue of the MAGAZINE [**2**]. Readers may be interested to know that the result illustrated is known as *Viviani's Theorem*, after Vicenzo Viviani (1622–1703) [**1**]. Another proof (also without words) of this theorem appeared previously in the MAGAZINE [**3**].

Roger Nelsen
Lewis and Clark College
Portland, Oregon 97219

## REFERENCES

1. Ross Honsberger, *Mathematical Gems*, MAA, 1973, p. 23.
2. James Tanton, Proof without words: Equilateral triangle, this MAGAZINE, **74**:4 (2001), 313.
3. Samuel Wolf, Proof without words, this MAGAZINE, **62**:3 (1989), 190.

# NOTES

## The Wallet Paradox Revisited

MAUREEN T. CARROLL
University of Scranton
Scranton, PA 18510
carrollm1@uofs.edu

MICHAEL A. JONES
Montclair State University
Upper Montclair, NJ 07043
jonesma@pegasus.montclair.edu

ELYN K. RYKKEN
Muhlenberg College
Allentown, PA 18104
elrykken@muhlenberg.edu

Put your wallet on the table next to mine. The game is this: The person whose wallet has less money wins all the money in the other person's wallet. Do you want to play?

You might think along these lines: "I don't know how much money that other wallet has, and I'm not even sure how much is in mine. If I have more money, then I'll lose it, but if I have less, I'll win the larger amount. I have no idea what the odds are, but since I stand to win more money than I can lose, it seems like a good game." Upon further thought, you realize that both players are probably thinking the same thing! Can both be correct?

How can a game favor both players? It can't! In any two-person, zero-sum game (where one person wins what the other person loses), it is not possible for the game to be advantageous to both players. Believing that the wallet game favors both players is a paradox, one discussed by Martin Gardner [**3**]. A variation of this game was originally posed by Kraitchik [**1**] where the person with the greater amount in her wallet gives the difference to the other.

What if you and I decide to play this game day after day? We will need to establish a few more rules, because it would not be a very interesting game if neither of us ever carried any money. Since we do not want to mandate a minimum amount that must be carried, we agree that on average (and in the long run) we will carry the same amount of money. How should you decide how much to carry each day?

Kraitchik shows that if the amount of money each player carries is uniformly (discretely) distributed between 0 and some large $x$ (he uses the total amount of money that has been minted to date), then the game is *fair* (each player's expected payoff is zero). Gardner notes that this does not explain the source of the paradox. Merryfield, Viet, and Watson [**2**] argue that the source of the apparent paradox is that the players do not take into account the probabilities of winning and losing. They argue that if the amounts of money in the players' wallets are determined by independent, identically distributed random variables, then the game is also fair. Hence, the game is fair when the players are required to use the same distributions.

It is natural to ask if requiring the players to carry the same amount on average might also ensure a fair game. When "on average" is interpreted as requiring both players' distributions to have the same mean, Merryfield, *et al.* point out that the game may not be fair. In fact, they give an example that shows it is possible to have a smaller mean than one's opponent and still be at a disadvantage (that is, have a negative expected payoff). At the end of their paper, they pose the following question:

> If we suppose that the distributions of players A and B are required to have the same means, is there a strategy that player A could adopt to have a winning edge? In other words, is there a preferred distribution (or a winning strategy)? [**2**]

The answer to this question depends upon whether knowledge of an opponent's strategy, *not just the mean*, is assumed. In this article, we show that if a player knows her opponent's strategy, then she can construct a winning strategy which has the same positive mean or median as her opponent. This implies that there is no optimal strategy (or Nash equilibrium) when players are restricted to use strategies with the same mean or median. We consider both the discrete and continuous cases. Throughout, we assume that players' distributions are independent.

**Discrete distributions** Consider an example using discrete distributions. Suppose players $A$, $B$, and $C$ use strategies given by independent random variables $X$, $Y$, and $Z$, respectively. Suppose $X$ places probability 1 on \$2, $Y$ places probability $1/2$ on both \$1 and \$3, while $Z$ places probability $3/4$ and $1/4$ on \$1 and \$5, respectively. Notice that the mean of each distribution is \$2. Using the notation developed by Merryfield, *et al.* [**2**], let $W_{A/B}$ be the random variable returning the amount of money that player $A$ wins (or loses) when playing against player $B$, that is,

$$W_{A/B} = \begin{cases} -X & \text{if } X > Y \\ Y & \text{if } X < Y \\ 0 & \text{if } X = Y. \end{cases}$$

If players $A$ and $B$ use strategies $X$ and $Y$, respectively, $X$ is preferred to $Y$, denoted $X \succ Y$, if and only if $E(W_{A/B}) > 0$.

Suppose players $A$ and $B$ play the Wallet Game. Player $A$ loses \$2 when player $B$ carries \$1 and wins \$3 when $B$ carries \$3. Player $A$'s expected payoff against player $B$ is $E(W_{A/B}) = \frac{1}{2}(-2) + \frac{1}{2}(3) = \frac{1}{2}$. Thus, strategy $X$ is preferred to strategy $Y$.

The following matrix, which is similar to one used by Kraitchik [**1**], is used to compute $E(W_{B/C})$. The $(i, j)^{\text{th}}$ entry of the matrix, $m_{ij}$, is the amount that player $B$ wins or loses when carrying $y_i$ in his wallet, while player $C$ is carrying $z_j$ in her wallet; this occurs with probability $p_i q_j$.

| B/C | $q_0 = 3/4$ $z_0 = 1$ | $q_1 = 1/4$ $z_1 = 5$ |
|---|---|---|
| $p_0 = 1/2$ $y_0 = 1$ | 0 | 5 |
| $p_1 = 1/2$ $y_1 = 3$ | $-3$ | 5 |

Calculating $E(W_{B/C})$ requires summing the products of the matrix entries and their probabilities; in this case,

$$E(W_{B/C}) = \sum_{i=0}^{1}\sum_{j=0}^{1} p_i q_j m_{ij} = \frac{3}{8}(0) + \frac{3}{8}(-3) + \frac{1}{8}(5) + \frac{1}{8}(5) = \frac{1}{8},$$

and $Y \succ Z$. Finally, if players $A$ and $C$ play, then $E(W_{A/C}) = \frac{3}{4}(-2) + \frac{1}{4}(5) = -\frac{1}{4}$ and $Z \succ X$.

The lack of transitivity of $X \succ Y$, $Y \succ Z$, and $Z \succ X$ suggests that there may not be a "best" strategy in the discrete case when both players are required to have the same positive mean. The following proposition confirms this, answering the question posed by Merryfield, *et al.* [**2**], by showing that there is no distribution that is preferred to all others, that is, there is no optimal distribution. (Recall that $Y$ has *finite support* if positive probability is placed on a finite number of values.)

PROPOSITION 1. *For any discrete random variable $Y$ with finite support, there exists a discrete random variable $X$ with $\mu_X = \mu_Y$ such that $X \succ Y$.*

*Proof.* Suppose player $A$ knows that player $B$ carries an amount of money given by the random variable $Y$ whose probabilities, $q_i$, are distributed on a finite set of monetary values $y_i$, such that $y_0 = 0$ and $y_i < y_{i+1}$ for all $i \leq n$. Since the mean of $Y$, $\mu_Y$, is positive, it follows that $q_0 \neq 1$.

We construct for player $A$ a random variable $X$ that defeats $Y$. Player $A$'s strategy is to win almost every game; however, when player $A$ loses, she forfeits a large amount of money. Interestingly, player $A$ need only place positive probability on three values, regardless of the complexity of player $B$'s distribution. Define $X$ by the distribution of probabilities $p_i$ on monetary values $x_i$ as follows: $p_0 = q_0$ on $x_0 = 0$, $p_1$ on $x_1 = \frac{1}{2}y_1$, and $p_2 = 1 - p_0 - p_1$ on $x_2$, where $p_1$ and $x_2$ satisfy the following conditions,

$$\frac{(1-p_0)\mu_Y}{\mu_Y + \frac{1}{2}(1-p_0)y_1} < p_1 < 1 - p_0 \quad \text{and} \quad x_2 = \frac{\mu_Y - \frac{1}{2}p_1 y_1}{1 - p_0 - p_1}.$$

Notice that $p_1$ exists since $p_0 = q_0 \neq 1$. Also, $x_2$ is defined such that $\mu_X = \mu_Y$.

As in the example above, it is convenient to view the Wallet Game in matrix form. Although we do not know how $x_2$ compares to the $y_i$s, we assume the worst-case scenario for player $A$, that is, $x_2$ is greater than the largest monetary value that player $B$ carries, $y_n$. As before, the matrix entries are payoffs to player $A$.

| A/B | | $q_0$<br>$y_0$ | $q_1$<br>$y_1$ | $q_2$<br>$y_2$ | $q_3$<br>$y_3$ | $\cdots$<br>$\cdots$ | $q_n$<br>$y_n$ |
|---|---|---|---|---|---|---|---|
| $p_0$ | $x_0$ | $0$ | $y_1$ | $y_2$ | $y_3$ | $\cdots$ | $y_n$ |
| $p_1$ | $x_1 = y_1/2$ | $-x_1$ | $y_1$ | $y_2$ | $y_3$ | $\cdots$ | $y_n$ |
| $p_2$ | $x_2$ | $-x_2$ | $-x_2$ | $-x_2$ | $-x_2$ | $\cdots$ | $-x_2$ |

The expected values of the first column and first row cancel because

$$p_0(q_1y_1 + q_2y_2 + q_3y_3 + \cdots + q_ny_n) + q_0(-p_1x_1 - p_2x_2) = p_0\mu_Y - q_0\mu_X = 0.$$

Since $x_1 < y_i$ when $i > 0$, the remaining entries in the second row yield a positive contribution to the expected value for player $A$ in the amount of $p_1(q_1y_1 + q_2y_2 + \cdots + q_ny_n)$, or $p_1\mu_Y$. In this worst-case scenario, $x_2 > y_n$ implies that the remaining entries in the third row contribute the following to the expected value of player $A$

$$-p_2(q_1 + q_2 + \cdots + q_n)x_2 = -(1 - p_0 - p_1)(1 - q_0)x_2 = -\left(\mu_Y - \frac{1}{2}p_1y_1\right)(1 - q_0).$$

Hence, we have

$$E(W_{A/B}) \geq p_1\mu_Y - [\mu_Y - \frac{1}{2}p_1y_1](1 - q_0) > 0,$$

by the definition of $p_1$. Therefore, $X \succ Y$. ■

In our earlier example, and as indicated in the above proposition, playing the mean with probability one can be defeated. However, it *is* a winning strategy against all other symmetric, discrete distributions. In this case, player A loses half of the time with loss $\mu_X$, but wins half of the time with a gain that is greater than $\mu_X$. Hence, the expected payoff is positive. In the next section we focus on continuous density functions and examine the roles of both the mean and median.

**Continuous Density Functions** Suppose that random variables $X$ and $Y$ have continuous density functions $f$ and $g$, respectively. Recall that a continuous density function never places positive probability on a single value; that is, the probability of a player carrying a specific amount of money is zero. As in the discrete case, if $g$ is a symmetric density function then playing the mean with probability one (or equivalently, the median) is preferred to $g$. Although playing the mean with probability one does not satisfy our restriction to continuous density functions, this idea is easily modified to show the existence of a continuous density function with the same mean (and median) that defeats the original symmetric density function.

We do this in the following proposition, considering nonsymmetric, continuous density functions where players are required to have the same median. Denote the median of the random variable $X$ as $m_X$. Thus it is equally likely that the player has more than or less than $m_X$.

PROPOSITION 2. *For any random variable $Y$ with a continuous density function, there exists a random variable $X$ with a continuous density function where $m_X = m_Y$ and $X \succ Y$.*

*Proof.* Suppose player $A$ knows that player $B$ carries an amount of money given by the random variable $Y$ with probability density function $g$. The discrete response $X = m_Y$, where $m_Y$ is the median of $Y$, is preferred to $Y$. This follows since the median $m_Y$ loses half of the time with a loss of $m_Y$ and wins half of the time, averaging a payoff greater than $m_Y$. Therefore, the expected payoff for player $A$ is positive. However, this is a discrete distribution. To construct a continuous distribution, playing the median can be considered as the limit of a sequence of uniform distributions where the variances tend to zero. Since the expected value of playing the median is positive, there exists a uniform distribution with $m_X = m_Y$ and positive expected value. ■

Since there is no optimal continuous density function when the distributions are required to have the same median, let's consider the case where they have the same mean. The following proposition shows that there is no optimal continuous density function in this case either. The proof is constructive, as in the discrete case, and the motivation for the strategy is similar. Once again, Player $A$'s strategy is to win more frequently than player $B$, while infrequently losing a large sum of money. We construct a density function that matches the opponent on $[0, m_Y]$, and is piecewise uniform on both $[m_Y, m_Y + \epsilon]$ and $[n - \epsilon, n]$, where $n$ and $\epsilon$ are selected such that $\mu_X = \mu_Y$ and $X \succ Y$.

PROPOSITION 3. *For any random variable $Y$ with a continuous density function, there exists a random variable $X$ with a continuous density function where $\mu_X = \mu_Y$ and $X \succ Y$.*

*Proof.* Suppose player $A$ knows that player $B$ carries an amount of money given by the random variable $Y$ with continuous density function $g$. Suppose $g$ has mean $\mu_Y$ and median $m_Y$. As in the discrete proof, the goal is to construct a density function that defeats $g$ while having the same mean. Let $\gamma$ be the average conditional expected value of $g$ conditioned on being in the interval $[m_Y, \infty)$, that is, $\gamma = \int_{m_Y}^{\infty} yg(y)\,dy / \int_{m_Y}^{\infty} g(y)\,dy = 2\int_{m_Y}^{\infty} yg(y)\,dy$.

Let $X$ be a random variable with density function $f$ defined by

$$f(x) = \begin{cases} g(x) & \text{on } [0, m_Y] \\ \frac{1-\epsilon}{2\epsilon} & \text{on } (m_Y, m_Y + \epsilon] \\ \frac{1}{2} & \text{on } [n-\epsilon, n] \\ 0 & \text{elsewhere,} \end{cases}$$

where $0 < \epsilon < 1$ is selected so that $n - \epsilon > m_Y + \epsilon$, where

$$n = \frac{\gamma}{\epsilon} + \epsilon - \frac{1}{2} - \frac{m_Y}{\epsilon} + m_Y,$$

and so that the following inequality holds:

$$2\gamma(1-\epsilon)\int_{m_Y+\epsilon}^{\infty} g(y)\,dy > \left(\gamma + \epsilon^2 - \frac{\epsilon}{2} - m_Y + m_Y\epsilon\right) + 2(m_Y+\epsilon)(1-\epsilon)\int_{m_Y}^{m_Y+\epsilon} g(y)\,dy. \qquad (1)$$

Notice that the left side of (1) converges to $\gamma$ as $\epsilon$ approaches zero, while the right side converges to $\gamma - m_Y$. Also, $n$ grows without bound as $\epsilon$ approaches zero. Therefore, a sufficiently small $\epsilon$ can be chosen to satisfy both inequalities. Although (1) and the definition of $n$ appear complex, selecting such an $\epsilon$ guarantees that $\mu_X = \mu_Y$ and $X \succ Y$ as shown below.

Since $f$ is equal to $g$ on $[0, m_Y]$ and $f$ is composed of piecewise horizontal line segments on $(m_Y, \infty)$, then, by the definition of $n$,

$$\mu_X = \int_0^{m_Y} yg(y)\,dy + \left(\frac{1-\epsilon}{2\epsilon}\right)\epsilon\left(m_Y + \frac{\epsilon}{2}\right) + \frac{1}{2}\epsilon\left(n - \frac{\epsilon}{2}\right)$$
$$= \int_0^{m_Y} yg(y)\,dy + \frac{\gamma}{2} = \int_0^{\infty} yg(y)\,dy = \mu_Y.$$

To see that $X \succ Y$, we employ a matrix again. As in the previous proof, we consider the worst-case scenario for player $A$. For example, when $X$ is in the interval $[n-\epsilon, n]$ and $Y$ is in $(m_Y + \epsilon, \infty)$, we assume that $X$ loses $n$. Also, when $X$ is in $[m_Y, m_Y + \epsilon]$ and $Y$ is in $(m_Y + \epsilon, \infty)$ then $Y$ loses, on average, more than $\gamma$. In the following matrix, the entries are payoffs to player $A$. Let $\gamma_1$ and $\gamma_2$ be the average conditional expected values of $g$ conditioned on being in $(m_Y, m_Y + \epsilon]$ and $(m_Y + \epsilon, \infty)$, respectively.

| | A/B | $\frac{1}{2}$ $[0, m_Y]$ | $\int_{m_Y}^{m_Y+\epsilon} g(y)\,dy$ $(m_Y, m_Y+\epsilon]$ | $\int_{m_Y+\epsilon}^{\infty} g(y)\,dy$ $(m_Y+\epsilon, \infty)$ |
|---|---|---|---|---|
| $\frac{1}{2}$ | $[0, m_Y]$ | $0$ | $\gamma_1$ | $\gamma_2$ |
| $\frac{1-\epsilon}{2}$ | $[m_Y, m_Y+\epsilon]$ | $-(m_y + \frac{\epsilon}{2})$ | $-(m_Y+\epsilon)$ | $\gamma$ |
| $\frac{\epsilon}{2}$ | $[n-\epsilon, n]$ | $-(n - \frac{\epsilon}{2})$ | $-n$ | $-n$ |

The contribution to the expected value from the first row is

$$\frac{\gamma_1}{2}\left(\int_{m_Y}^{m_Y+\epsilon} g(y)\,dy\right) + \frac{\gamma_2}{2}\left(\int_{m_Y+\epsilon}^{\infty} g(y)\,dy\right) = \frac{\gamma}{4},$$

which cancels with the contribution from the first column since

$$\frac{\gamma}{4} - \left(\frac{1-\epsilon}{4}\right)\left(m_Y + \frac{\epsilon}{2}\right) - \frac{\epsilon}{4}\left(n - \frac{\epsilon}{2}\right) = 0,$$

by definition of $n$. Computing the contribution to the expected payoff from the remaining entries of the matrix, player $A$ wins (or loses)

$$\frac{-n\epsilon}{4} - (m_Y + \epsilon)\left(\frac{1-\epsilon}{2}\right)\int_{m_Y}^{m_Y+\epsilon} g(y)\,dy + \gamma\left(\frac{1-\epsilon}{2}\right)\int_{m_Y+\epsilon}^{\infty} g(y)\,dy. \quad (2)$$

Substituting for $n$, (2) is positive if

$$2\gamma(1-\epsilon)\int_{m_Y+\epsilon}^{\infty} g(y)\,dy > \left(\gamma + \epsilon^2 - \frac{\epsilon}{2} - m_Y + m_Y\epsilon\right) + 2(m_Y+\epsilon)(1-\epsilon)\int_{m_Y}^{m_Y+\epsilon} g(y)\,dy.$$

This inequality holds by the selection of $\epsilon$. Therefore, $X \succ Y$. ∎

**Game-theoretic conclusion** Let's interpret the propositions in this paper game-theoretically. A pair of strategies is a Nash equilibrium if neither player, given knowledge of her opponent's strategy, can improve her outcome by deviating from her strategy. Since the Wallet Game is a zero-sum game, at least one player must have a nonpositive expected payoff. Using the constructions in the propositions, this player can change her (discrete or continuous) distribution to yield a positive expected payoff. So, there does not exist a Nash equilibrium in any of the cases we considered.

In game theory, the fundamental solution concept is the Nash equilibrium. Consequently, the fact that there is no optimal strategy, hence no Nash equilibrium, may seem troubling. It is interesting to note that while the existence of Nash equilibria is often proved by variations or extensions of the Kakutani Fixed Point Theorem, this theorem does not apply here as the hypotheses require the set of strategies to be compact [**4**]. Neither the space of all discrete random variables with fixed means nor the space of all continuous distributions with fixed medians or means are compact.

So what should you do when someone suggests playing the Wallet Game? Since the standard game-theoretic assumption of knowing your opponent's strategy is highly unlikely, the authors advise readers to play the game at their own risk.

## REFERENCES


1. M. Kraitchik, *Mathematical Recreations*, 2nd edition, Dover, New York, 1953.
2. K. Merryfield, N. Viet, and S. Watson, The wallet paradox, *Amer. Math. Monthly* **104** (1997), 647–649.
3. M. Gardner, *Aha! Gotcha*, W. H. Freeman and Company, New York, 1982.
4. D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, Cambridge, MA, 1991.

# Extriangles and Excevians

LARRY HOEHN
Austin Peay State University
Clarksville, TN 37044

The configuration of three squares drawn on the sides of a right triangle is familiar from the Pythagorean theorem. This configuration is shown in FIGURE 1 along with three shaded triangles; however, as we continue, we relax the assumption that $\triangle ABC$ is a right triangle. We call the shaded triangles *extriangles* of $\triangle ABC$ in a manner somewhat analogous to excircles of a triangle. In this paper we consider selected cevians of these extriangles and investigate their relationships to $\triangle ABC$.

Recall that *cevians* are lines through one vertex of a triangle and one point of the opposite side. Named in honor of Giovanni Ceva, these include medians, altitudes, and angle bisectors. The perpendicular bisectors of the sides are not, strictly speaking, cevians; however, they, like the other triples of lines mentioned, are concurrent, with an interesting point of intersection, the circumcenter of the triangle. Along with the cevians of extriangles, we also consider the perpendicular bisectors of the newly created sides.

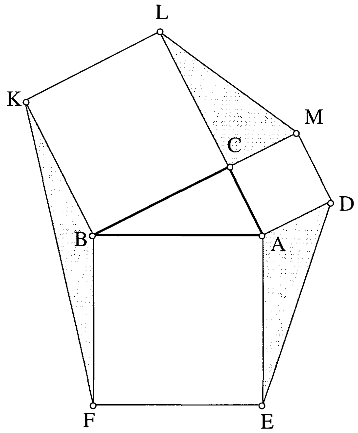



**Figure 1** Pythagorean configuration

First we recall some well-known properties of cevians of triangles. *The three medians, the three angle bisectors, the three altitudes, and the three perpendicular bisectors of the sides of any triangle are concurrent, respectively, at the centroid, the incenter, the orthocenter, and the circumcenter of any triangle. Moreover, the orthocenter, centroid, and circumcenter are collinear and determine the Euler line of the triangle.* These properties and similar facts can be found in standard geometry books such as Coxeter and Greitzer [**1**, p. 19].

Each extriangle enjoys these concurrency and collinearity properties, of course, but we are interested in finding relationships that exist when we consider one cevian from each extriangle. In most cases, we consider cevians of the extriangles that pass through the vertices of the original triangle $\triangle ABC$. For instance, an *exaltitude* of $\triangle ABC$ is an altitude of an extriangle through one of the points $A$, $B$, or $C$. An *exangle bisector* is also an angle bisector of the original triangle. We would have liked to use the term *exmedian*, but one referee pointed out that this term is already used for a line through the vertex of a triangle which is parallel to the opposite side [**3**, p. 176], so we will use *exomedian* for the median of an extriangle.

An *experpendicular bisector* of $\triangle ABC$ does not in general pass through a vertex of $\triangle ABC$, but is the perpendicular bisector of the side of an extriangle that is opposite a vertex of $\triangle ABC$. In FIGURE 5, $AN$ is an exaltitude, $AP$ is an exomedian, $JP$ is an experpendicular bisector, and $AI$ (not drawn) is an exangle bisector of $\triangle ABC$ for extriangle $\triangle ADE$. All of these various lines will be called *excevians* of $\triangle ABC$.

### Concurrency of excevians

THEOREM 1. *The three exaltitudes of any $\triangle ABC$ are concurrent at the centroid of $\triangle ABC$.*

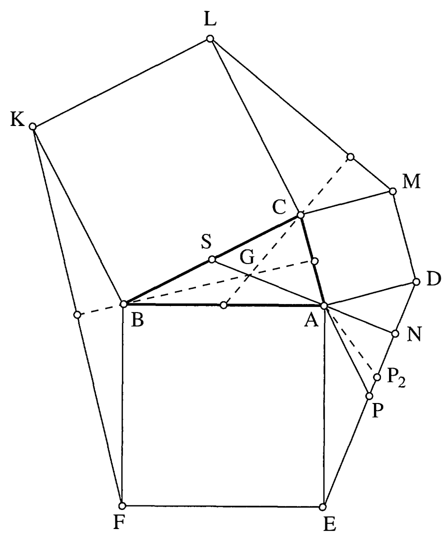



**Figure 2** Concurrency of exaltitudes

*Proof.* Let $AN$ be an exaltitude of $\triangle ABC$ (that is, an altitude of extriangle $\triangle ADE$), and let it meet $BC$ at $S$ as shown in FIGURE 2. We wish to prove that $AS$ is a median of $\triangle ABC$. Since $\angle BAS$ and $\angle EAN$ are complements and since $\angle AEN$ and $\angle EAN$ are complements, then $\angle BAS \cong \angle AEN$. Next construct a point $P$ on $DE$ such that $EP = AS$. Then $\triangle BAS \cong \triangle AEP$ by the side-angle-side congruence for triangles. Therefore $\angle BSA \cong \angle APE$ and $BS = AP$.

In the same manner, note that $\angle CAS \cong \angle ADN$ and construct a point $P_2$ on $DE$ such that $P_2D = AS$. Then $\triangle CAS \cong \triangle ADP_2$, which implies that $\angle CSA \cong \angle AP_2D$ and $CS = AP_2$. Therefore $m\angle APE + m\angle AP_2D = m\angle BSA + m\angle CSA = 180°$ so that $P = P_2$. Since $EP = AS = PD$, then $P$ is the midpoint of $DE$.

Since $BS = AP$ and $SC = AP_2 = AP$, then $BS = SC$. Thus $S$ is the midpoint of $BC$ which implies that $AS$ is a median of $\triangle ABC$.

Repeating this argument, we find that the exaltitudes at $B$ and $C$ of $\triangle ABC$ are also the medians of $\triangle ABC$. Since the medians of any triangle are concurrent at the centroid $G$, the three exaltitudes are concurrent at $G$. ■

It is worth noting that since $\triangle BAS \cong \triangle AEP$ and $\triangle CAS \cong \triangle ADP$, the extriangle $\triangle ADE$ and the original $\triangle ABC$ have equal areas. Indeed, all three extriangles have the same area.

As an aside we note that this proof is reminiscent of a theorem of Brahmagupta [**1**, p. 59]: *If a quadrilateral is cyclic and has perpendicular diagonals [orthodiagonal],* *then the perpendicular from the point of intersection of the diagonals to a side bisects the opposite side.* It is easy to show that quadrilateral $BCDE$ of FIGURE 2 is orthodiagonal, but it is not necessarily cyclic.

THEOREM 2. *The three exomedians of any $\triangle ABC$ are concurrent at the orthocenter of $\triangle ABC$.*

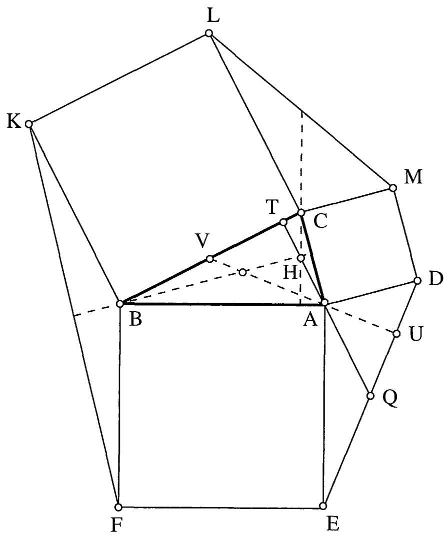



**Figure 3** Concurrency of exomedians

*Proof.* Let $AQ$ be the exomedian of $\triangle ABC$ that is also the median of extriangle $\triangle ADE$; let $AQ$ meet $BC$ at $T$ (see FIGURE 3). Construct the exaltitude $AU$ that meets $BC$ at $V$. By the proof of Theorem 1, $AV$ is a median of $\triangle ABC$ and $\triangle BVA \cong \triangle AQE$. Thus $\angle BVA \cong \angle AQE$ so that $m\angle VTA + m\angle TAV = m\angle QUA + m\angle UAQ$. But $\angle TAV$ is congruent to $\angle UAQ$ (vertical angles). Therefore $\angle VTA \cong \angle QUA$. Since $\angle QUA$ is a right angle, $\angle VTA$ is as well, so $AT$ is an altitude of $\triangle ABC$.

Similarly, the exomedians through $B$ and $C$ are shown to be altitudes of $\triangle ABC$. Since the altitudes of any triangle are concurrent at the orthocenter $H$, then the three exomedians are concurrent at $H$. The referees observed that because of symmetry, Theorem 1 and Theorem 2 are equivalent. ■

THEOREM 3. *The three exangle bisectors of any $\triangle ABC$ are concurrent at the incenter $I$ of $\triangle ABC$.*

The proof of Theorem 3 is trivial so we turn to the fourth of our cevians.

THEOREM 4. *The three experpendicular bisectors of any $\triangle ABC$ are concurrent.*

*Proof.* We begin with $\triangle ABC$, as in FIGURE 4. Let $M'$ be the midpoint of $AB$ and extend median $CM'$ to a point $\tilde{C}$ such that $M'\tilde{C} = CM'$. Join $\tilde{C}$ and $A$ to form $\triangle CA\tilde{C}$ which has sidelengths of $b$, $a$, and $2m_c$. Next rotate $\triangle CA\tilde{C}$ 90° clockwise about center $A$ to get $\triangle SAT$. Thus $AT \perp BC$, $AS \perp AC$, and $TS \perp CM'$. Translate $\triangle SAT$ so that $A$ ends at $G$ forming $\triangle B'GA'$. Thus we note that if we start with $\triangle ABC$ and construct $GA' \perp BC$ with $GA' = BC$ and construct $GB' \perp AC$ with $GB' = AC$, then $A'B' \perp GC$ and $A'B' = 2m_c$.

If we repeat the above process on the other sides of $\triangle ABC$, we obtain $\triangle A'B'C'$ with sides of length $2m_a$, $2m_b$, and $2m_c$. Therefore $\triangle A'GB'$ is obtained from extriangle $\triangle LCM$ (see FIGURE 5) by translating it in the direction $CG$ so that $C$ ends

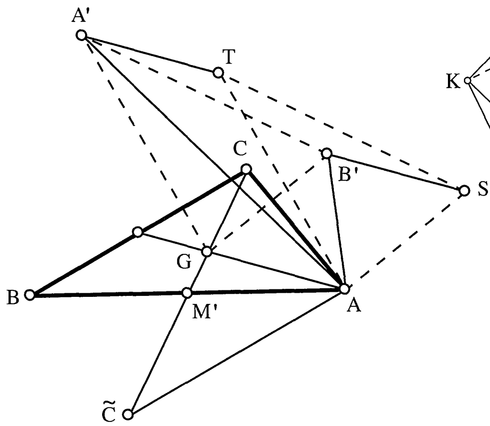



**Figure 4** Rotation and translation

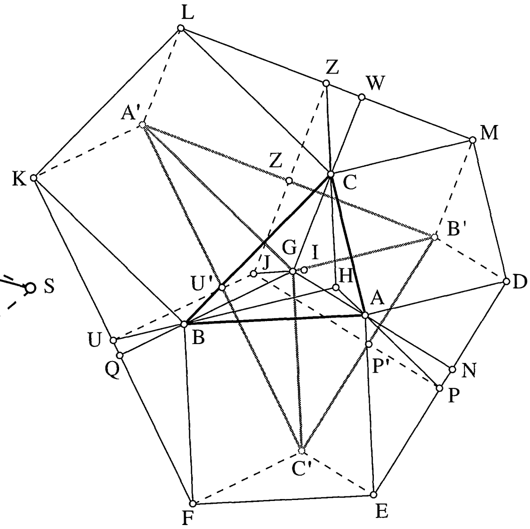



**Figure 5** Concurrency of experpendicular bisectors

at $G$. Since $A'B' \perp GC$, then $LA'B'M$ is a rectangle. In the same manner $\triangle B'GC'$ and $\triangle C'GA'$ are translates of extriangles $\triangle DAE$ and $\triangle FBK$, respectively, so that $DB'C'E$ and $FC'A'K$ are rectangles.

Next construct the perpendicular bisectors of $DE$, $FK$, and $LM$. Since these are perpendicular bisectors of the sides of $\triangle A'B'C'$, then they are concurrent at the circumcenter $J$. Hence the experpendicular bisectors of $\triangle ABC$ are concurrent. ∎

As an aside we note the nice dual relationship between $\triangle ABC$ and $\triangle A'B'C'$. They have a common centroid, the sides of each are perpendicular to the medians of the other, and the lengths of the medians of each are proportional to the lengths of the sides of the other. It is of interest to compare this relationship with the dual relationships of triangles described by Pedoe [**4**].

The discussion also provides a neat proof of the relationship between the area of a triangle and the area of the triangle formed by its medians. Since $\triangle A'B'C'$ has sides of length $2m_a$, $2m_b$, and $2m_c$, we see that the area of $\triangle A'B'C'$ is four times that of the triangle whose sides have the lengths of the medians of $\triangle ABC$. Since $area(\triangle A'B'C') = 3\, area(\triangle ABC)$, then the triangle whose sides are the medians has 3/4 the area of $\triangle ABC$. Moreover, if $\sigma = (m_a + m_b + m_c)/2$, then Heron's area formula applied to the triangle of medians leads to the expression for the area of $\triangle ABC$ in terms of its medians; namely,

$$area(\triangle ABC) = \frac{4}{3}\sqrt{\sigma(\sigma - m_a)(\sigma - m_b)(\sigma - m_c)}.$$

This formula can also be found in Hobson [**2**, p. 201], with a trigonometric proof.

**Collinearity** Recall that the orthocenter $H$, centroid $G$, and circumcenter $O$ of any triangle are collinear, all lying on the *Euler line*. What collinearity might be possible among the points of concurrency for the exaltitudes, $G$, exomedians, $H$, exangle bisectors, $I$, and experpendicular bisector $J$? It is known that in general $I$ does not lie on the Euler line [**5**]; might there be a line that contains $I$, $G$, and $J$? In FIGURE 5, these points appear collinear.

Our experiments with the software *Geometer's Sketchpad* seem to indicate that the incenter $I$ lies on the line $JG$. Altering $\triangle ABC$ to have sides of various lengths we find that $JG + GI - JI = 0.000$ inches, most of the time, and 0.001 inches, rarely. It is tempting to guess that the 0.001 must be due to roundoff errors.

Alas, empirical evidence leads us astray. A referee provided the following excellent counterexample: If $\triangle ABC$ has vertices $(6, 0)$, $(0, 0)$, and $(0, 12)$, then $G = (2, 4)$, $J = (1, 11)$, and $I = (9 - 3\sqrt{5}, 9 - 3\sqrt{5})$. Since the slopes of $GJ$ and $JI$ are not equal, then $G$, $J$, and $I$ cannot be collinear.

For readers interested in investigating the location of $J$, we mention this: Since $J$ and $G$ are the circumcenter and centroid, respectively, of $\triangle A'B'C'$, $JG$ is the Euler line of this triangle.

**Acknowledgment.** The author wishes to thank the referees for their knowledge and assistance.

## REFERENCES

1. H. S. M. Coxeter and S. L. Greitzer, *Geometry Revisited*, Random House, 1967.
2. E. W. Hobson, *A Treatise on Plane and Advanced Trigonometry*, Dover Publications, Inc., New York, 1928.
3. R. A. Johnson, *Modern Geometry: An Elementary Treatise on the Geometry of the Triangle and the Circle*, Houghton Mifflin, Boston, 1929.
4. D. Pedoe, On (what should be) a well-known theorem in geometry, *Amer. Math. Monthly*, **74** (1967), 839–841.
5. Anthony Varilly, Location of incenters and Fermat points in variable triangles, this MAGAZINE, **74**:2 (2001), 123–129.

# Boxlike Domains in the Complex Plane

JOHN A. FEROE
BENJAMIN A. LOTTO
CHARLES I. STEINHORN
Vassar College
Poughkeepsie, NY 12604

The interplay between the geometry of a domain in the complex plane and the analytic properties of holomorphic (sometimes called analytic) functions defined on that domain is central in complex analysis. A fundamental example of this interplay is the following characterization of simply connected domains given by one version of Cauchy's Theorem and Morera's Theorem: A domain $G$ is simply connected if and only if $\int_\gamma f = 0$ for every holomorphic function $f$ defined on $G$ and every simple closed rectifiable curve $\gamma$ in $G$. Other geometric properties of domains can also be characterized analytically. For example, convex and starlike domains can be identified by examining the analytic properties of their Riemann maps [**1**, pp. 40–43].

In this paper, we introduce a new class of domains that arise naturally from a generalization of one proof of Cauchy's Theorem. We call these *boxlike* domains. After making the appropriate definitions and proving the generalization, we give a geometric characterization of boxlike domains. Finally, we derive from this characterization a heuristic for identifying boxlike domains just by looking at them.

The inspiration for the definition of boxlike domains comes from the treatment of Cauchy's Theorem in one widely used Complex Analysis text. It starts with the following Lemma [**2**, pp. 123, 125–127]. We adopt the convention for the rest of this paper that *rectangle* means *rectangle with sides parallel to the coordinate axes.*

LEMMA 1. *Suppose that $\gamma$ is a rectangle and that $f$ is analytic on a domain that includes $\gamma$ and its interior. Then $\int_\gamma f = 0$.*

The proof of Lemma 1 uses an elegant idea due to Goursat, in which the rectangle is subdivided repeatedly. In the limit, the subdivided rectangles intersect in a single point $z$. The existence of $f'(z)$, combined with the right estimates, leads to the desired result.

From Lemma 1, we can easily derive Cauchy's Theorem for a disk.

THEOREM 1. *Suppose that $f$ is holomorphic on a disk $D$. Then $f$ has an antiderivative on $D$ and $\int_\gamma f = 0$ for any closed curve $\gamma$ that lies in $D$.*

The proof of Theorem 1 starts by defining a function $F(z)$ at a point $z$ in $D$ as the integral of $f$ along a rectangular path from the center of the disk to $z$. Lemma 1 enables us to write the difference quotient $(F(z + \Delta z) - F(z))/\Delta z$ as an integral of $f$ over a rectangular path from $z$ to $z + \Delta z$. Standard estimates show that this converges to $f(z)$ as $\Delta z$ tends to zero. Hence $F'(z) = f(z)$. An application of the Fundamental Theorem of Calculus now shows that integrals of $f$ over closed paths must equal zero.

What property of the disk makes this proof work? When we draw a rectangle connecting the center of the disk to any other point in the disk, the rectangle and its interior lie entirely within the disk and Lemma 1 can be applied. If $G$ is a domain that enjoys this same property, then the proof of Theorem 1 applies verbatim and yields Cauchy's Theorem for $G$. Here are the relevant definitions.

DEFINITION 1. *Given two points $z$ and $w$ in the complex plane, the rectangle spanned by $z$ and $w$ is the rectangle (with sides parallel to the coordinate axes) that has opposite corners at $z$ and $w$. We write $R(z, w)$ for this rectangle together with its interior.*

If $z$ and $w$ lie on a line parallel to the one of the coordinate axes, then $R(z, w)$ degenerates into a line segment. We still consider this to be a rectangle, which happens to have zero height or width.

DEFINITION 2. *A domain $G$ is said to be boxlike with respect to a point $z$ if $R(z, w) \subset G$ for every $w$ in $G$. In this case, we will call $z$ the boxlike center of $G$.*

*Finally, we say that a domain $G$ is boxlike if it is boxlike with respect to some point.*

Note that a domain may have many boxlike centers. A simple example illustrating this is an open rectangle, for which every point in the domain is a boxlike center.

In view of the remarks preceding Definition 2, the proof of Theorem 1 yields the following, stronger result.

THEOREM 2. *Suppose that $f$ is holomorphic on a boxlike domain $G$. Then $f$ has an antiderivative on $G$ and $\int_\gamma f = 0$ for any closed curve $\gamma$ that lies in $G$.*

Have we really accomplished anything by making this generalization? We address this question by deriving a characterization of boxlike domains. As one consequence of this characterization, we see that the class of boxlike domains includes a whole lot more than just disks. Hence Theorem 2 is a significant generalization of Theorem 1. As a second consequence, we derive a simple, practical, and effective method for recognizing whether a given domain is boxlike, making Theorem 2 easy to apply to particular domains.

Before proceeding to the characterization, let's get a better idea of what boxlike domains look like. First, recall that a domain is *starlike* with respect to a point $z$ if any segment that connects $z$ to another point in the domain lies entirely within the domain. Clearly, a domain that is boxlike with respect to $z$ is also starlike with respect to $z$.

To see this, let $w$ be any other point in the domain. Then $R(z, w)$ is contained in the domain. But the segment that connects $z$ to $w$ is a diagonal of $R(z, w)$. Hence that segment is contained in the domain, which tells us that the domain is starlike.

Second, recall that a domain is *convex* if any segment that connects two points in the domain lies entirely in the domain. A convex domain is clearly starlike with respect to any point in the domain. Does convexity necessarily imply a boxlike nature, or conversely? In fact, neither implies the other as the examples in FIGURE 1 show.

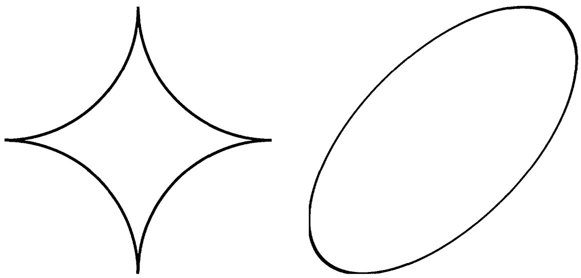


**Figure 1** The domain on the left is boxlike (with respect to its "center") but not convex. The domain on the right is convex but not boxlike. To see this, observe that the rectangle spanned by any fixed point and a suitable point near the boundary of the domain is not contained in the domain.

We are now ready to state and prove the above-mentioned characterization of boxlike domains.

The actual statement is somewhat technical looking, but the idea of the characterization is quite simple. It says that a boxlike domain lies between the graphs of two functions that are defined on a common interval. (This way of thinking about domains in the plane is familiar from finding limits of integration in iterated integrals.) The function describing the top must rise and then fall and the function describing the bottom must fall and then rise, with the two functions reversing course at the same point.

THEOREM 3. *Let $G$ be a domain in the complex plane. Then $G$ is boxlike with respect to the point $z_0 = x_0 + iy_0$ if and only if there exist extended real numbers $a < b$ and extended real-valued functions $f(x)$ and $g(x)$ defined on $(a, b)$ such that*

1. $-\infty \leq a < x_0 < b \leq \infty$;
2. $g(x) < y_0 < f(x)$ *on* $(a, b)$;
3. $f(x)$ *is monotonic increasing on* $(a, x_0]$ *and monotonic decreasing on* $[x_0, b)$;
4. $g(x)$ *is monotonic decreasing on* $(a, x_0]$ *and monotonic increasing on* $[x_0, b)$; *and*
5. $G = \{z : z = x + iy, a < x < b \text{ and } g(x) < y < f(x)\}$.

*Proof.* Note that $G$ is boxlike with respect to $z_0$ if and only if $G - z_0 = \{z - z_0 : z \in G\}$ is boxlike with respect to 0, so we may assume without loss of generality that $z_0 = 0$.

To prove necessity, suppose that there exist $a < b$ and $f(x)$, $g(x)$ that satisfy conditions 1–5. Let $z = x + iy$ be a given point in $G$. Assume first that $x \geq 0$ and $y \geq 0$. Then $x < b$ and $y < f(x)$ by 5. Let $w = u + iv$ be any point in $R(0, z)$. Then $0 \leq u \leq x$ and $0 \leq v \leq y$. Since $f$ is decreasing on $[0, b)$, it follows that $f(u) \geq f(x)$. Hence $g(u) < 0 \leq v \leq y < f(x) \leq f(u)$, the first inequality holding by 2. It follows that $w \in G$ by 5. The other three cases, where $x \leq 0$ and/or $y \leq 0$, are handled similarly.

As to sufficiency, suppose we are given a domain $G$ that is boxlike with respect to 0. Let $z = x + iy$ be any point in $G$. Since $R(0, z) \subset G$, it follows that the closed interval with endpoints 0 and $x$ along the real axis is contained in $G$. We conclude that $G \cap \mathbb{R}$ is an open interval. Define $a$ and $b$ to be the left and right endpoints of this interval; we allow for the possibility that $a = -\infty$ and/or $b = \infty$. Since $0 \in G$, we have $a < 0 < b$.

Fix $x \in (a, b)$. Using the definition of boxlike as in the preceding paragraph, it follows that the set $\{y : x + iy \in G\}$ is an open interval. Define $g(x)$ and $f(x)$ to be the left and right endpoints of this interval. Again we allow for the possibility that $g(x) = -\infty$ and/or $f(x) = \infty$. Since $x$ itself is in $G$, we have $g(x) < 0 < f(x)$.

We have established conditions 1 and 2 of the theorem; condition 5 follows immediately from the previous two paragraphs. To finish the proof, we show that $f(x)$ is decreasing on $[0, b)$; similar arguments establish the other monotonicity properties stated in conditions 3 and 4. Suppose that $0 \leq x < x' < b$. If $y < f(x')$, then the definition of $f(x')$ yields that $x' + iy \in G$. But $x + iy \in R(0, x' + iy) \subset G$ since $G$ is boxlike. Hence $f(x) \geq y$ by definition. Since $y$ is arbitrary subject to the condition $y < f(x')$, it follows that $f(x) \geq f(x')$. The proof is thus complete. ∎

Note that we can reverse the roles of $x$ and $y$ to obtain a characterization of boxlike domains in terms of functions of $y$ bounding the domain on the left and right; this is similar to reversing limits of integration in iterated integrals. We leave the details to the reader.

Theorem 3 allows us to construct more exotic boxlike domains. One such example is given in FIGURE 2.

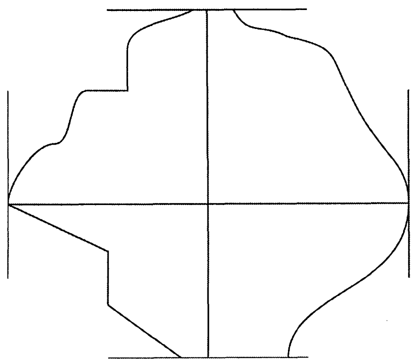


**Figure 2** A complicated boxlike domain

Now, using Theorem 3, we describe the heuristic promised above for recognizing a boxlike domain $G$ as boxlike or not just by inspection.

Assume for the moment that $G$ is a bounded domain. First, find vertical lines that bound $G$ as closely as possible on the left and right, a step familiar from writing a double integral as an iterated integral. Let these best-bounding lines be $x = a$ and $x = b$ (if $G$ is boxlike). The boundary of $G$ will touch each of these vertical lines in one or more points. If there are a pair of points that lie on the intersection of these lines with the boundary of $G$ and share the same $y$-coordinate $y_0$, then $y_0$ is the $y$-coordinate of a possible boxlike center. If no such pair exists, then $G$ cannot be boxlike.

Next we apply a similar procedure using horizontal lines that bound $G$ as closely as possible on the top and bottom. As before, we look for a pair of points that are on the boundary of $G$, on these lines, and line up vertically. If such a pair exists, they identify the $x$-coordinate of a possible boxlike center; if not, then $G$ is not boxlike.

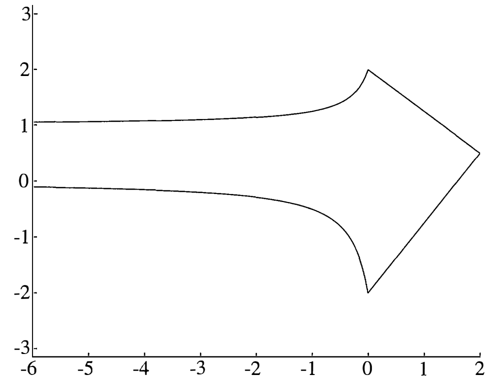



**Figure 3** This unbounded boxlike domain touches $x = -\infty$, the best vertical bounding line on the left, at all $y$ in $[0, 1]$. The boxlike center is identified by the heuristic to be $0.5i$.

If we've gotten this far, then we've identified a candidate for the boxlike center of $G$. The final step is to check that the boundary of $G$ can be described by graphs of functions $f$ above and $g$ below. If so, and if these functions obey the proper monotonicity conditions, then $G$ is boxlike!

As one example of this heuristic, consider the domain on the right in FIGURE 1. In this case, the best-bounding vertical lines touch the domain at points that don't align horizontally; hence, the domain cannot be boxlike. On the other hand, the domain in Figure 2 shows how easily a boxlike domain is identified using this heuristic—the best-bounding horizontal and vertical lines are shown, as are resulting lines that cross at a (in this case unique) boxlike center.

An important point is that the functions $f$ and $g$ from Theorem 3 need not be continuous. The example in FIGURE 2 uses discontinuous bounding functions. It is easy to see that a set $G$ defined as in Theorem 3 is open if and only if $f$ is lower semicontinuous and $g$ is upper semicontinuous.

The same procedure works for unbounded domains, so long as we interpret the above terms appropriately. For example, if $G$ is unbounded on the left, then the "best-bounding line" for $G$ on the left is $x = -\infty$. The domain is said to touch this line at $y = c$ if there is a sequence $x_n + iy_n$ in $G$ such that $x_n \to -\infty$ and $y_n \to c$. See FIGURE 3 for an example.

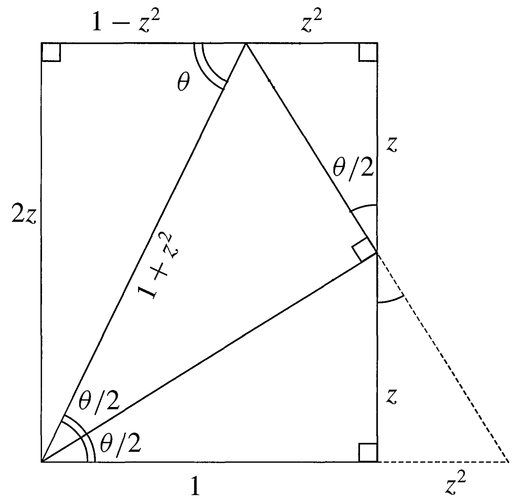
As a final note, we should point out that Lemma 1 can be proved for triangles instead of rectangles, in which case the proof of Theorem 1 immediately yields Cauchy's Theorem for all starlike (and hence all boxlike) domains (see [**3**, pp. 205–6]). The pedagogical price paid for the triangular approach is that one needs to develop the finite intersection property for compact sets in $\mathbb{R}^2$. The rectangular approach, on the other hand, requires only the nested interval property in $\mathbb{R}$.

## REFERENCES

1. Peter L. Duren, *Univalent Functions*, Grundlehren der mathematischen Wissenschaften **259**, Springer-Verlag, New York, 1983.
2. Jerrold E. Marsden and Michael J. Hoffman, *Basic Complex Analysis*, W. H. Freeman, New York, 3rd edition, 1999.
3. Walter Rudin, *Real And Complex Analysis*, McGraw-Hill, New York, 3rd edition, 1987.

# Proof Without Words: The Weierstrass Substitution

SIDNEY H. KUNG
University of North Florida
Jacksonville, FL 32224-2645




$$z = \tan\left(\frac{\theta}{2}\right) \quad \Rightarrow \quad \sin(\theta) = \frac{2z}{1+z^2}, \quad \cos(\theta) = \frac{1-z^2}{1+z^2}$$

# Integrals of Periodic Functions

SEAN F. ELLERMEYER
DAVID G. ROBINSON
Kennesaw State University
Kennesaw, GA 30144-5591
sellerme@kennesaw.edu
drobinso@kennesaw.edu

Computing integrals of powers of the sine function is a standard exercise in calculus. Using integration by parts or some basic trigonometric identities, the student discovers that

$$\int \sin^2 t \, dt = -\frac{1}{2}\cos t \sin t + \frac{1}{2}t + C \tag{1}$$

and

$$\int \sin^3 t \, dt = -\frac{1}{3}\sin^2 t \cos t - \frac{2}{3}\cos t + C. \tag{2}$$

Further investigation reveals that all even powers of sine have integrals containing periodic terms and a (nontrivial) linear term, whereas all odd powers of sine have integrals with only periodic terms.

In this note, we show that the first integral is representative of the integral of any periodic function. Although this fact (Proposition 1) is not pointed out in any of the calculus textbooks we have studied, we feel that it is elementary enough and of sufficient usefulness in applications to be given attention in an elementary calculus course. As we show by examples, Proposition 1 can come in handy not only in knowing what to expect when we integrate a periodic function, but also in enabling us to introduce some qualitative aspects of differential equations (the existence of periodic solutions) at the elementary calculus level.

Throughout, when we say that $f$ *is periodic of period* $T > 0$, we mean that $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(t+T) = f(t)$ for all $t$, but not necessarily that $T$ is the smallest positive period possessed by $f$. If $f$ has period $T$, then we define the *average value of* $f$ in the usual way, as

$$\overline{f} = \frac{1}{T}\int_t^{t+T} f(s)\,ds.$$

Since $f$ has period $T$, the value of $\overline{f}$ does not depend on the choice of $t \in \mathbb{R}$.

PROPOSITION 1. *If* $f : \mathbb{R} \to \mathbb{R}$ *is continuous and periodic of period* $T > 0$, *then*

$$\int f(t)\,dt = g(t) + \overline{f}t + C, \tag{3}$$

*where* $g$ *is a periodic function of period* $T$.

To prove the proposition, we let

$$g(t) = \int_0^t f(s)\,ds - \overline{f}t.$$

Then

$$\frac{d}{dt}\big(g(t) + \overline{f}t\big) = f(t),$$

which implies (3). Also, for any $t \in \mathbb{R}$, we have

$$\begin{aligned} g(t+T) &= \int_0^{t+T} f(s)\,ds - \overline{f}T - \overline{f}t \\ &= \int_0^{t+T} f(s)\,ds - \int_t^{t+T} f(s)\,ds - \overline{f}t \\ &= g(t), \end{aligned}$$

which shows that $g$ is periodic of period $T$. ■

Since Proposition 1 is qualitative (rather than quantitative) in nature, it should be expected that applications of the Proposition will yield mainly qualitative information.

**Example 1 (Integrals of Powers of Sine):** By Proposition 1, if $n$ is a positive integer, then

$$\int \sin^n t\,dt = g(t) + mt + C,$$

where $g$ is periodic of period $2\pi$ and $m$ is the average value of $f(t) = \sin^n t$. Since even powers of sine have positive average value and odd powers of sine have zero average value, the constant $m$ that appears in the linear term will be positive when $n$ is even and zero when $n$ is odd. This agrees with the results shown in (1) and (2).

**Example 2 (The Harmonic Oscillator):** The motion of a unit mass attached to a spring is described by the differential equation

$$y'' + ay' + by = 0, \tag{4}$$

where $y(t)$ is the position of the spring at time $t$, $-by$ is the force exerted on the mass by the spring, and $-ay'$ the damping force exerted on the mass by the medium through which it is moving. (No other force, including gravity, is assumed to be present.) In elementary differential equations courses, the model (4) is often used to motivate the study of linear differential equations and linear systems. The usual approach first handles the case $a = 0$ (the undamped case), showing that the mass exhibits sustained oscillations about its rest position (assuming also that $b > 0$). Once this is done, it is shown that when $a > 0$, the mass eventually comes to rest (possibly via decaying oscillations) due to the damping force. In the setting of a differential equations course, the fact that the real parts of the roots of the characteristic equation of (4) are negative shows that the mass approaches equilibrium when $a > 0$. In what follows, we use Proposition 1 to explain why equation (4) cannot have nontrivial periodic solutions if $a \neq 0$. Our argument does not delve into the specific forms of solutions of (4) and hence does not yield the quantitative information that would follow from a more detailed analysis, but our approach does reveal the qualitative effect of a damping force.

If $y$ is a periodic solution of (4) with period $T > 0$, then $y'$ is also periodic of period $T$. Multiplying both sides of (4) by $y'$, we obtain

$$y'y'' + a\left(y'\right)^2 + byy' = 0.$$

Integrating this from time 0 to any positive time $t$ gives

$$\left(y'(t)\right)^2 - \left(y'(0)\right)^2 + 2a\int_0^t \left(y'(s)\right)^2 ds + b\left(y(t)\right)^2 - b\left(y(0)\right)^2 = 0.$$

Since $(y')^2$ is periodic of period $T$, Proposition 1 implies

$$\int_0^t \left(y'(s)\right)^2 ds = g(t) - g(0) + mt,$$

where $g$ is periodic of period $T$ and $m$ is the average value of $(y')^2$. This gives us

$$\left(y'(t)\right)^2 - \left(y'(0)\right)^2 + 2ag(t) - 2ag(0) + b\left(y(t)\right)^2 - b\left(y(0)\right)^2 = -2amt. \tag{5}$$

Since the left-hand side of (5) is periodic of period $T$, then so is $-2amt$. Of course, this is only possible if either $a = 0$ or $m = 0$. If $a$ isn't zero, then $m$ is, and in fact $y'$ is identically zero (because $m$ is the average value of $(y')^2$, which is nonnegative). This in turn implies that $y$ is constant (i.e., $y$ is trivially periodic of period $T$). But then (4) reduces to $by = 0$. Hence, if we assume in addition that $b \neq 0$ (which would be reasonable in a spring problem), then it follows that $y(t) = 0$ is the only periodic solution of (4).

Finally, if $a = 0$, $b \neq 0$, and $y$ is a nontrivial periodic solution of (4), then $y$ cannot be of constant sign for all $t$ (because integration of equation (4) from 0 to $T$ would yield a contradiction). Hence, we may assume without loss of generality that $y(0) > 0$

is the maximum value of $y$. Equation (5) then gives us

$$(y(t))^2 = (y(0))^2 - \frac{1}{b}\left(y'(t)\right)^2.$$

If $b < 0$, then we would have $(y(t))^2 > 0$ for all $t$, which would mean that $y$ cannot change sign. We conclude that the differential equation (4) has nontrivial periodic solutions only if $a = 0$ and $b > 0$.

The same approach can be extended to the study of similar nonlinear differential equations such as the unforced *Duffing's equation*,

$$y'' + ay' + by + cy^3 = 0,$$

which is usually considered in more advanced differential equations courses. (See, for example, Hale [**2**, p. 168].)

**Example 3 (Linear Systems with Constant Coefficients):** We conclude by using Proposition 1 to show that if $abcd < 0$, then the linear system of differential equations

$$\begin{aligned} x' &= ax + by \\ y' &= cx + dy \end{aligned} \tag{6}$$

has no nontrivial periodic solutions. Conditions that determine the qualitative nature of solutions of (6) are usually given in differential equations courses in terms of the trace $(a + d)$ and the determinant $(ad - bc)$ of the coefficient matrix of (6) [**1**, p. 312]. These are obtained by analyzing the characteristic equation of the system. As a contrast, we argue as we did for the harmonic oscillator.

If we assume that $abcd < 0$ and that $(x, y)$ is a periodic solution of (6) of period $T$, then we can write

$$\begin{aligned} cxx' &= acx^2 + bcxy \\ -byy' &= -bcxy - bdy^2 \end{aligned}$$

to obtain

$$cxx' - byy' = acx^2 - bdy^2.$$

Integration of both sides of the above equation from 0 to $t$ gives

$$c\,(x(t))^2 - c\,(x(0))^2 - b\,(y(t))^2 + b\,(y(0))^2 = g(t) - g(0) + mt, \tag{7}$$

where $g$ is periodic of period $T$ and $m$ is the average value of $2acx^2 - 2bdy^2$. Since all terms in equation (7) must be periodic of period $T$, we conclude that $m = 0$. Also, since $(ac)(-bd) > 0$, then $ac$ and $-bd$ must have the same sign, which means that $x$ and $y$ must both be identically 0 on $[0, T]$, and hence on all of $\mathbb{R}$.

## REFERENCES

1. P. Blanchard, R. L. Devaney, and G. R. Hall, *Differential Equations*, Brooks/Cole, Boston, 1998.
2. J. K. Hale, *Ordinary Differential Equations*, Robert E. Krieger, Malabar, FL, 1980.

# A Special Case of Dirichlet's Theorem on Primes in an Arithmetic Progression

HILLEL GAUCHMAN
Eastern Illinois University
Charleston IL 61920

Dirichlet's theorem asserts that *every* arithmetic progression $m, m+n, m+2n, \ldots$, with $m$ and $n$ relatively prime, contains infinitely many primes. The simplest proofs are analytic, using properties of Dirichlet L-series [**1**], although Atle Selberg gave a complicated elementary proof in 1949 [**5**]. Certain individual cases, such as $3, 3+4, 3+8, \ldots$ and $5, 5+6, 5+12, \ldots$, are easy to prove. Other special cases, notably $1, 1+4, 1+8, \ldots$, can be proved using simple properties of quadratic residues [**1**]. In this note, we use elementary arguments to cover an infinite number of cases. While these have been given other elementary proofs (see, for instance, Dickson [**2**, vol. 1, p. 418] or Ribenboim [**4**, p. 268]), the proof presented here is the simplest and shortest that the author knows.

In his recent paper in this MAGAZINE [**3**], Lionel Levine proved the following interesting theorem:

THEOREM A. *Let $f$ be any function from a set $S$ to itself such that $f^n$ (the nth iterate of $f$) has finitely many fixed points for every $n$. If $T(n)$ is the number of points fixed under $f^n$, then*

$$n \mid \sum_{d|n} \mu\left(\frac{n}{d}\right) T(d)$$

*for all positive integers $n$. (Here, $\mu$ is the Möbius function: $\mu(p) = -1$ when $p$ is prime, $\mu(p^m) = 0$ for $m \geq 2$, and $\mu(ab) = \mu(a)\mu(b)$ when $a$ and $b$ are coprime).*

Levine used Theorem A to prove a generalized form of Fermat's Little Theorem: For all positive integers $n$ and $k$,

$$n \mid \sum_{d|n} \mu\left(\frac{n}{d}\right) k^d.$$

He did so by taking $S$ to be the set of complex numbers and $f(z) = z^k$. In this note we show that for a different choice of $S$ and $f : S \to S$, Theorem A gives a simple and elementary proof of a special case of Dirichlet's theorem on primes in an arithmetic progression. Namely, we prove that every arithmetic progression with first term 1 contains infinitely many primes.

We begin with a lemma that may be of interest in itself.

LEMMA. *Let $a$ and $n$ be integers greater than 1, let $n = p_1^{\alpha_1} \ldots p_k^{\alpha_k}$ be a representation of $n$ as a product of primes, and let $q$ be a common divisor of*

$$\frac{a^n - 1}{a^{n/p_1} - 1}, \ldots, \frac{a^n - 1}{a^{n/p_k} - 1}.$$

*Then $n$ divides $(a^n - 1)(q - 1)/q$.*

*Proof.* For integers $x_1, \ldots x_n$ satisfying $0 \le x_i \le a-1$ for $i = 1, \ldots, n$, the expression $(x_1 \ldots x_n)_a := x_1 a^{n-1} + \cdots + x_{n-1}a + x_n$ is called an *$n$-digit number written in the base $a$*. Let us define $S$ to be the set of all $n$-digit numbers $x = (x_1 \ldots x_n)_a$ written in the base $a$ that are not divisible by $q$. For any $n$-digit number $x = (x_1 \ldots x_n)_a$ written in base $a$ we define $f(x) = (x_2 \ldots x_n x_1)_a$. Then $f(x) = ax - x_1(a^n - 1)$. Since $q \mid a^n - 1$, $f$ is a map from $S$ into itself.

Assume that for some $n$-digit number $x$ written in the base $a$, we have $f^{n/p_i}(x) = x$ for $1 \le i \le k$. Then $x$ has the form

$$x = \Big(\underbrace{x_1 \ldots x_{\frac{n}{p_i}}}\,\underbrace{x_1 \ldots x_{\frac{n}{p_i}}} \ldots \underbrace{x_1 \ldots x_{\frac{n}{p_i}}}\Big)_a$$

$$= \left(x_1 \ldots x_{\frac{n}{p_i}}\right)_a \left(1 + a^{\frac{n}{p_i}} + \cdots + a^{(p_i-1)\frac{n}{p_i}}\right) = \left(x_1 \ldots x_{\frac{n}{p_i}}\right)_a \cdot \frac{a^n - 1}{a^{\frac{n}{p_i}} - 1}.$$

Hence $x$ is divisible by $q$, so $x \notin S$. If $d$ is a divisor of $n$ such that $d < n$ and if $f^d(x) = x$, then for some $i$ with $1 \le i \le k$, we have $f^{n/p_i}(x) = x$, and once more we get that $x \notin S$. Thus, using the notation from Theorem A, we find that $T(d) = 0$ for all divisors $d$ of $n$ such that $d < n$, and Theorem A implies that $n \mid T(n)$. It is easy to see that

$$T(n) = |S| = a^n - 1 - \frac{a^n - 1}{q} = \frac{(a^n - 1)(q - 1)}{q}.$$

Hence $n$ divides $(a^n - 1)(q - 1)/q$, and the proof is complete. ■

We can now prove our main result.

THEOREM 1. *Let $n$ be a positive integer. There are infinitely many primes of the form $1 + ny$, where $y$ is a positive integer.*

*Proof.* For $n = 1$ the theorem is obvious. Let $n > 1$. Assume that the theorem is false and let $q_1, \ldots, q_s$ be all primes of the form $1 + ny$. Let $n = p_1^{\alpha_1} \ldots p_k^{\alpha_k}$ be the representation of $n$ as a product of primes. Consider the polynomials

$$\frac{x^n - 1}{x^{\frac{n}{p_1}} - 1}, \ldots, \frac{x^n - 1}{x^{\frac{n}{p_k}} - 1}. \tag{1}$$

Since all polynomials (1) have a mutual root $x = e^{2\pi i/n}$, there exists a polynomial $g(x)$ of positive degree, with integer coefficients and positive leading coefficient, such that $g(x)$ divides each polynomial (1). For $x = 0$, the value of each polynomial in (1) is 1, so $g(0) = \pm 1$, and it follows that $\gcd(g(x), x) = 1$ for each integer $x$. Since the leading coefficient of $g(x)$ is positive, there exists a positive integer $t$ such that $g(x) > 1$ for each $x > t$. Set $a = ntq_1 \ldots q_s$. Since $a > t$, $g(a) > 1$. Let $q$ be a prime divisor of $g(a)$. The numbers $a$, $n$, and $q$ satisfy the conditions of the Lemma, so $n \mid (a^n - 1)(q - 1)$. Since $n \mid a$, we have $\gcd(a^n - 1, n) = 1$. Therefore $n \mid q - 1$; that is, $q = 1 + ny$. Since $q \mid g(a)$ and $\gcd(g(a), a) = 1$, we have $\gcd(q, a) = 1$, and it follows that $\gcd(q, q_1 \ldots q_s) = 1$. Thus $q$ is not one of the numbers $q_1, \ldots, q_s$, which contradicts the assumption that $q_1, \ldots, q_s$ are the only primes of the form $1 + ny$. This completes the proof. ■

**Acknowledgment.** The author wants to thank the referee for valuable suggestions.

## REFERENCES

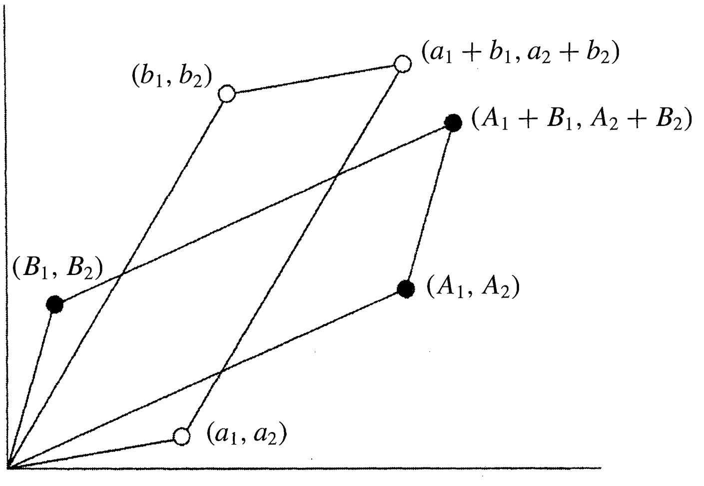
1. T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, New York, 1976.
2. L. E. Dickson, *History of the Theory of Numbers*, Chelsea Publishing Company, New York, 1952.
3. L. Levine, Fermat's Little Theorem: a proof by function iteration, this MAGAZINE **72** (1999), 308–309.
4. P. Ribenboim, *The New Book of Prime Number Records*, Springer-Verlag, New York, 1996.
5. A. Selberg, An elementary proof of Dirichlet's theorem about primes in an arithmetic progression, *Ann. of Math.* **50** (1949), 297–304.


# Proof Without Words: Simpson's Paradox


JERZY KOCIK
Southern Illinois University
Carbondale, IL 62901


Popularity of a candidate is greater among women than men in each town, yet popularity of the candidate in the whole district is greater among men.

Procedure $A$ has greater succes than procedure $B$ in each hospital, yet, in general, procedure B has greater success than $A$.




$$\frac{a_2}{a_1} < \frac{A_2}{A_1} \quad \text{and} \quad \frac{b_2}{b_1} < \frac{B_2}{B_1}, \quad \text{yet} \quad \frac{a_2 + b_2}{a_1 + b_1} > \frac{A_2 + B_2}{A_1 + B_1}$$

For more about Simpson's paradox, see


1. Thomas R. Knapp, Instances of Simpson's paradox, *College Math. J.*, **16**:3, 209–211.
2. A. Tan, A geometric interpretation of Simpson's paradox, *College Math. J.*, **17**:4, 340–341.

# A Nowhere Differentiable Continuous Function Constructed Using Cantor Series

LIU WEN
Hebei University of Technology
Tianjin 300130, China

The examples of continuous nowhere differentiable functions given in most analysis texts involve the uniform convergence of a series of functions (see Hobson [**1**, pp. 401–412]). In the last twenty years interest in this subject has been renewed ([**2**]–[**7**]). In this note we construct a new elementary example by using the Cantor series, which is very accessible and needs only the basic notion of limit. Let $q_n \geq 2$ be a sequence of positive integers. A *Cantor series*, or *Cantor expansion*, for a real number $x \in [0, 1]$ is analogous to a decimal expansion, where numbers other than powers of ten can serve as denominators:

$$x = \sum_{n=1}^{\infty} \frac{x_n}{q_1 \dots q_n}. \tag{1}$$

Here the $n^{\text{th}}$ digit $x_n$ can take on the values $0, 1, \dots, q_n - 1$. It is known that every $x \in [0, 1]$ has a Cantor expansion, although the expansion may not be unique. Our function is defined by

$$u = f(x) = \sum_{n=1}^{\infty} \frac{u_n}{n(n+1)}, \tag{2}$$

where the numbers $u_n$ are defined as follows: $u_1 = 1$, and when $n \geq 1$,

$$u_{n+1} = \begin{cases} -\dfrac{u_n}{n}, & \text{if } x_{n+1} = 0 \text{ but } x_n \neq 0, \\ & \text{or if } x_{n+1} = q_{n+1} - 1 \text{ but } x_n \neq q_n - 1; \\ u_n, & \text{otherwise}. \end{cases} \tag{3}$$

The condition in the first line identifies situations where either adding or subtracting in the next digit past the $n^{\text{th}}$ would affect the $n^{\text{th}}$ digit. We must check to see that $f$ is well defined by these formulas in the case when there are two distinct Cantor expansions for $x$. Suppose that $x \in (0, 1]$ has two Cantor expansions; they must have the following form (Thinking of a number like 0.47999... might be helpful.):

$$x = \sum_{k=1}^{n} \frac{x_k}{q_1 \dots q_k}, \; x_n > 0, \qquad \text{and}$$

$$x = \sum_{k=1}^{n-1} \frac{x_k}{q_1 \dots q_k} + \frac{x_n - 1}{q_1 \dots q_n} + \sum_{k=n+1}^{\infty} \frac{q_k - 1}{q_1 \dots q_k}.$$

Since $u_k$ depends only on the first $k$ digits of the Cantor series of $x$, the corresponding values of $f(x)$ are

$$u = \sum_{k=1}^{n} \frac{u_k}{k(k+1)} - \frac{u_n}{n} \sum_{k=n+1}^{\infty} \frac{1}{k(k+1)} = \sum_{k=1}^{n-1} \frac{u_k}{k(k+1)}, \qquad \text{and}$$

$$u' = \sum_{k=1}^{n-1} \frac{u_k}{k(k+1)} + \frac{u'_n}{n(n+1)} - \frac{u'_n}{n} \sum_{k=n+1}^{\infty} \frac{1}{k(k+1)} = \sum_{k=1}^{n-1} \frac{u_k}{k(k+1)},$$

so the two values are equal. (Note that the series $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ is a telescoping series that sums to $1/(n+1)$.) Hence $f(x)$ is well defined. For convenience, we agree that if x has two Cantor expansions, only the one with infinitely many 0s is used.

We first prove that $f(x)$, as defined by (2) and (3) is right-continuous. Given $\varepsilon > 0$, there exists a positive integer $n$ such that $2/(n+1) < \varepsilon$.

We define a number just slightly larger than $x$:

$$x^* = \sum_{k=1}^{n} \frac{x_k}{q_1 \ldots q_k} + \sum_{k=n+1}^{\infty} \frac{q_k - 1}{q_1 \ldots q_k}.$$

Then for any $x' \in (x, x^*)$ we have

$$x' = \sum_{k=1}^{n} \frac{x_k}{q_1 \ldots q_k} + \sum_{k=n+1}^{\infty} \frac{x'_k}{q_1 \ldots q_k},$$

$$f(x') = \sum_{k=1}^{n} \frac{u_k}{k(k+1)} + \sum_{k=n+1}^{\infty} \frac{u'_k}{k(k+1)}, \quad \text{and}$$

$$\left|f(x') - f(x)\right| \le \sum_{k=n+1}^{\infty} \frac{\left|u'_k - u_k\right|}{k(k+1)} \le \sum_{k=n+1}^{\infty} \frac{2}{k(k+1)} = \frac{2}{n+1} < \varepsilon.$$

Hence $f$ is right-continuous at $x$. Now we prove that $f$ does not have a finite right-derivative at any $x \in [0, 1)$, provided $q_n \ge 3$ $(n \ge 1)$ and

$$\lim_{n\to\infty} \frac{q_1 \ldots q_n}{n!} = \infty. \tag{4}$$

By our agreement about duplicate Cantor expansions, there exists a positive integer $n$, which may be arbitrarily large, such that $x_n < q_n - 1$. Let

$$a_n = \sum_{k=1}^{n-1} \frac{x_k}{q_1 \ldots q_k} + \frac{x_n + 1}{q_1 \ldots q_n};$$

$$b_n = \sum_{k=1}^{n-1} \frac{x_k}{q_1 \ldots q_k} + \frac{x_n + 1}{q_1 \ldots q_n} + \sum_{k=n+1}^{\infty} \frac{q_k - 2}{q_1 \ldots q_k}.$$

Then $x < a_n < b_n$. Noting that $\sum_{k=n+1}^{\infty} (q_k - 1)/(q_1 \ldots q_k) = 1/(q_1 \ldots q_n)$, we derive several useful inequalities:

$$b_n - a_n < \frac{1}{q_1 \ldots q_n} \tag{5}$$

$$b_n - x \le \frac{1}{q_1 \ldots q_n} + \sum_{k=n+1}^{\infty} \frac{q_k - 2}{q_1 \ldots q_k} < \frac{2}{q_1 \ldots q_n}$$

$$b_n - a_n = \sum_{k=n+1}^{\infty} \frac{q_k - 2}{q_1 \ldots q_k} = \frac{1}{2} \sum_{k=n+1}^{\infty} \frac{2q_k - 4}{q_1 \ldots q_k} \ge \frac{1}{2} \sum_{k=n+1}^{\infty} \frac{q_k - 1}{q_1 \ldots q_k} = \frac{1}{2q_1 \ldots q_n}$$

Hence

$$b_n - a_n > (b_n - x)/4 > (a_n - x)/4. \tag{6}$$

From the definition of $f$, we compute the following values:

$$f(a_n) = \sum_{k=1}^{n} \frac{u_k}{k(k+1)} - \frac{\overline{u}_n}{n} \sum_{k=n+1}^{\infty} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \frac{u_k}{k(k+1)} - \frac{\overline{u}_n}{n(n+1)}$$

$$f(b_n) = \sum_{k=1}^{n} \frac{u_k}{k(k+1)} + \sum_{k=n+1}^{\infty} \frac{\overline{u}_n}{k(k+1)} = \sum_{k=1}^{n} \frac{u_k}{k(k+1)} + \frac{\overline{u}_n}{(n+1)}$$

Now $|\overline{u}_n| \geq 1/(n-1)!$, so computation shows

$$|f(b_n) - f(a_n)| = \frac{|\overline{u}_n|}{n} \geq \frac{1}{n!}. \tag{7}$$

Combining (4), (5), and (7), we have

$$\frac{|f(b_n) - f(a_n)|}{b_n - a_n} \geq \frac{q_1 \cdots q_n}{n!} \to \infty \quad \text{as} \quad n \to \infty. \tag{8}$$

Using the triangle inequality and (6), we find

$$\begin{aligned} \frac{|f(b_n) - f(a_n)|}{b_n - a_n} &\leq \frac{|f(b_n) - f(x)|}{b_n - a_n} + \frac{|f(a_n) - f(x)|}{b_n - a_n} \\ &\leq \frac{4\,|f(b_n) - f(x)|}{b_n - x} + \frac{4\,|f(a_n) - f(x)|}{a_n - x}. \end{aligned} \tag{9}$$

Since (8) shows that the left-hand side grows without bound as $n \to \infty$, (9) says that $f$ does not have a finite right-derivative at $x$. To obtain the analogous result on the left, we agree that if $x$ has two Cantor expansions, only the one with $x_k = q_k - 1$ for all $k \geq n$ is used. Arguments similar to those given can then be used to show that f is left-continuous and does not have a finite left-derivative at any $x \in (0, 1]$. Thus $f$ is continuous and nowhere differentiable.

## REFERENCES


1. E. W. Hobson, *The Theory of Functions of a Real Variable*, Vol. 2, Dover, New York, 1957.
2. P. Billingsley, Van der Warden's continuous nowhere differentiable function, *Amer. Math. Monthly*, **89** (1982), 691.
3. F. S. Cater, On Van der Warden's nowhere differentiable function, *Amer. Math. Monthly*, **91** (1984), 307–308.
4. M. Lynch, A continuous, nowhere differentiable function, *Amer. Math. Monthly*, **99** (1992), 8–9.
5. D. P. Minassian and J. W. Gaisser, A simple Weierstrass function, *Amer. Math. Monthly*, **91** (1984), 254–256.
6. H. Sagan, An elementary proof that Schoenberg's space-filling curve is nowhere differentiable, this MAGAZINE, **65**:2 (1992), 125–128.
7. A. Shidear, On the continuity of Van der Warden's function in the Holder sense, *Amer. Math. Monthly*, **93** (1986), 375–376.

# PROBLEMS

ELGIN H. JOHNSTON, *Editor*
Iowa State University

*Assistant Editors:* RĂZVAN GELCA, Texas Tech University; ROBERT GREGORAC, Iowa State University; GERALD HEUER, Concordia College; PAUL ZEITZ, The University of San Francisco

## Proposals

*To be considered for publication, solutions should be received by May 1, 2002.*

**1633.** *Proposed by Emeric Deutsch, Polytechnic University, Brooklyn, NY.*

A palindromic composition of a positive integer $n$ is a palindromic finite sequence of positive integers whose sum is $n$. As examples, 1, 2, 2, 1 and 2, 1, 1, 2 are different palindromic compositions of 6, and 10, 3, 10 is a palindromic composition of 23. Find the number of palindromic compositions of the positive integer $n$.

**1634.** *Proposed by Constantin P. Niculescu, University of Craiova, Craiova, Romania.*

Find the smallest constant $k > 0$ such that

$$\frac{ab}{a+b+2c}+\frac{bc}{b+c+2a}+\frac{ca}{c+a+2b}\le k(a+b+c)$$

for every $a, b, c > 0$.

**1635.** *Proposed by Larry Hoehn, Austin Peay State University, Clarksville, TN.*

Prove or disprove: If two cevians of a triangle are congruent and divide their respective sides in the same proportion, then the triangle is isosceles.

**1636.** *Proposed by Leroy Quet, Denver, CO.*

For any positive integer $m$, define $R(m)=\prod_{k=1}^{m}k^{2k-1-m}$.

(a) Prove that $R(m)$ is an integer that is divisible by every prime less than or equal to $m$ if and only if either $m+1$ is prime or $\frac{m+1}{p^k}>p$, where $p$ is the largest prime dividing $m+1$ and $p^k$ is the largest power of $p$ that divides $m+1$.

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.

Solutions and new proposals should be mailed to Elgin Johnston, Problems Editor, Department of Mathematics, Iowa State University, Ames IA 50011, or mailed electronically (ideally as a LaTeX file) to `ehjohnst@iastate.edu`. All communications should include the readers name, full address, and an e-mail address and/or FAX number.

(b) Prove that if $R(m)$ is not divisible by every prime less than or equal to $m$, then there is exactly one prime less than or equal to $m$ that does not divide $R(m)$.

**1637.** *Proposed by Erwin Just (Emeritus), Bronx Community College, Bronx, NY.*

Prove that the circle with equation $x^2 + y^2 = 1$ contains an infinite number of points with rational coordinates such that the distance between each pair of the points is irrational.

## Quickies

*Answers to the Quickies are on page 409.*

**Q915.** *Proposed by Emeric Deutsch, Polytechnic University, Brooklyn, NY.*

Let $S_n$, $n \geq 2$, denote the set of all permutations of $\{1, 2, \ldots, n\}$. Obviously $1 \leq \max_{1 \leq i \leq n} |\sigma_i - i| \leq n - 1$ for each permutation $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ in $S_n$.

(a) For how many permutations $\sigma$ in $S_n$ do we have $\max_{1 \leq i \leq n} |\sigma_i - i| = n - 1$?

(b) For how many permutations $\sigma$ in $S_n$ do we have $\max_{1 \leq i \leq n} |\sigma_i - i| = 1$?

**Q916.** *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

Do there exist any integers $k$ such that there are an infinite number of relatively prime positive integer triples $(x, y, z)$ satisfying the Diophantine equation

$$x^2y^2 = k^2(x + y + z)(y + z - x)(z + x - y)(x + y - z)?$$

## Solutions

**Onederful, Onederful!** **December 2000**

**1608.** *Proposed by William D. Weakley, Indiana-Purdue University at Fort Wayne, Fort Wayne, IN.*

Let $b$ be a positive integer, $b > 1$. We call a positive integer *onederful* in the base $b$ if it divides some integer whose base $b$ representation is all ones. Which positive integers are onederful in the base $b$?

*Solution by William P. Wardlaw, U.S. Naval Academy, Annapolis, MD.*

A positive integer $n$ is onederful in the base $b$ if and only if $n$ is relatively prime to $b$. First observe that if $n$ divides $(111\ldots11)_b = \sum_{k=0}^{r} b^k$ for some $r$, then clearly $n$ is relatively prime to $b$. Conversely, if $n$ is relatively prime to $b$, then $m = (b-1)n$ is also relatively prime to $b$. Hence, by the Euler-Fermat theorem, $m$ divides $b^{\phi(m)} - 1$. It follows that $n$ divides

$$\frac{b^{\phi(m)} - 1}{b - 1} = b^{\phi(m)-1} + \cdots + b^2 + b + 1 = (11\ldots111)_b.$$

*Also solved by Roy Barbara (Lebanon), Jany C. Binz (Switzerland), Michel Bataille (France), Jean Bogaert (Belgium), Molly Brazill and Michele Renehan, Marc A. Brodie, Con Amore Problem Group (Denmark), Jim Delany, Daniele Donini (Italy), Robert L. Doucette, Marty Getz and Dixon Jones, Jerrold W. Grossman, Elmer K. Hayashi, Tom Jager, Kelly Jahns, Victor Y. Kutsenok, Stephen Maguire, Reiner Martin, Tyrel McQueen, José H. Nieto (Venezuela), Bill Stone, John S. Sumner and Kevin L. Dove, Ajaj A. Tarabay and Bassem B. Ghalayini (Lebanon), Michael Woltermann, and the proposer. There was one incorrect submission.*

**A Triangle Inequality** **December 2000**

**1609.** *Proposed by Yanir A. Rubenstein, student, Technion–Israel Institute of Technology, Haifa, Israel.*

Evaluate

$$\inf_{\substack{a,b\in\mathbb{C}\\ \mathrm{Im}(a\bar{b})\neq 0}} \frac{(|a|+|b|)(|a|+|b|+|a+b|)}{\left|\mathrm{Im}\,(a\bar{b})\right|}.$$

*Solution by John S. Sumner and Kevin L. Dove, University of Tampa, Tampa, FL.*

Let $F$ denote the expression whose infimum we wish to find, and write $a = |a|e^{i\alpha}$, $b = |b|e^{i\beta}$, where $\alpha - \beta \neq \pi k$ ($k$ an integer). Then $|\mathrm{Im}(a\bar{b})| = |a||b||\sin(\alpha - \beta)|$ is twice the area of the triangle $T$ with sides $|a|$, $|b|$, and $|a+b|$. We then have

$$F = \frac{p(p-|c|)}{2\,\mathrm{Area}(T)},$$

where $p$ is the perimeter of triangle $T$. Because the value of $F$ is unchanged when $a$ and $b$ are multiplied by the same constant, we may assume that $p$ is fixed. For fixed $p$ and given $|c| < p$, the area of triangle $T$ is maximized when $|a| = |b|$. Setting $|a| = |b|$ and using Heron's formula for the area of a triangle, we have

$$F = \frac{p(p-|c|)}{2\sqrt{s(s-|a|)(s-|a|)(s-|c|)}} = \frac{2p(p-|c|)}{\sqrt{pc^2(p-2|c|)}} = \frac{2\left(1-\frac{|c|}{p}\right)}{\sqrt{\left(\frac{|c|}{p}\right)^2\left(1-2\frac{|c|}{p}\right)}},$$

where $s = p/2$ is the semiperimeter of the triangle. Because $T$ is a nondegenerate triangle, $0 < |c|/p < 1/2$. It is easy to verify that $F$ is minimized for $\frac{|c|}{p} = \frac{3-\sqrt{5}}{2}$ and that the corresponding value of $F$ is $\sqrt{22+10\sqrt{5}}$.

*Also solved by Roy Barbara (Lebanon), Michel Bataille (France), Jean Bogaert (Belgium), Con Amore Problem Group (Denmark), Daniele Donini (Italy), Robert L. Doucette, Tom Jager, Ajaj A. Tarabay and Bassem B. Ghalayini (Lebanon), Michael Woltermann, Li Zhou, and the proposer. There were two incorrect submissions.*

**Balanced Segments on a Necklace** **December 2000**

**1610.** *Proposed by Hassan A. Shah Ali, Tehran, Iran.*

Place $n$ black pieces and $n$ white pieces on distinct points on the circumference of a circle.

(a) Prove that for each natural number $k \leq n$, there exists a chain of $2k$ consecutive pieces on the circle of which exactly $k$ are black.

(b) Prove that there are at least two such chains that are disjoint if

$$k \leq \sqrt{2n+2} - 2.$$

*Solution by Roy Barbara, Lebanese University, Fanar, Lebanon.*

For convenience a chain of $2k$ consecutive pieces on the circle is called a $2k$-chain. A $2k$-chain is called balanced if it consists of $k$ black pieces and $k$ white pieces. If $C$ is a $2k$-chain, define

$$f(C) = \tfrac{1}{2}(\text{number of black pieces in } C - \text{number of white pieces in } C).$$

Note that if $C$ is shifted either direction by one piece, then $f(C)$ changes by $-1$ or $0$ or $1$. Thus if $C_a$ and $C_b$ are $2k$-chains with $f(C_a) \geq 0$ and $f(C_b) \leq 0$, then there is a (balanced) $2k$-chain $C$ with $f(C) = 0$. In particular, such a chain is encountered by shifting $C_a$ one piece at a time until it coincides with $C_b$.

(a) Let $C_1, C_2, \ldots, C_{2n}$ be all of the distinct $2k$-chains on the circle. Then

$$\sum_{k=1}^{2n} f(C_i) = k(\text{number of black piece} - \text{number of white pieces}) = 0.$$

Thus there exist $2k$-chains $C_a$ and $C_b$ with $f(C_a) \geq 0$ and $f(C_b) \leq 0$. As shown above, it follows that there is a balanced $2k$-chain.

(b) Let $n = kq + r$, with $q \geq 2$ and $0 \leq r < k$. We show that there exist two disjoint balanced $2k$-chains if either

$$\text{(i) } k < 2q - 1 \qquad \text{or} \qquad \text{(ii) } r < q - 1.$$

We show later that condition (i) is satisfied if $k \leq \sqrt{2n+2} - 2$. Condition (ii) allows large values of $k$, including all proper divisors of $n$.

Let $C$ be a balanced $2k$-chain, and denote by $A$ the complementary $(2n - 2k)$-chain. Because $2n - 2k = (q-1)2k + 2r$, we can divide $A$ into $q$ disjoint pieces $C_0, C_1, \ldots, C_{q-1}$, where $C_1, C_2, \ldots, C_{q-1}$ are $2k$-chains and $C_0$ is a $2r$-chain. Extending the definition of $f$ in a natural way to $2r$-chains, we have

$$\sum_{i=0}^{q-1} f(C_i) = 0. \qquad (*)$$

Assume that there are no balanced $2k$-chains in $A$. We may then assume, without loss of generality, that $f(C') \geq 1$ for any $2k$-chain $C'$ contained in $A$. (If $A$ contains $2k$-chains $C_a$ and $C_b$ with $f(C_a) \geq 1$ and $f(C_b) \leq 0$, then we can argue as in part (a) to find a balanced $2k$-chain in $A$.)

Case 1: $r < q - 1$. Because $f(C_0) \geq -r$, we have $f(C_0) + f(C_1) + \cdots + f(C_{q-1}) \geq -r + (q-1) > 0$. This contradicts $(*)$. Hence, if $r < q - 1$, then there is a balanced $2k$-chain in $A$. This chain is disjoint from $C$.

Case 2: $k < 2q - 1$. By $(*)$,

$$f(C_0) = -\sum_{i=1}^{q-1} f(C_i) \leq -(q-1).$$

This means that $C_0$ contains at least $2(q-1)$ white pieces, and hence at least $k$ white pieces. Let $D_0$ be the $2k$-chain formed by appending to $C_0$ the next adjacent chain of $2k - 2r$ pieces from the adjacent chain $C_1$. Because $D_0$ contains at least $k$ white pieces, $f(D_0) \leq 0$. This contradicts the assumption that $f(C') \geq 1$ for all $2k$-chains $C'$ in $A$. Thus, if $k \leq 2q - 2$, then there is a balanced $2k$-chain disjoint from $C$.

Finally, note that (i) is true if $k \leq \sqrt{2n+2} - 2$. Indeed if this is the case, then $k < \sqrt{2n+6} - 2$. This in turn implies that $k^2 + 4k - 2(n+1) < 0$, from which

$$k \leq \frac{2n + 2 - 4k}{k} \leq \frac{2n - 2k - 2r}{k} = 2q - 2 < 2q - 1.$$

*Also solved by Con Amore Problem Group (Denmark), Robert L. Doucette, Georgi D. Gospodinov and Jeff Lutgen, Jerrold W. Grossman, Victor Y. Kutsenok, Reiner Martin, José H. Nieto (Venezuela), Edward Schmeichel, Harry Sedinger, David Smyth, John S. Sumner and Kevin L. Dove, Ajaj A. Tarabay and Bassem B. Ghalayini (Lebanon), Li Zhou, and the proposer. There was one incorrect submission.*

**A Centroid Condition** **December 2000**

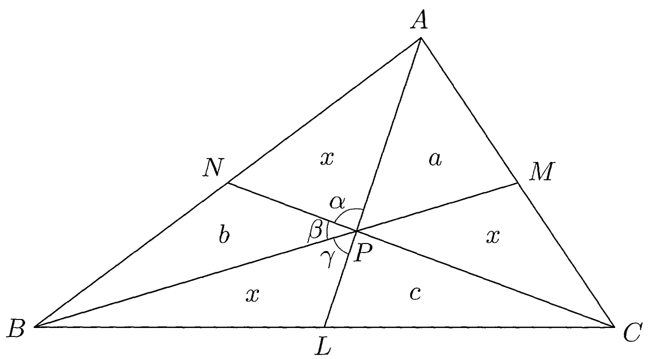
**1611.** *Proposed by Ho-joo Lee, student, Kwangwoon University, Seoul, South Korea.*

Let $P$ be in the interior of $\triangle ABC$, and let lines $AP$, $BP$, $CP$ intersect the sides $BC$, $CA$, $AB$ in $L$, $M$, $N$, respectively. Prove that $P$ is the centroid of $\triangle ABC$ if

$$[APN] = [BPL] = [CPM],$$

where $[\,\cdot\,]$ denotes area.

*Solution by Ajaj A. Tarabay and Bassam B. Ghalayini, Notre Dame University, Zouk Mikael, Lebanon.*




Referring to the figure, let $x = [APN]$, $a = [MPA]$, $b = [NPB]$, and $c = [LPC]$. Then

$$\begin{aligned} x^3 &= \left(\frac{1}{2}PA \cdot PN \sin\alpha\right)\left(\frac{1}{2}PB \cdot PL \sin\gamma\right)\left(\frac{1}{2}PC \cdot PM \sin\beta\right) \\ &= \left(\frac{1}{2}PM \cdot PA \sin\gamma\right)\left(\frac{1}{2}PN \cdot PB \sin\beta\right)\left(\frac{1}{2}PL \cdot PC \sin\alpha\right) \\ &= abc. \end{aligned}$$

Without loss of generality, we may assume that $a \geq b \geq c$, so $a \geq x \geq c$. Now

$$\frac{BL}{LC} = \frac{[BPL]}{[LPC]} = \frac{[ABL]}{[ACL]},$$

and it follows that

$$\frac{x}{c} = \frac{2x+b}{a+x+c} = \frac{2x+b-x}{a+x+c-c} = \frac{x+b}{x+a}.$$

Because $\frac{x}{c} \geq 1$ and $\frac{x+b}{x+a} \leq 1$, we have $\frac{x}{c} = \frac{x+b}{x+a} = 1$. Therefore $x = c$ and $a = b$. Because $abc = x^3$, it follows that $a = b = c = x$. Hence $BL = LC$ and $CM = MA$, so $P$ is the centroid of triangle $ABC$.

*Also solved by Herb Bailey, Roy Barbara (Lebanon), Michel Bataille (France), Jany C. Binz (Switzerland), Minh Can, Daniele Donini (Italy), Robert L. Doucette, Petar Drianov (Canada), Ovidiu Furdui, Michael Golomb, Geoffrey A. Kandall, Victor Y. Kutsenok, Laurel and Hardy Problem Group, Heinz-Jürgen Seiffert (Germany), Raul A. Simon (Chile), Achilleas Sinefakopoulos (Greece), John S. Sumner and Kevin L. Dove, Michael Woltermann, Li Zhou, and the proposer.*

**Not a Centroid Condition** **December 2000**

**1612.** *Proposed by Ho-joo Lee, student, Kwangwoon University, Seoul, South Korea.*

Let $P$ be in the interior of $\triangle ABC$, and let lines $AP$, $BP$, $CP$ intersect the sides $BC$, $CA$, $AB$ in $L$, $M$, $N$, respectively. Prove that $P$ is the centroid of $\triangle ABC$ if

$$[APN]+[BPL]+[CPM]=[APM]+[BPN]+[CPL],$$

where $[\,\cdot\,]$ denotes area.

*Solution by Robert L. Doucette, McNeese State University, Lake Charles, LA.*

The given statement is not true. For a counterexample, choose $\triangle ABC$ with $AB = AC$. Let $P$ be any point other than the centroid on the median from $A$. Note that $[APN] = [APM]$, $[BPL] = [CPL]$, and $[CPM] = [BPN]$. We then have $[APN]+[BPL]+[CPM]=[APM]+[BPN]+[CPL]$.

One conclusion that is valid from the given hypotheses is that at least one of the segments $AL$, $BM$, $CN$ is a median of $\triangle ABC$. To prove this, let $x = [APN]$, $y = [BPL]$, $z = [CPM]$, $a = [APM]$, $b = [BPN]$, and $c = [CPL]$. Let $s$ be the common value of $x + y + z$ and $a + b + c$. Because $\triangle ACN$ and $\triangle BCN$ share an altitude from $C$,

$$\frac{AN}{BN}=\frac{x+z+a}{b+c+y}=\frac{s-(y-a)}{s+(y-a)}. \qquad (*)$$

Similarly,

$$\frac{BL}{CL}=\frac{x+y+b}{a+c+z}=\frac{s-(z-b)}{s+(z-b)} \quad \text{and} \quad \frac{CM}{AM}=\frac{y+z+c}{a+b+x}=\frac{s-(x-c)}{s+(x-c)}.$$

By Ceva's Theorem,

$$\frac{AN}{BN}\cdot\frac{BL}{CL}\cdot\frac{CM}{AM}=1.$$

Letting $d_1 = x - c$, $d_2 = y - a$, and $d_3 = z - b$, it follows that

$$(s-d_1)(s-d_2)(s-d_3)=(s+d_1)(s+d_2)(s+d_3).$$

Expanding, and keeping in mind that $d_1 + d_2 + d_3 = 0$, we find $d_1d_2d_3 = 0$. Thus at least one of the $d_i$'s is 0. If $d_2 = 0$, then it follows from $(*)$ that $AN = BN$ so $CN$ is a median. Similar conclusions follow if $d_1 = 0$ or $d_3 = 0$.

*Note.* The converse of this result holds. Indeed, if $AL$ is a median and $P$ is a point on $AL$, then use an affine transformation to map $\triangle ABC$ to $\triangle A'B'C'$ with $A'B' = A'C'$. Because medians and ratios of areas are preserved, this reduces the situation to the counterexample described earlier.

*Also solved by Herb Bailey, Roy Barbara (Lebanon), Michel Bataille (France), Daniele Donini (Italy), Petar Drianov (Canada), Michael Golomb, Geoffrey A. Kandall, Victor Y. Kutsenok, Heinz-Jürgen Seiffert (Germany), Achilleas Sinefakopoulos (Greece), John S. Sumner and Kevin L. Dove, Michael Woltermann, and Li Zhou. There was one solution with no name and four incorrect submissions.*

# Answers

*Solutions to the Quickies from page 404.*

**A915.**

(a) In this case we must have either $\sigma_1 = n$ or $\sigma_n = 1$, or both. By the Inclusion/Exclusion Principle there are $2(n-1)! - (n-2)! = (2n-3)(n-2)!$ such permutations.

(b) Let $\alpha_n$ denote the number of permutations in $S_n$ satisfying the maximality condition. If $\max_{1\le i\le n}|\sigma_i - i| = 1$, then either $\sigma_n = n$ or $\sigma_n = n-1$. The number of such permutations with $\sigma_n = n$ is $\alpha_{n-1}$. On the other hand, if $\sigma_n = n-1$, then $\sigma_{n-1} = n$, and $\max_{1\le i\le n-2}|\sigma_i - i| = 0$ or $1$; there are $\alpha_{n-2}+1$ such permutations. Thus

$$\alpha_n = \alpha_{n-1} + \alpha_{n-2} + 1 \qquad (*)$$

with $\alpha_1 = 0$ and $\alpha_2 = 1$. Adding 1 to both sides of $(*)$ we find that $\alpha_n = F_{n+1} - 1$, where $F_k$ denotes the $k$th Fibonacci number.

**A916.** In order for the right-hand side of the equation to be positive, $x$, $y$, $z$ must be the lengths of the sides of a triangle. If $A$ is the area of the triangle and $\alpha$ is the angle between the sides of lengths $x$ and $y$, then

$$\frac{4A^2}{\sin^2\alpha} = 16k^2A^2,$$

so $\sin\alpha = \frac{1}{2k}$. By the Law of Cosines,

$$z^2 = x^2 + y^2 - 2xy\cos\alpha = x^2 + y^2 \pm \frac{xy}{k}\sqrt{4k^2-1}.$$

Because the left-hand side is an integer and the right-hand side is irrational, there are no solutions.

50 Years Ago in the MAGAZINE (Vol. **25**, No. 2, Nov.–Dec., 1951):

Editor's Note

The editors of the Mathematics Magazine like either to accept submitted articles or give good and sufficient reasons for not accepting them.

However papers on classic problems that have been proved unsolvable, such as trisecting all angles and squaring the circle, have become so numerous that it is not feasible for us to unravel them. Hence we have adopted the policy that authors will be required to show us some error in the classic proofs that such problems are unsolvable before we will weigh their attempts to solve them.

# 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.*

Pickover, Clifford A., *The Zen of Magic Squares, Circles, and Stars*, Princeton University Press, 2002; xx + 399 pp, $29.95. ISBN 0–691–07041–5.

"An ardent magic-square explorer is a 'mathematical samurai,' someone who trains to be neither fearful of defeat nor hopeful of victory and thus enters combat with a neutral attentiveness, indifferent to—but prepared for—the difficult demands of each instant." If you have always wanted to think of yourself as a mathematical samurai, this may be for you. This fascinating book offers a mystical take on the phenomena of magic squares and similar objects; it collects together much lore about magic squares, from the famous book by W.S. Andrews and from other sources. In particular, author Pickover successfully conveys the feeling of transcendence and wonder that magic squares bring him. But the mysticism, which will nevertheless hook some readers into the book's enjoyable mathematical adventures, still seems a little overemphasized, or at best unnecessary. "Many researchers . . . have noted the prevalence of integer patterns in geometry and mysticism. Perhaps there is something special about integers in the fabric of the universe." Perhaps indeed, but the remainder of that long paragraph on p. 371 makes no further mention of mysticism. That paragraph also repeats again, from Pickover's *The Loom of God: Mathematical Tapestries at the Edge of Time* (Plenum, 1997), what is surely the most mangled version ever ("The primary source of all mathematics are [sic] the integers") of Kronecker's famous quote about the integers ("Die ganze Zahl schuf der liebe Gott, alles Uebrige ist Menschenwerk"); but perhaps Pickover's omission of Kronecker's attribution of the integers to God is more suited to this book's theme. [The uncorrected advance proof of this book sent to reviewers did not have an index.]

Cipra, Barry, Number fun with Ben, *Science* 292 (4 May 2001), `http://www.academicpress.com/inscight/04302001/grapha.htm` . Pasles, Paul C., The lost squares of Dr. Franklin, *American Mathematical Monthly* 108 (6) (June-July 2001), 489–511.

As a youth, Benjamin Franklin "amused" himself with making magic squares, though he later had misgivings about not spending the time "more usefully" than on such "*difficiles nugæ* [laborious trifles], incapable of any useful application." Franklin published just three squares, and Paul Pasles (Villanova University) has now republished three more from Franklin's papers plus a previously unknown one—$16 \times 16$—from Franklin's letters. What was Franklin's method? Pasles tries to reconstruct it, since Franklin never described it: "[N]o one has desired me to show him my method of disposing the numbers. It seems they wish rather to investigate it themselves."

MagPortal.com. Magazine articles on mathematics. `http://www.magportal.com/c/sci/math/` .

This Web site adds one or two listings each week of magazine articles about mathematics, mostly from *Science News*, *New Scientist*, *Scientific American*, and *Lingua Franca*, with links to the full text. It's a handy location to find the text of an article that you remember but whose bibliographic details you don't.

Ganter, Susan L., *Changing Calculus: A Report on Evaluation Efforts and National Impact from 1988–1998*, MAA, 2001; xi + 78 pp, $24.95 (P) (discount available to MAA members). ISBN 0–88385–534–8.

After 11 years, 3 million students, and $27 million of NSF money, what has the effort to reform calculus accomplished? The goal was "to help students have a better understanding of and appreciation for mathematics." The 18 objectives, however, were not about outcomes but were just to adopt some features from a specified list of instructional techniques (laboratory experience, discovery learning, cooperative learning), content (applications, real-world modeling, approximation, differential equations), and technology (computer use, graphing calculators). Hence it is perhaps not surprising that only 43 published articles resulted from the 127 projects and only 34 projects reported results on student achievement. In 88% of the studies with such results, there was positive impact on at least one measure of achievement; in 15% there was some negative impact, commonly on computational skills. This report suggests a pattern of positive impact particularly in projects that used technology and group work. Also, some students respond with a positive attitude to reform courses but others do not, and "perhaps calculus reform generates more interest in mathematics [in terms of taking further courses]." However, the author cautions: "[A]ll results ... reported are suspect, since no overall evaluation design was used to collect even the stronger data in a manner that is consistent across [the 127] projects." One wishes that the projects had started with agreement on appropriate course goals and means for evaluating success; then we might by now have results more definitive and satisfying than the vague mild successes summarized here.

Kenschaft, Patricia, Math Medley. One-hour radio show. Saturdays, KFNX 1100 Phoenix 11 A.M., WALE 990 Providence 9 A.M.; possibly one hour later in summer. Schedule of future and past shows at `http://www.csam.montclair.edu/~kenschaft/WALEsched.html` ; indexes by topic and person, with links for listening to past programs, at `http://www.webct.com/math/viewpage?name=math_medley` .

Chemists have long had a nationally syndicated public radio show, Men and Molecules, and National Public Radio has its Science Friday talk show. Unknown to most of us, mathematics too has had its own radio show for several years. Math Medley is a one-hour call-in radio talk show with 10,000 listeners, primarily in Arizona and Rhode Island. Pat Kenschaft (Montclair State University) interviews a different person every Saturday about "education, parenting, equity, and environmental issues, with an underlying theme of mathematics." The past four years of shows have included many mathematicians prominent in the MAA (Pres. Ann Watkins was on in late September about The Joy of Statistics) as well as authorities on elementary and secondary mathematics education and practitioners in various areas involving mathematics (e.g., the mathematics of pesticides, the mathematics of world energy usage). Scheduled for soon after you receive this issue (Dec. 29) is Paul Sally (University of Chicago) on a topic to be announced. I hope you will tune in, either live or asynchronously over the Web; it's time for a national audience for this show.

Stewart, Ian, What could go wrong? *New Scientist* (15 September 2001), 35–39.

Murphy's Law: "Anything that can go wrong, will go wrong." Author Ian Stewart ponders Gödel's undecidability results and wonders why mathematics seems to have escaped Murphy's law unscathed. Any one of the four-color theorem (1976), Fermat's Last Theorem (1994), and the Kepler conjecture (2000) could have turned out to be undecidable propositions but they didn't. "Where are all the undecidable problems? Why are there only a few, lurking in the meta-mathematical fringes?" Then again, there is the enormous question of whether $\mathcal{P} = \mathcal{NP}$. If it is provably undecidable, there would have to be a model of mathematics in which $\mathcal{P} = \mathcal{NP}$ is true, and in that model all kinds of problems would have "quick and easy answers. This would be a mathematician's paradise, and surely Murphy could not allow such a thing to exist." If $\mathcal{P} = \mathcal{NP}$ is undecidable but we cannot prove it, "the question of whether it is undecidable is also undecidable. ... First change the axioms so that $\mathcal{P} = \mathcal{NP}$ is provably undecidable, then change them again so it's true. But if Murphy's master plan is an infinite regression of undecidability .... If it can go meta-wrong, then it will go meta-wrong. Murphy metamurphosed."

# NEWS AND LETTERS

## Acknowledgments

*In addition to our Associate Editors, the following referees have assisted the* MAGAZINE *during the past year. We thank them for their time and care.*

Allman, Elizabeth S., *University of Southern Maine, Falmouth, ME*

Andrews, Peter G., *Eastern Illinois University, Charleston, IL*

Appleby, Glenn D., *Santa Clara University, Santa Clara, CA*

Axler, Sheldon, *San Francisco State University, San Francisco, CA*

Axtell, Michael C., *Wabash College, Crawfordsville, IN*

Barbeau, Jr., Edward J., *University of Toronto, Toronto, Canada*

Barksdale, Jr., James B., *Western Kentucky University, Bowling Green, KY*

Bauer, Craig P., *Northwestern State University, Natchitoches, LA*

Baxley, John V., *Wake Forest University, Winston-Salem, NC*

Beals, Richard W., *Yale University, New Haven, CT*

Beauregard, Raymond A., *University of Rhode Island, Kingston, RI*

Beezer, Robert A., *University of Puget Sound, Tacoma, WA*

Bergum, Gerald E., *South Dakota State University, Brookings, SD*

Berndt, Bruce C., *University of Illinois, Urbana, IL*

Borwein, Jonathan M., *Simon Fraser University, Burnaby, Canada*

Brawner, James N., *Armstrong Atlantic State University, Savannah, GA*

Brown, Peter, *University of New South Wales, Sydney, Australia*

Bullington, Grady D., *University of Wisconsin, Oshkosh, Oshkosh, WI*

Burton, David M., *University of New Hampshire, Durham, NH*

Callan, C. David, *University of Wisconsin, Madison, WI*

Cammack, Larry A., *Central Missouri State University, Warrensburg, MO*

Case, Jeremy S., *Taylor University, Upland, IN*

Chakerian, G. Don, *University of California, Davis, Davis, CA*

Culp-Ressler, Wendell, *Franklin & Marshall College, Lancaster, PA*

DeTemple, Duane W., *Washington State University, Pullman, WA*

Dodge, Clayton W., *University of Maine, Orono, ME*

Drobot, Vladimir, *San Jose State University, San Jose, CA*

Dunne, Edward G., *American Mathematical Society, Providence, RI*

Eisenberg, Bennett, *Lehigh University, Bethelem, PA*

Elderkin, Richard H., *Pomona College, Claremont, CA*

Emert, John W., *Ball State University, Muncie, IN*

Ensley, Douglas E., *Shippensburg University, Shippensburg, PA*

Feroe, John A., *Vassar College, Poughkeepsie, NY*

Firey, William, *Corvallis, OR*

Fisher, Evan, *Lafayette College, Easton, PA*

Fisher, J. Chris, *University of Regina, Regina, Canada*

Fleron, Julian F., *Westfield State College, Westfield, MA*

Flint, Donna L., *South Dakota State University, Brookings, SD*

Foster, Lorraine L. T., *California State University, Northridge, Northridge, CA*

Frantz, Marc, *Indiana University, Bloomington, IN*

Fredricks, Gregory A., *Lewis & Clark College, Portland, OR*

Fung, Maria G., *University of Wisconsin–Stout, Menomonie, WI*

Gallian, Joseph A., *University of Minnesota–Duluth, Duluth, MN*

Ghahramani, Saeed, *Western New England College, Springfield, MA*

Gillman, Leonard, *University of Texas at Austin, Austin, TX*

Gluchoff, Alan Daniel, *Villanova University, Downingtown, PA*

Grabiner, Judith V., *Pitzer College, Claremont, CA*

Green, Euline I., *Abilene Christian University, Abilene, TX*

Grosshans, Frank D., *West Chester University of Pennsylvania, Baltimore, MD*

Guichard, David R., *Whitman College, Walla Walla, WA*

Guy, Richard K., *University of Calgary, Calgary, Canada*

Harborth, Heiko, *Tech University Carolo-Wilhelmina, Braunschweig, Germany*

Hayashi, Elmer K., *Wake Forest University, Winston-Salem, NC*

Hirschhorn, Michael D., *University of New South Wales, Sydney, Australia*

Hoffman, Frederick, *Florida Atlantic University, Boca Raton, FL*

Howard, Fredric T., *Wake Forest University, Winston-Salem, NC*

Isaac, Richard, *Herbert H. Lehman College CUNY, Bronx, NY*

Jafari, Farhad, *University of Wyoming, Laramie, WY*

Janke, Steven J., *Colorado College, Colorado Springs, CO*

Jepsen, Charles H., *Grinnell College, Grinnell, IA*

Johnson, Diane L., *Humboldt State University, Arcata, CA*

Johnson, Warren P., *University of Wisconsin–Madison, Madison, WI*

Jones, Michael A., *Montclair State University, Upper Montclair, NJ*

Jordan, James H., *Washington State University, Pullman, WA*

Kerckhove, Michael G., *University of Richmond, Richmond, VA*

Kiltinen, John O., *Northern Michigan University, Marquette, MI*

Kimberling, Clark, *University of Evansville, Evansville, IN*

Kuzmanovich, James J., *Wake Forest University, Winston-Salem, NC*

Larson, Dean S., *Gonzaga University, Spokane, WA*

Larson, Loren C., *St. Olaf College, Northfield, MN*

Lautzenheiser, Roger G., *Rose-Hulman Institute of Technology, Terre Haute, IN*

Lo Bello, Anthony J., *Allegheny College, Meadville, PA*

Loud, Warren S., *Minneapolis, MN*

Lynch, Mark, *Millsaps College, Jackson, MS*

Mackin, Gail, *Northern Kentucky University, Highland Heights, KY*

Mackiw, George B., *Loyola College in Maryland, Baltimore, MD*

Martin, Greg, *The University of British Columbia, Vancouver, Canada*

McCleary, John H., *Vassar College, Poughkeepsie, NY*

McDermot, Richard F., *Allegheny College, Meadville, PA*

McGuire, Linda E., *Muhlenberg College, Allentown, PA*

Meda, Gowri, *Pacific University, Forest Grove, OR*

Merrill, Kathy D., *Colorado College, Colorado Springs, CO*

Michaels, John G., *State University of New York at Brockport, Brockport, NY*

Mills, Mark A., *Central College, Pella, IA*

Moore, Thomas L., *Grinnell College, Grinnell, IA*

Morics, Steven W., *University of Redlands, Redlands, CA*

Needham, Tristan, *University of San Francisco, San Francisco, CA*

Neidinger, Richard D., *Davidson College, Davidson, NC*

Nelsen, Roger B., *Lewis and Clark College, Portland, OR*

Nicodemi, Olympia E., *State University of New York, Geneseo, NY*

Oakley, Patricia A., *Goshen College, Goshen, IN*

Ostrov, Daniel N., *Santa Clara University, Santa Clara, CA*

Otero, Daniel E., *Xavier University, Cincinnati, OH*

Park, Efton L., *Texas Christian University, Ft. Worth, TX*

Pedersen, Jean J., *Santa Clara University, Santa Clara, CA*

Perline, Ronald K., *Drexel University, Philadelphia, PA*

Pfiefer, Richard E., *San Jose State University, San Jose, CA*

Pinsky, Mark A., *Northwestern University, Evanston, IL*

Ratliff, Thomas C., *Wheaton College, Norton, MA*

Rodger, Chris, *Auburn University, Montgomery, AL*

Ross, Kenneth A., *University of Oregon, Eugene, OR*

Ross, Peter, *Santa Clara University, Santa Clara, CA*

Roth, Richard L., *University of Colorado, Boulder, CO*

Roy, Ranjan, *Beloit College, Beloit, WI*

Sadek, Jawad, *Northwest Missouri State University, Maryville, MO*

Samuels, Stephen M., *Purdue University, West Lafayette, IN*

Sander, Evelyn, *George Mason University, Fairfax, VA*

Saunders, Sam C., *Scientific Consulting Services, Inc, Kirkland, WA*

Saxe, Karen, *Macalester College, St. Paul, MN*

Schumer, Peter D., *Middlebury College, Middlebury, VT*

Scott, Richard A., *Santa Clara University, Santa Clara, CA*

Shahriari, Shahriar, *Pomona College, Claremont, CA*

Shell, Amy E., *United States Military Academy, West Point, NY*

Sherman, Gary, *Rose-Hulman Institute of Technology, Terre Haute, IN*

Shifrin, Theodore, *University of Georgia, Athens, GA*

Simoson, Andrew J., *King College, Bristol, TN*

Sloughter, Daniel C., *Furman University, Greenville, SC*

Smolarski, S. J., Dennis C., *Santa Clara University, Santa Clara, CA*

Snell, J. Laurie, *Dartmouth College, Hanover, NH*

Solheid, Ernie S., *California State University, Fullerton, Fullerton, CA*

Stone, Alexander P., *University of New Mexico, Albuquerque, NM*

Szabò, Tamàs, *Weber State University, Ogden, UT*

Townsend, Douglas W., *Indiana University-Purdue University at Fort Wayne, Ft. Wayne, IN*

Towse, Chris, *Scripps College, Claremont, CA*

Tucker, Thomas W., *Colgate University, Hamilton, NY*

Whitehead, Tamsen, *Santa Clara University, San Jose, CA*

Worner, Tamara S., *Wayne State College, Wayne, NE*

Zeleke, Melkamu, *William Patterson University, Wayne, NJ*

Zullo, Holly S., *Carroll College, Helena, MT*

---

# Index to Volume 74

## Authors

TITLES

## PROBLEMS

*The letters P, Q, and S refer to Proposals, Quickies, and Solutions, respectively; page numbers appear in parentheses. For example, P1622(155) refers to Proposal 1622, which appears on page 155.*

*Thanks are also due to those who have refereed proposals for the Problems Column:*
Clifford Bergman, Alexander Burstein, Irvin Hentzel, Roger Maddux, and James Wilson of Iowa State University, Ames, *and* Vania Mascioni of Western Washington University, Bellingham.

Letters to the Editor appear on pages 368 and 377.

United States Postal Service

### Statement of Ownership, Management, and Circulation

| 1. Publication Title | 2. Publication Number | 3. Filing Date |
|---|---|---|
| Mathematics Magazine | 0 0 2 5 - 5 7 0 x | 9/30/01 |
| **4. Issue Frequency** | **5. Number of Issues Published Annually** | **6. Annual Subscription Price** |
| Bimonthly, except July and August | 5 | $20 |

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*

MAA, 1529 18th St., NW, Washington, DC 20036

Contact Person: H. Waldman

Telephone: 202-387-5200

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

Same

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*

MAA, 1529 18th St., NW, Washington, DC 20036-1385

Editor *(Name and complete mailing address)*

Frank Farris, Santa Clara University, Santa Clara, CA 95053

Managing Editor *(Name and complete mailing address)*

Harry Waldman

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

| Full Name | Complete Mailing Address |
|---|---|
| The Mathematical ASSN of America | 1529 18th St., NW, Washington, DC 20036 |
| | |
| | |
| | |
| | |

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

| Full Name | Complete Mailing Address |
|---|---|
| | |
| | |
| | |
| | |
| | |

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes
☒ Has Not Changed During Preceding 12 Months
☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

PS Form **3526**, October 1999 *(See Instructions on Reverse)*

| 13. Publication Title | | 14. Issue Date for Circulation Data Below | |
|---|---|---|---|
| Mathematics Magazine | | June 2001 | |
| **15. Extent and Nature of Circulation** | | **Average No. Copies Each Issue During Preceding 12 Months** | **No. Copies of Single Issue Published Nearest to Filing Date** |
| a. Total Number of Copies *(Net press run)* | | 14,000 | 15,000 |
| b. Paid and/or Requested Circulation | (1) Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. *(Include advertiser's proof and exchange copies)* | 12,500 | 13,500 |
| | (2) Paid In-County Subscriptions Stated on Form 3541 *(Include advertiser's proof and exchange copies)* | 0 | 0 |
| | (3) Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution | 0 | 0 |
| | (4) Other Classes Mailed Through the USPS | 0 | 0 |
| c. Total Paid and/or Requested Circulation *[Sum of 15b (1), (2), (3), and (4)]* ▶ | | 12,500 | 13,500 |
| d. Free Distribution by Mail *(Samples, complimentary, and other free)* | (1) Outside-County as Stated on Form 3541 | 0 | 0 |
| | (2) In-County as Stated on Form 3541 | 0 | 0 |
| | (3) Other Classes Mailed Through the USPS | 1,500 | 1,500 |
| e. Free Distribution Outside the Mail *(Carriers or other means)* | | 0 | 0 |
| f. Total Free Distribution *(Sum of 15d. and 15e.)* ▶ | | 1,500 | 1,500 |
| g. Total Distribution *(Sum of 15c. and 15f)* ▶ | | 14,000 | 15,000 |
| h. Copies not Distributed | | 0 | 0 |
| i. Total *(Sum of 15g. and h.)* ▶ | | 14,000 | 15,000 |
| j. Percent Paid and/or Requested Circulation *(15c. divided by 15g times 100)* | | 90% | 90% |

16. Publication of Statement of Ownership
☒ Publication required. Will be printed in the December issue of this publication. ☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner: Donald J. Albers, Director of Publications — Date: 10/1/01

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

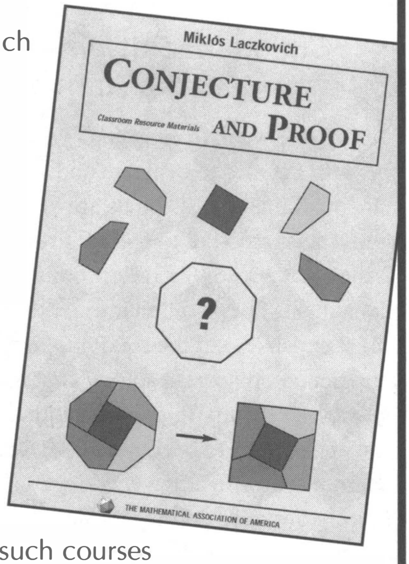
#### Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. Item 15h., Copies not Distributed, must include (1) newsstand copies originally stated on Form 3541, and returned to the publisher, (2) estimated returns from news agents, and (3), copies for office use, leftovers, spoiled, and all other copies not distributed.
5. If the publication had Periodicals authorization as a general or requester publication, this Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
6. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
7. Item 17 must be signed.

***Failure to file or publish a statement of ownership may lead to suspension of Periodicals authorization.***

PS Form **3526**, October 1999 *(Reverse)*

# COMING IN DECEMBER...




***Conjecture and Proof***/Miklós Laczkovich is a compilation of the lecture notes for a course designed and initiated by Paul Erdös, Lásszló Lovász, Vera Sós, and László Babai for a one-semester course given by the Budapest Seminars in Mathematics.

By introducing a variety of advanced topics, the book functions, in part, as a survey of topics from number theory, geometry, measure theory, and set theory. It can be used as a supplement in courses that introduces abstract mathematics to undergraduates. The ideas that are presented are deeper and more sophisticated than those typically encountered in sophomore-level "transition" courses. However, talented students in such courses should find this book to be an exciting excursion into new areas of mathematics—and more importantly, new ways of thinking about mathematical problems. Because of its unusual depth and the fact that some of the sections can stand alone or be combined with a few others to form a unit, this book is ideally suited for upper-level undergraduate seminars or capstone courses.

Although the text discusses questions from various fields including number theory, algebra and geometry, it is centered around the real number system and the problem of measure. Thus, the number theoretic sections are concerned with rational and irrational and with algebraic and transcendental numbers; the problems of geometric constructions clarify the nature of constructible numbers (as a subset of algebraic numbers), and the questions of geometric dissections serve as motivation for general problems of equidecomposability.

**Catalog Code: CAP/MM • 140 pp., Paperbound, 2001 • ISBN 0-88385-722-7**
**List $24.95 • MAA Member $19.95**

**from the**

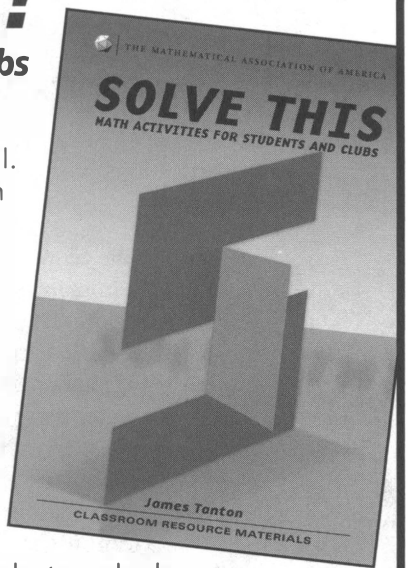
**Mathematical Association of America...**

**Call 800-331-1622 to order!**

# SOLVE THIS!

## *Math Activities for Students & Clubs*

## by James Tanton




Sophisticated mathematics is accessible to all. This book proves it! *Solve This* is a collection of intriguing mathematical problems and activities linked by common themes that involve working with objects from our everyday experience. Learn about the mathematical mysteries of a bagel, a checkerboard, and a pile of laundry, for example. Discover for yourself that wheels need not be round, that braids need not have free ends, that it is always best to turn around twice—and more! Mathematics is all around us. We all do mathematics everyday. This book irresistibly tempts the reader to embark on a journey of investigation and discovery. All the activities are immediate, catchy, and fun, but upon investigation begin to unfold into surprising layers of depth and new perspectives. The necessary mathematics in increasing levels of sophistication is fully explained along the way, but readers may amend the journey in any way to match their mathematical abilities. Elementary, middle and high school students, college students and mathematics majors, faculty from all departments and professional mathematicians, as well as self-described math phobics have all enjoyed these activities and have all attained a sense of satisfaction and accomplishment from them. Mathematics educators will find this an invaluable resource of fresh and innovative approaches to topics in mathematics. *This book is for everyone!*. No preparatory learning is needed to dive right in! You will enjoy doing these activities even more as part of a group event. Whether you are a student or an instructor, anxious or comfortable with math, or just curious about mathematics, you will be challenged and delighted by the activities described herein.

**Catalog Code: HOM/MM • 232 pp., Paperbound, 2001 • ISBN 0-88385-717-0**
**List $23.95 • MAA Member $29.95**

**from the**

**Mathematical Association of America...**

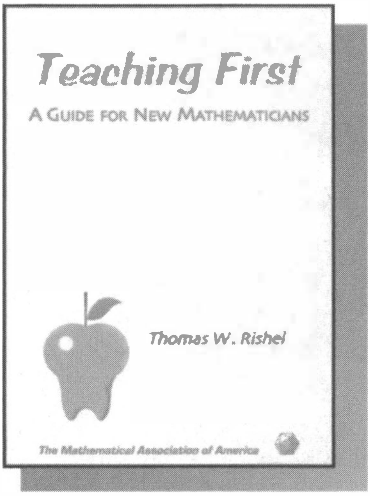
**Call 800-331-1622 to order!**

The Mathematical Association of America

# *Teaching First: A Guide for New Mathematicians*

## Thomas W. Rishel

### Series: MAA Notes




In this volume Thomas Rishel draws on his nearly forty years of teaching experience to address the "nuts and bolts" issues of teaching college mathematics. This book is written for the mathematics TA or young faculty member who may be wondering just where and how to start. Rishel opens the readers' eyes to pitfalls they may never have considered, and offers advice for balancing an obligation "to the student" with an obligation "to mathematics." Throughout, he provides answers to seemingly daunting questions shared by most new TAs, such as how to keep a classroom active and lively; how to prepare writing assignments, tests, and quizzes; how exactly to write a letter of recommendation; and how to pace, minute by minute, the "mathematical talks" one will be called upon to give.

This book is Rishel's answer to those who may suggest that good teaching is innate and cannot be taught. This he emphatically denies, and he insists that solid teaching starts with often overlooked "seeming trivialites" that one needs to master before exploring theories of learning. Along the way he also covers the general issues that teachers of all subjects eventually experience: fairness in grading, professionalism among students and colleagues, identifying and understanding student "types", technology in the classroom. All of the subjects in this book are considered within the context of Rishel's experience as a mathematics teacher. All are illustrated with anecdotes and suggestions specific to the teaching of mathematics.

*Teaching First* is a comprehensive guide for a mathematics TA, from the first semester preparations through the unforseen challenges of accepting a faculty position. Its aim is to prepare the new TA with clear suggestions for rapidly improving their teaching abilities.

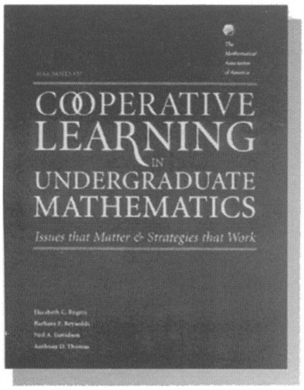
Catalog Code: NTE-54/JR 150 pp., Paperbound, 2000 ISBN 088385-165-2 List: $19.00 MAA Member: $15.00

Name________ Credit Card No.________
Address________ Signature________ Exp. Date___/___
City________ Qty____ Price $____ Amount $____
State____ Zip ____ Shipping and Handling $____
Phone ________ Catalog Code: NTE-54/JR Total $____

Shipping and Handling: USA orders (shipped via UPS): $3.00 for the first book, and $1.00 for each additional book. Canadian orders: $4.50 for the first book and $1.50 for each additional book. Canadian orders will be shipped within 2-3 weeks of receipt of order via the fastest available route. We do not ship via UPS into Canada unless the customer specially requests this service. Canadian customers who request UPS shipment will be billed an additional 7% of their total order. Overseas Orders: $4.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $10.00 per book. Foreign orders must be paid in US dollars through a US bank or through a New York clearinghouse. Credit card orders are accepted for all customers. All orders must be prepaid with the exception of books purchased for resale by bookstores and wholesalers.

Order Via:

Phone: 1 (800) 331.1622
Fax: (301) 206.9789
Mail: Mathematical Association of America
PO Box 91112
Washington, DC 20090-1112
Web: www.maa.org


COOPERATIVE
LEARNING
IN
UNDERGRADUATE
MATHEMATICS
Issues that Matter & Strategies that Work

# CONTENTS

## ARTICLES



## NOTES



## PROBLEMS



## REVIEWS



## NEWS AND LETTERS



**THE MATHEMATICAL ASSOCIATION OF AMERICA**
1529 Eighteenth Street, NW
Washington, DC 20036