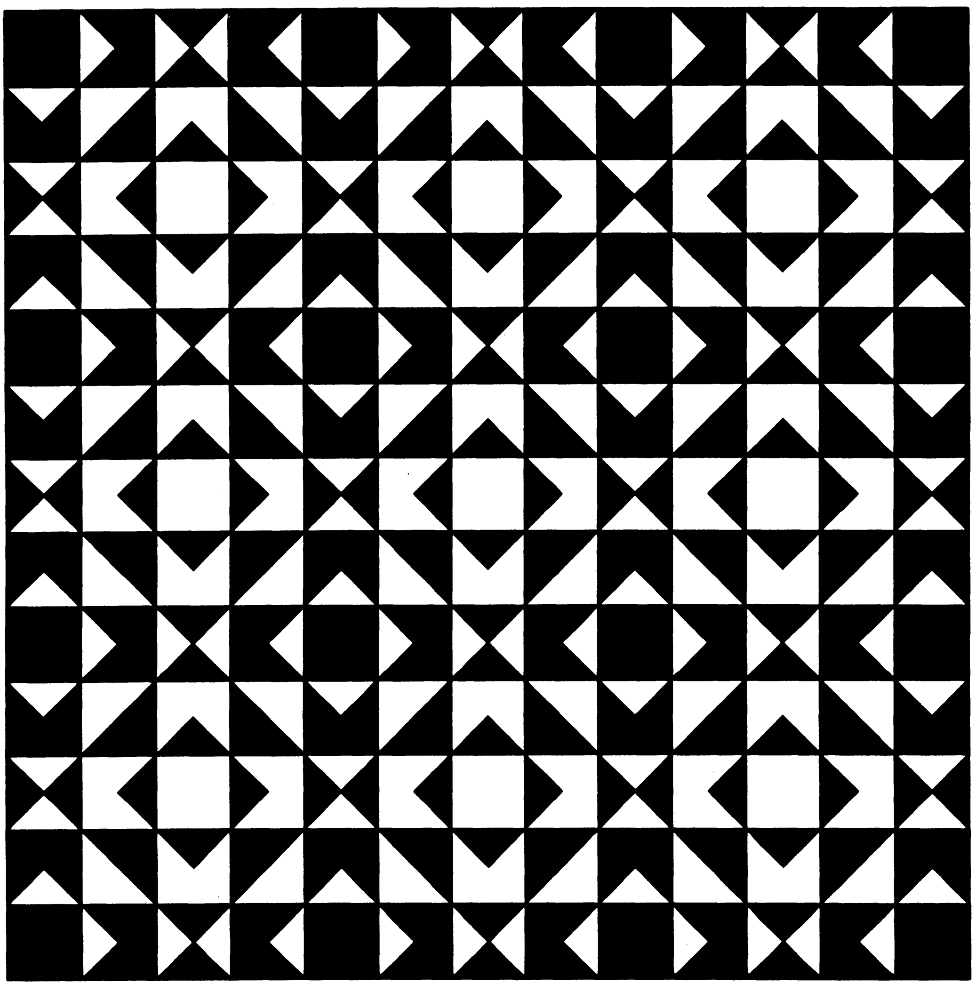
Vol. 72, No. 4, October 1999

MATHEMATICS MAGAZINE



- Farmer Ted Goes Natural
- Integer Antiprisms and Integer Octahedra
- Integral Triangles

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. 71, pp. 76–78, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Send new manuscripts to Paul Zorn, Editor, Department of Mathematics, St. Olaf College, 1520 St. Olaf Avenue, Northfield, MN 55057-1098. Manuscripts should be laser-printed, with wide line-spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit three copies and keep one copy. In addition, authors should supply the full five-symbol Mathematics Subject Classification number, as described in *Mathematical Reviews*, 1980 and later. Copies of figures should be supplied on separate sheets, both with and without lettering added.

Cover illustration. The 16 combinations of black and white triangles inside a square are arranged in a regularly repeating, 4×4 array. For each pair of adjacent squares, the colorings are mirror images of one another, except that their adjacent triangles are oppositely colored. The cover graphic and those on page 286 were created by Barry Cipra.

AUTHORS

Ray Beauregard received his Ph.D. at the University of New Hampshire in 1968, under the direction of Richard E. Johnson. Since that time he has been at the University of Rhode Island, where he is a professor of mathematics. His research interests include ring theory and number theory. He is co-author (with John Fraleigh) of a textbook, *Linear Algebra*, currently in its third edition. Although he has had little formal training for any of them, he enjoys babysitting for his granddaughters, sailing his sloop, *Aleph One*, playing the piano, and toying with triangles.

Jim Jordan is a true West Coast person—born in California, raised in California and Oregon, and spent most of his life in Washington. He has his B.S. degree in elementary education from Southern Oregon College, M.A. in mathematics from U. of Oregon, and Ph.D. in mathematics (number theory) from U. of Colorado. He has taught elementary school in Grants Pass, Oregon (3 years), spent 5.5 years as a teaching assistant (Oregan and Colorado), and 37.5 years with Washington State University. He was an associate editor of *Mathematics Magazine* when Roy Dubisch was editor. He plans to retire at the first stroke of midnight December 31, 1999. He brags about his children, his former graduate students, and his Mathematical Scholars Training Program.

Greg Martin grew up in Spring, Texas (near Houston) and attended Stanford University as an undergraduate, receiving his bachelor's degree in 1992. He completed his doctorate in analytic number theory at the University of Michigan in 1997, under the supervision of Trevor Wooley and Hugh Montgomery (and thus is probably the only person to have two Salem Prize winners as advisors). He worked for a year at the Institute for Advanced Study in Princeton, New Jersey until, fearing that his movements were becoming too predictable, he fled the country for a postdoctoral position at the University of Toronto. He supplements his mathematical interests with a passion for collegiate *a cappella* music.

Blake E. Peterson received a B.A. from Utah State University in 1986 and a Ph.D. from Washington State University in 1993, working under James H. Jordan. Although his Ph.D. was in pure mathematics, his experience as a high school teacher led him to a mathematics education position in the mathematics department at Oregon State University, where he taught until 1996. Currently teaching at Brigham Young University, he works closely with prospective secondary mathematics teachers.

E. R. Suryanarayan taught at the Universities of Gauhati and Pune in India before receiving his Ph.D. at the University of Michigan in 1961 under the direction of Nathaniel Coburn. He has been at the University of Rhode Island since 1960, where he is professor of mathematics. His research interests include applied mathematics, crystallography, and the history of mathematics with emphasis on Indian history. His hobbies include music, languages, and frequent summer escapades to India.

Vol. 72, No. 4, October 1999

MATHEMATICS MAGAZINE

The *MATHEMATICS MAGAZINE* (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, bimonthly except July/August.

The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $16 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $64. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 40% dues discount for the first two years of membership.) The nonmember/library subscription price is $68 per year.

Subscription correspondence and notice of change of address should be sent to the Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

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

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

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

Printed in the United States of America

ARTICLES

Farmer Ted Goes Natural

GREG MARTIN
University of Toronto
Toronto, Ontario
Canada M5S 3G3

1. Setting the stage

We've all been given a problem in a calculus class remarkably similar to the following one:

> Farmer Ted is building a chicken coop. He decides he can spare 190 square feet of his land for the coop, which will be built in the shape of a rectangle. Being a practical man, Farmer Ted wants to spend as little as possible on the chicken wire for the fence. What dimensions should he make the chicken coop?

By solving a simple optimization problem, we learn that Farmer Ted should make his chicken coop a square with side lengths $\sqrt{190}$ feet. And that, according to the solution manual, is that.

But the calculus books don't tell the rest of the story:

> So Farmer Ted went over to Builders Square and told the salesman, "I'd like $4\sqrt{190}$ feet of chicken wire, please." The salesman, however, replied that he could sell one foot or two feet or a hundred feet of chicken wire, but what the heck was $4\sqrt{190}$ feet of chicken wire? Farmer Ted was taken aback, explaining heatedly that his family had been buying as little chicken wire as possible for generations, and he really wanted $4\sqrt{190}$ feet of chicken wire measured off for him immediately! But the salesman, fearing more irrational behavior from Farmer Ted, told him, "I don't want to hear about your roots. We do business in a natural way here, and if you don't like it you can leave the whole store." Well, Farmer Ted didn't feel that this treatment was commensurate with his request, but he left Builders Square to rethink his coop from square one.

At first, Farmer Ted thought his best bet would be to make a $10' \times 19'$ chicken coop, necessitating the purchase of 58 feet of chicken wire—certainly this was better than 86 feet of chicken wire for a $5' \times 38'$ coop, say. But then he realized that he could be more cost-effective by not using all of the 190 square feet of land he had reserved for the coop. For instance, he could construct an $11' \times 17'$ coop (187 square feet) with only 56 feet of chicken wire; this would give him about 3.34 square feet of coop space per foot of chicken wire purchased, as opposed to only 3.28 square feet per chicken-wire-foot for the $10' \times 19'$ coop. Naturally, the parsimonious farmer wondered: could he do even better?

2. Posing the problem

Jon Grantham posed the following problem at the 1998 SouthEast Regional Meeting On Numbers in Greensboro, North Carolina: given a positive integer N, find the dimensions of the rectangle with integer side lengths and area at most N whose area-to-perimeter ratio is largest among all such rectangles. In the story above, Farmer Ted is trying to solve this problem for $N = 190$.

Let's introduce some notation so we can formulate Grantham's problem more precisely. For a positive integer n, let $s(n)$ denote the least possible semiperimeter (length plus width) of a rectangle with integer side lengths and area n. (Since the area-to-semiperimeter ratio of a rectangle is always twice the area-to-perimeter ratio, it doesn't really change the problem if we consider semiperimeters instead of perimeters; this will eliminate annoying factors of 2 in many of our formulas.) In other (and fewer) words,

$$s(n) = \min_{cd=n} (c+d) = \min_{d \mid n} (d + n/d),$$

where $d|n$ means that d divides n.

Let $F(n) = n/s(n)$ denote the area-to-semiperimeter ratio in which we are interested. We want to investigate the integers n such that $F(n)$ is large, and so we define the set $\mathscr{A}$ of "record-breakers" for the function F as follows:

$$\mathscr{A} = \{ n \in \mathbb{N} : F(k) \le F(n) \text{ for all } k \le n \}. \tag{1}$$

(Well, the "record-tiers" are also included in $\mathscr{A}$.) Then it is clear after a moment's thought that to solve Grantham's problem for a given number N, we simply need to find the largest element of $\mathscr{A}$ not exceeding N.

By computing all possible factorizations of the numbers up to 200 by brute force, we can make a list of the first 59 elements of $\mathscr{A}$:

$$\begin{aligned}\mathscr{A} = \{&1, 2, 3, 4, 6, 8, 9, 12, 15, 16, 18, 20, 24, 25, 28, 30, 35, 36, 40, 42, 48, 49, 54, 56, 60, 63, 64, 70, 72,\\ &77, 80, 81, 88, 90, 96, 99, 100, 108, 110, 117, 120, 121, 130, 132, 140, 143, 144, 150, 154, 156, 165,\\ &168, 169, 176, 180, 182, 192, 195, 196, \dots\}\end{aligned}$$

If we write, in place of the elements $n \in \mathscr{A}$, the dimensions of the rectangles with area n and least semiperimeter, we obtain

$$\begin{aligned}\mathscr{A} = \{&1\times1, 1\times2, 1\times3, 2\times2, 2\times3, 2\times4, 3\times3, 3\times4, 3\times5, 4\times4, 3\times6, 4\times5, 4\times6, 5\times5, 4\times7,\\ &5\times6, 5\times7, 6\times6, 5\times8, 6\times7, 6\times8, 7\times7, 6\times9, 7\times8, 6\times10, 7\times9, 8\times8, 7\times10, 8\times9,\\ &7\times11, 8\times10, 9\times9, 8\times11, 9\times10, 8\times12, 9\times11, 10\times10, 9\times12, 10\times11, 9\times13, 10\times12,\\ &11\times11, 10\times13, 11\times12, 10\times14, 11\times13, 12\times12, 10\times15, 11\times14, 12\times13, 11\times15,\\ &12\times14, 13\times13, 11\times16, 12\times15, 13\times14, 12\times16, 13\times15, 14\times14, \dots\},\end{aligned}$$

a list that exhibits a tantalizing promise of pattern! The interested reader is invited to try to determine the precise pattern of the set $\mathscr{A}$ before reading into the next section, in which the secret will be revealed. One thing we notice immediately, though, is that the dimensions of each of these rectangles are almost (or exactly) equal. For this reason, we will call the elements of $\mathscr{A}$ *almost-squares*. This supports our intuition about what the answers to Grantham's problem should be, since after all, Farmer Ted would build his rectangles with precisely equal sides if he weren't hampered by the integral policies of (the ironically-named) Builders Square.

From the list of the first 59 almost-squares, we find that 182 is the largest almost-square not exceeding 190. Therefore, Farmer Ted should build a chicken coop with area 182 square feet; and indeed, a $13' \times 14'$ coop would give him about 3.37 square feet of coop space per foot of chicken wire purchased, which is more cost-effective than the options he thought of back in Section 1. But what about next time, when Farmer Ted wants to build a supercoop on the 8,675,309 square feet of land he has to spare, or even more? Eventually, computations will need to give way to a better understanding of $\mathscr{A}$.

Our specific goals in this paper are to answer the following questions:

1. Can we describe $\mathscr{A}$ more explicitly? That is, can we characterize when a number n is an almost-square with a description that refers only to n itself, rather than all the numbers smaller than n? Can we find a formula for the number of almost-squares not exceeding a given positive number x?
2. Can we quickly compute the largest almost-square not exceeding N, for a given number N? We will describe more specifically what we mean by "quickly" in the next section, but for now we simply say that we'll want to avoid both brute force searches and computations that involve factoring integers.

In the next section, we will find that these questions have surprisingly elegant answers.

3. Remarkable results

Have you uncovered the pattern of the almost-squares? One detail you might have noticed is that all numbers of the form $m \times m$ and $(m-1) \times m$, and also $(m-1) \times (m+1)$, seem to be almost-squares. (If not, maybe we should come up with a better name for the elements of $\mathscr{A}$!) This turns out to be true, as we will see in Lemma 3 below. The problem is that there are other almost-squares than these—3×6, 4×7, 5×8, 6×9, 6×10—and the "exceptions" seem to become more and more numerous.... Even so, it will be convenient to think of the particular almost-squares of the form $m \times m$ and $(m-1) \times m$ as "punctuation" of a sort for $\mathscr{A}$. To this end, we will define a *flock* to be the set of almost-squares between $(m-1)^2 + 1$ and $m(m-1)$, or between $m(m-1)+1$ and m^2, including the endpoints in both cases.

If we group the rectangles corresponding to the almost-squares into flocks in this way, indicating the end of each flock by a semicolon, we obtain:

$$\mathscr{A} = \{1 \times 1;\ 1 \times 2;\ 1 \times 3, 2 \times 2;\ 2 \times 3;\ 2 \times 4, 3 \times 3;\ 3 \times 4;\ 3 \times 5, 4 \times 4;\ 3 \times 6, 4 \times 5;\ 4 \times 6, 5 \times 5;$$
$$4 \times 7, 5 \times 6;\ 5 \times 7, 6 \times 6;\ 5 \times 8, 6 \times 7;\ 6 \times 8, 7 \times 7;\ 6 \times 9, 7 \times 8;\ 6 \times 10, 7 \times 9, 8 \times 8;\ 7 \times 10,$$
$$8 \times 9;\ 7 \times 11, 8 \times 10, 9 \times 9;\ 8 \times 11, 9 \times 10;\ 8 \times 12, 9 \times 11, 10 \times 10;\ 9 \times 12, 10 \times 11;\ 9 \times 13,$$
$$10 \times 12, 11 \times 11;\ 10 \times 13, 11 \times 12;\ 10 \times 14, 11 \times 13, 12 \times 12;\ 10 \times 15, 11 \times 14, 12 \times 13;$$
$$11 \times 15, 12 \times 14, 13 \times 13;\ 11 \times 16, 12 \times 15, 13 \times 14;\ 12 \times 16, 13 \times 15, 14 \times 14;\ \ldots\}$$

It seems that all of the rectangles in a given flock have the same semiperimeter; this also turns out to be true, as we will see in Lemma 4 below. The remaining question, then, is to determine which rectangles of the common semiperimeter a given flock contains. At first it seems that all rectangles of the "right" semiperimeter will be in the flock as long as their area exceeds that of the last rectangle in the preceding flock, but then we note a few omissions—2×5, 3×7, 4×8, 5×9, 5×10—which also become more numerous if we extend our computations of $\mathscr{A}$....

But as it happens, this question can be resolved, and we can actually determine exactly which numbers are almost-squares, as our main theorem indicates. Recall that $\lfloor x \rfloor$ denotes the greatest integer not exceeding x.

MAIN THEOREM. *For any integer $m \geq 2$, the set of almost-squares between $(m-1)^2 + 1$ and m^2 (inclusive) consists of two flocks, the first of which is*

$$\{(m + a_m)(m - a_m - 1), (m + a_m - 1)(m - a_m), \ldots, (m+1)(m-2), m(m-1)\}$$

where $a_m = \lfloor (\sqrt{2m-1} - 1)/2 \rfloor$, and the second of which is

$$\{(m + b_m)(m - b_m), (m + b_m - 1)(m - b_m + 1), \ldots, (m+1)(m-1), m^2\}$$

where $b_m = \lfloor \sqrt{m/2} \rfloor$.

The Main Theorem allows us to easily enumerate the almost-squares in order, but if we simply want an explicit characterization of almost-squares without regard to their order, there turns out to be one that is extremely elegant. To describe it, we recall that the *triangular numbers* $\{0, 1, 3, 6, 10, 15, \ldots\}$ are the numbers $t_n = \binom{n}{2} = n(n-1)/2$ (Conway and Guy **[1]** describe many interesting properties of these and other "figurate" numbers). We let $T(x)$ denote the number of triangular numbers not exceeding x. (Notice that in our notation, $t_1 = \binom{1}{2} = 0$ is the first triangular number, so that $T(1) = 2$, for instance.) Then an alternative interpretation of the Main Theorem is the following:

COROLLARY 1. *The almost-squares are precisely those integers that can be written in the form $k(k+h)$, for some integers $k \geq 1$ and $0 \leq h \leq T(k)$.*

It is perhaps not so surprising that the triangular numbers are connected to the almost-squares—after all, adding t_m to itself or to t_{m+1} yields almost-squares of the form $m(m-1)$ or m^2, respectively (FIGURE 1 illustrates this for $m = 6$). In any case,

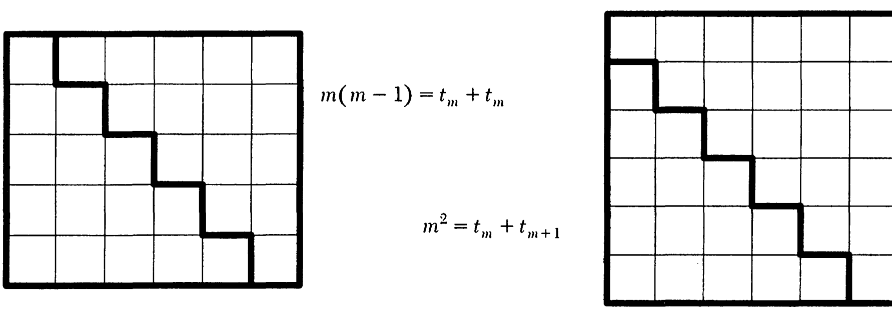


FIGURE 1
Two triangular integers invoke an almost-square.

the precision of this characterization is quite attractive and unexpected, and it is conceivable that Corollary 1 has a direct proof that doesn't use the Main Theorem. We leave this as an open problem for the reader.

In a different direction, we can use the Main Theorem's precise enumeration of the almost-squares in each flock to count the number of almost-squares quite accurately.

COROLLARY 2. *Let $A(x)$ denote the number of almost-squares not exceeding x. Then for $x \geq 1$,*

$$A(x) = \frac{2\sqrt{2}}{3}x^{3/4} + \frac{1}{2}x^{1/2} + R(x),$$

where $R(x)$ is an oscillating term whose order of magnitude is $x^{1/4}$.

A graph of $A(x)$ (see FIGURE 2) exhibits a steady growth with a little bit of a wiggle. When we isolate $R(x)$ by subtracting the main term $2\sqrt{2}x^{3/4}/3 + x^{1/2}/2$ from $A(x)$, the resulting graph (FIGURE 2, where we have plotted a point every time x passes an almost-square) is a pyrotechnic, almost whimsical display that seems to suggest that our computer code needs to be rechecked. Yet this *is* the true nature of $R(x)$. When we prove Corollary 2 (in a more specific and precise form) in Section 6, we will see that there are two reasons that the "remainder term" $R(x)$ oscillates: there are oscillations on a local scale because the almost-squares flock towards the right half of

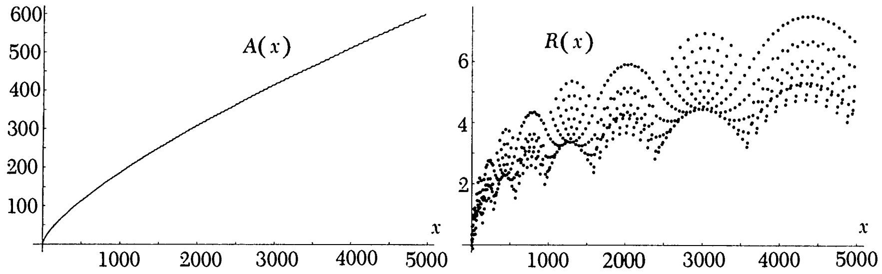


FIGURE 2
Superficial steadiness of $A(x)$, mesmerizing meanderings of $R(x)$.

each interval of the form $((m-1)^2, m(m-1)]$ or $(m(m-1), m^2]$, and oscillations on a larger scale for a less obvious reason.

These theoretical results about the structure of the almost-squares address question 1 nicely, and we turn our attention to the focus of question 2, the practicality of actually computing answers to questions about almost-squares. Even simple tasks like printing out a number or adding two numbers together obviously take time for a computer to perform, and they take longer for bigger numbers. To measure how the computing time needed for a particular computation increases as the size of the input grows, let $f(k)$ denote the amount of time it takes to perform the calculation on a k-digit number. Of course, the time could depend significantly on which k-digit number we choose; what we mean is the worst-case scenario, so that the processing time is at most $f(k)$ no matter which k-digit number we choose.

We say that a computation runs in *polynomial time* if this function $f(k)$ grows only as fast as a polynomial in k, i.e., if there are positive constants A and B such that $f(k) < Ak^B$. Generally speaking, the computations that we consider efficient to perform on very large inputs are those that run in polynomial time. (Because we are only concerned with this category of computations as a whole, it doesn't matter if we write our numbers in base 10 or base 2 or any other base, since this only multiplies the number of digits by a constant factor, like $\log_2 10$.)

All of our familiar arithmetic operations $+$, $-$, $\times$, $\div$, $\sqrt{\cdot}$, $\lfloor\cdot\rfloor$ and so on have polynomial-time algorithms. On the other hand, performing a calculation on each of the numbers from 1 to the input n, or even from 1 to $\sqrt{n}$, etc., is definitely not polynomial-time. Thus computing almost-squares by their definition, which involves comparing $F(n)$ with all of the preceding $F(k)$, is not efficient for large n. Furthermore, the obvious method of factoring numbers—testing all possible divisors in turn—is not polynomial-time for the same reason. While there are faster ways to factor numbers, at this time there is no known polynomial-time algorithm for factoring numbers; so factoring even a single number would make an algorithm inefficient. (Dewdney **[2]** writes about many facets of algorithms, including this property of running in polynomial time, while Pomerance **[4]** gives a more detailed discussion of factoring algorithms and their computational complexity.)

Fortunately, the Main Theorem provides a way to compute almost-squares that avoids both factorization and brute-force enumeration. In fact, we can show that all sorts of computations involving almost-squares are efficient:

COROLLARY 3. *There are polynomial-time algorithms to perform each of the following tasks, given a positive integer N:*

(a) *determine whether N is an almost-square, and, if so, determine the dimensions of the optimal rectangle;*

(b) *find the greatest almost-square not exceeding N, including the dimensions of the optimal rectangle;*
(c) *compute the number $A(N)$ of almost-squares not exceeding N;*
(d) *find the Nth almost-square, including the dimensions of the optimal rectangle.*

We reiterate that these algorithms work without ever factoring a single integer. Corollary 3, together with our lack of a polynomial-time factoring algorithm, has a rather interesting implication: for large values of N, it is much faster to compute the most cost-effective chicken coop (in terms of area-to-semiperimeter ratio) with area *at most* N than it is to compute the most cost-effective chicken coop with area *equal to* N, a somewhat paradoxical state of affairs! Nobody ever said farming was easy. . . .

4. The theorem thought through

Before proving the Main Theorem, we need to build up a stockpile of easy lemmas. The first of these simply confirms our expectations that the most cost-effective rectangle of a given area is the one whose side lengths are as close together as possible, and also provides some inequalities for the functions $s(n)$ and $F(n)$. Let us define $d(n)$ to be the largest divisor of n not exceeding $\sqrt{n}$ and $d'(n)$ the smallest divisor of n that is at least $\sqrt{n}$, so that $d'(n) = n/d(n)$.

LEMMA 1. *The rectangle with integer side lengths and area n that has the smallest semiperimeter is the one with dimensions $d(n) \times d'(n)$. In other words,*

$$s(n) = d(n) + d'(n).$$

We also have the inequalities

$$s(n) \geq 2\sqrt{n} \quad \text{and} \quad F(n) \leq \sqrt{n}/2.$$

Proof. For a fixed positive number n, the function $f(t) = t + n/t$ has derivative $f'(t) = 1 - n/t^2$, which is negative for $1 \leq t < \sqrt{n}$. Therefore $t + n/t$ is a decreasing function of t in that range. Thus if we restrict our attention to those t such that both t and n/t are positive integers (in other words, t is an integer dividing n), we see that the expression $t + n/t$ is minimized when $t = d(n)$. We therefore have

$$s(n) = d(n) + n/d(n) \geq 2\sqrt{n},$$

where the last inequality follows from the Arithmetic Mean–Geometric Mean inequality. The inequality for $F(n)$ then follows directly from the definition of F. ■

If we have a number n written as $c \times d$, where c and d are pretty close to $\sqrt{n}$, when can we say that there isn't some better factorization out there, so that $s(n)$ is really equal to $c + d$? The following lemma gives us a useful criterion.

LEMMA 2. *If a number n satisfying $(m-1)^2 < n \leq m(m-1)$ has the form $n = (m-a-1)(m+a)$ for some number a, then $s(n) = 2m - 1$, and $d(n) = m - a - 1$ and $d'(n) = m + a$. Similarly, if a number n satisfying $m(m-1) < n \leq m^2$ has the form $n = m^2 - b^2$ for some number b, then $s(n) = 2m$, and $d(n) = m - b$ and $d'(n) = m + b$.*

Proof. First let's recall that, for any positive real numbers α and β, the pair of equations $r + s = \alpha$ and $rs = \beta$ has a unique solution (r, s) with $r \leq s$, as long as the Arithmetic Mean–Geometric Mean inequality $\alpha/2 \geq \sqrt{\beta}$ holds. This is because r

and s will be the roots of the quadratic polynomial $t^2 - \alpha t + \beta$, which has real roots when its discriminant $\alpha^2 - 4\beta$ is nonnegative, i.e., when $\alpha/2 \geq \sqrt{\beta}$.

Now if $n = (m-a-1)(m+a)$, then clearly $s(n) \leq (m-a-1)+(m+a) = 2m-1$ by the definition of $s(n)$. On the other hand, by Lemma 1 we know that $s(n) \geq 2\sqrt{n} > 2(m-1)$, and so $s(n) = 2m-1$ exactly. We now know that $d(n)d'(n) = n = (m-a-1)(m+a)$ and

$$d(n) + d'(n) = s(n) = 2m - 1 = (m-a-1) + (m+a),$$

and of course $d(n) \leq d'(n)$ as well; by the argument of the previous paragraph, we conclude that $d(n) = m-a-1$ and $d'(n) = m+a$. This establishes the first assertion of the lemma, and a similar argument establishes the second assertion. ■

Of course, if a number n satisfies $s(n) = 2m$ for some m, then n can be written as $n = cd$ with $c \leq d$ and $c + d = 2m$; and letting $b = d - m$, we see that $n = cd = (2m-d)d = (m-b)(m+b)$. A similar statement is true if $s(n) = 2m-1$, and so we see that the converse of Lemma 2 also holds. We also remark that in the statement of the lemma, the two expressions $m(m-1)$ can be replaced by $(m-1/2)^2 = m(m-1)+1/4$ if we wish.

Lemma 2 implies in particular that for $m \geq 2$,

$$s(m^2) = 2m, \quad s(m(m-1)) = 2m-1, \quad \text{and} \quad s((m-1)(m+1)) = 2m,$$

and so

$$F(m^2) = \frac{m}{2}, \quad F(m^2 - m) = \frac{m(m-1)}{2m-1}, \quad \text{and} \quad F(m^2-1) = \frac{m^2-1}{2m}.$$

Using these facts, we can verify our theory that these numbers are always almost-squares.

LEMMA 3. *Each positive integer of the form m^2, $m(m-1)$, or $m^2 - 1$ is an almost-square.*

It is interesting to note that these are precisely those integers n that are divisible by $\lfloor \sqrt{n} \rfloor$ (see **[3]**), one of the many interesting things that can be discovered by referring to Sloane and Plouffe **[6]**.

Proof. We verify directly that such numbers satisfy the condition in the definition (1) of $\mathscr{A}$. If $k < m^2$, then by Lemma 1 we have $F(k) \leq \sqrt{k}/2 < m/2 = F(m^2)$, and so m^2 is an almost-square. Similarly, if $k < m(m-1)$, then again

$$F(k) \leq \frac{\sqrt{k}}{2} \leq \frac{\sqrt{m^2-m-1}}{2} < \frac{m(m-1)}{2m-1} = F(m(m-1)),$$

where the strict inequality can be verified as a "fun" algebraic exercise. Thus $m(m-1)$ is also an almost-square. A similar argument shows that $m^2 - 1$ is also an almost-square. ■

Now we're getting somewhere! Next we show that the semiperimeters of the rectangles corresponding to the almost-squares in a given flock are all equal, as we observed at the beginning of Section 3.

LEMMA 4. *Let $m \geq 2$ be an integer. If n is an almost-square satisfying $(m-1)^2 < n \leq m(m-1)$, then $s(n) = 2m-1$; similarly, if n is an almost-square satisfying $m(m-1) < n \leq m^2$, then $s(n) = 2m$.*

Proof. If $n = m(m-1)$, we have already shown that $s(n) = 2m-1$. If n satisfies $(m-1)^2 < n < m(m-1)$, then by Lemma 1 we have $s(n) \geq 2\sqrt{n} > 2(m-1)$. On the other hand, since n is an almost-square exceeding $(m-1)^2$, we have

$$\frac{m-1}{2} = F\big((m-1)^2\big) \leq F(n) = \frac{n}{s(n)} < \frac{m(m-1)}{s(n)},$$

and so $s(n) < 2m$. Therefore $s(n) = 2m-1$ in this case.

Similarly, if n satisfies $m(m-1) < n < m^2$, then $s(n) \geq 2\sqrt{n} \geq 2\sqrt{m^2-m+1} > 2m-1$; on the other hand,

$$\frac{m(m-1)}{2m-1} = F(m(m-1)) \leq F(n) = \frac{n}{s(n)} \leq \frac{m^2-1}{s(n)},$$

and so $s(n) < (m+1)(2m-1)/m < 2m+1$. Therefore $s(n) = 2m$ in this case. ∎

Finally, we need to exhibit some properties of the sequences a_m and b_m defined in the statement of the Main Theorem.

LEMMA 5. *Define* $a_m = \lfloor(\sqrt{2m-1}-1)/2\rfloor$ *and* $b_m = \lfloor\sqrt{m/2}\rfloor$. *For any integer* $m \geq 2$:

(a) $a_m \leq b_m \leq a_m + 1$;
(b) $b_m = \lfloor m/\sqrt{2m-1}\rfloor$;
(c) $a_m + b_m = \lfloor\sqrt{2m}\rfloor - 1$.

We omit the proof of this lemma since it is tedious but straightforward. The idea is to show that in the sequences a_m, b_m, $\lfloor m/\sqrt{2m-1}\rfloor$, and $\lfloor\sqrt{2m}\rfloor$, two consecutive terms either are equal or else differ by 1, and then to determine precisely for what values of m the differences of 1 occur.

Armed with these lemmas, we are now ready to furnish a proof of the Main Theorem.

Proof of the Main Theorem. Fix an integer $m \geq 2$. By Lemma 4, every almost-square n with $(m-1)^2 < n \leq m(m-1)$ satisfies $s(n) = 2m-1$; while by Lemma 2, the integers $(m-1)^2 < n \leq m(m-1)$ satisfying $s(n) = 2m-1$ are precisely the elements of the form $n_a = (m-a-1)(m+a)$ that lie in that interval. Thus it suffices to determine which of the n_a are almost-squares.

Furthermore, suppose that n_a is an almost-square for some $a \geq 1$. Then $F(n_a) \geq F(n)$ for all $n < n_a$ by the definition of $\mathscr{A}$, while $F(n_a) > F(n)$ for all $n_a < n < n_{a-1}$ since we've already concluded that no such n can be an almost-square. Moreover, $n_{a-1} > n_a$ and $s(n_{a-1}) = 2m-1 = s(n_a)$, so $F(n_{a-1}) > F(n_a)$, and thus n_{a-1} is an almost-square as well. Therefore it suffices to find the largest value of a (corresponding to the smallest n_a) such that n_a is an almost-square.

By Lemma 3, we know that $(m-1)^2$ is an almost-square, and so we need to find the largest a such that $F(n_a) \geq F((m-1)^2)$, i.e.,

$$\frac{(m-a-1)(m+a)}{2m-1} \geq \frac{m-1}{2},$$

which is the same as $2a(a+1)+1 \leq m$. By completing the square and solving for a, we find that this inequality is equivalent to

$$\frac{-\sqrt{2m-1}-1}{2} \leq a \leq \frac{\sqrt{2m-1}-1}{2}, \tag{2}$$

and so the largest integer a satisfying the inequality is exactly $a = \lfloor (\sqrt{2m-1} - 1)/2 \rfloor = a_m$, as defined in the statement of the Main Theorem. This establishes the first part of the theorem.

By the same reasoning, it suffices to find the largest value of b such that $F(m^2 - b^2) \geq F(m(m-1))$, i.e.,

$$\frac{m^2 - b^2}{2m} \geq \frac{m(m-1)}{2m-1},$$

which is the same as

$$b^2 \leq m^2/(2m-1) \tag{3}$$

or $b \leq \lfloor m/\sqrt{2m-1} \rfloor$. But by Lemma 5(b), $\lfloor m/\sqrt{2m-1} \rfloor = b_m$ for $m \geq 2$, and so the second part of the theorem is established. ■

With the Main Theorem now proven, we remark that Lemma 5(c) implies that for any integer $m \geq 2$, the number of almost-squares in the two flocks between $(m-1)^2 + 1$ and m^2 is exactly $(1 + a_m) + (1 + b_m) = 1 + \lfloor \sqrt{2m} \rfloor$, while Lemma 5(a) implies that there are either equally many in the two flocks or else one more in the second flock than in the first.

5. Taking notice of triangular numbers

Our next goal is to derive Corollary 1 from the Main Theorem. First we establish a quick lemma giving a closed-form expression for $T(x)$, the number of triangular numbers not exceeding x.

LEMMA 6. *For all $x \geq 0$, we have $T(x) = \lfloor \sqrt{2x + 1/4} + 1/2 \rfloor$.*

Proof. $T(x)$ is the number of positive integers n such that $t_n \leq x$, or $n(n-1)/2 \leq x$. This inequality is equivalent to $(n - 1/2)^2 \leq 2x + 1/4$, or $-\sqrt{2x + 1/4} + 1/2 \leq n \leq \sqrt{2x + 1/4} + 1/2$. The left-hand expression never exceeds $1/2$, and so $T(x)$ is simply the number of positive integers n such that $n \leq \sqrt{2x + 1/4} + 1/2$; in other words, $T(x) = \lfloor \sqrt{2x + 1/4} + 1/2 \rfloor$ as desired. ■

Proof of Corollary 1. Suppose first that $n = k(k+h)$ for some integers $k \geq 1$ and $h \leq T(k)$. Let $k' = k + h$, and define

$$\begin{cases} m = k + (h+1)/2 \quad \text{and} \quad a = (h-1)/2 & \text{if } h \text{ is odd;} \\ m = k + h/2 \quad \text{and} \quad b = h/2, & \text{if } h \text{ is even,} \end{cases}$$

so that

$$\begin{cases} k = m - a - 1 \quad \text{and} \quad k' = m + a & \text{if } h \text{ is odd;} \\ k = m - b \quad \text{and} \quad k' = m + b & \text{if } h \text{ is even.} \end{cases}$$

We claim that

$$\begin{cases} (m-1)^2 < (m-a-1)(m+a) \leq (m-1/2)^2 & \text{if } h \text{ is odd;} \\ (m-1/2)^2 < m^2 - b^2 \leq m^2 & \text{if } h \text{ is even.} \end{cases} \tag{4}$$

To see this, note that in terms of k and h, these inequalities become

$$\left(k+\frac{h-1}{2}\right)^2 < k(k+h) \le \left(k+\frac{h}{2}\right)^2.$$

A little bit of algebra reveals that the right-hand inequality is trivially satisfied, while the left-hand inequality is true provided that $h < 2\sqrt{k} + 1$. However, from Lemma 6 we see that

$$T(k) = \lfloor \sqrt{2k+1/4} + 1/2 \rfloor \le \sqrt{2k+1/4} + 1/2 < 2\sqrt{k} + 1$$

for $k \ge 1$. Since we are assuming that $h \le T(k)$, this shows that the inequalities (4) do indeed hold.

Because of these inequalities, we may apply Lemma 2 (see the remarks following the proof of the lemma) and conclude that

$$\begin{cases} s(n) = 2m-1,\ d(n) = m-a-1, \text{ and } d'(n) = m+a & \text{if } h \text{ is odd;} \\ s(n) = 2m,\ d(n) = m-b, \text{ and } d'(n) = m+b & \text{if } h \text{ is even.} \end{cases}$$

Consequently, the Main Theorem asserts that n is an almost-square if and only if

$$\begin{cases} a \le a_m & \text{if } h \text{ is odd;} \\ b \le b_m & \text{if } h \text{ is even,} \end{cases} \tag{5}$$

which by the definitions of a, b, and m is the same as

$$\begin{cases} (h-1)/2 \le \lfloor (\sqrt{2m-1} - 1)/2 \rfloor = \lfloor (\sqrt{2k+h} - 1)/2 \rfloor & \text{if } h \text{ is odd;} \\ h/2 \le \lfloor \sqrt{m/2} \rfloor = \lfloor \sqrt{k/2+h/4} \rfloor & \text{if } h \text{ is even.} \end{cases}$$

Since in either case, the left-hand side is an integer, the greatest-integer brackets can be removed from the right-hand side, whence both cases reduce to $h \le \sqrt{2k+h}$. From here, more algebra reveals that this inequality is equivalent to $h \le \sqrt{2k+1/4} + 1/2$; and since h is an integer, we can add greatest-integer brackets to the right-hand side, thus showing that the inequality (5) is equivalent to $h \le T(k)$ (again using Lemma 6). In particular, n is indeed an almost-square.

This establishes half of the characterization asserted by Corollary 1. Conversely, suppose we are given an almost-square n, which we can suppose to be greater than 1 since 1 can obviously be written as $1(1+0)$. If we let $h = d'(n) - d(n)$, then the Main Theorem tells us that

$$\begin{cases} n = (m-a-1)(m+a),\ d(n) = m-a-1, \text{ and } d'(n) = m+a & \text{if } h \text{ is odd;} \\ n = m^2 - b^2,\ d(n) = m-b, \text{ and } d'(n) = m+b & \text{if } h \text{ is even} \end{cases}$$

for some integers $m \ge 2$ and either a with $0 \le a \le a_m$ or b with $0 \le b \le b_m$. If we set $k = d(n)$, then certainly $n = k(k+h)$. Moreover, the algebraic steps showing that the inequality (5) is equivalent to $h \le T(k)$ are all reversible; and (5) does in fact hold, since we are assuming that n is an almost-square. Therefore n does indeed have a representation of the form $k(k+h)$ with $0 \le h \le T(k)$. This establishes the corollary. ■

Pinpointing the pioneers We take a slight detour at this point to single out some special almost-squares. Let us make the convention that the kth flock refers to the flock of almost-squares with semiperimeter k, so that the first flock is actually empty, the second and third poor flocks contain only $1 = 1 \times 1$ and $2 = 1 \times 2$, respectively, the fourth flock contains $3 = 1 \times 3$ and $4 = 2 \times 2$, and so on. The Main Theorem tells us that a_m and b_m control the number of almost-squares in the odd-numbered and even-numbered flocks, respectively; thus, every so often, a flock will have one more almost-square than the preceding flock of the same "parity." We'll let a *pioneer* be an almost-square that begins one of these suddenly-longer flocks.

For instance, from the division of $\mathscr{A}$ into flocks on page 261, we see that the 4th flock $\{1 \times 3, 2 \times 2\}$ is longer than the preceding even-numbered flock $\{1 \times 1\}$, so $1 \times 3 = 3$ is the first pioneer; the 9th flock $\{3 \times 6, 4 \times 5\}$ is longer than the preceding odd-numbered flock $\{3 \times 4\}$, so $3 \times 6 = 18$ is the second pioneer; and so on, the next two pioneers being 6×10 in the 16th flock and 10×15 in the 25th flock. Now if this isn't a pattern waiting for a proof, nothing is! The following lemma shows another elegant connection between the almost-squares and the squares and triangular numbers.

COROLLARY 4. *For any positive integer j, the jth pioneer equals $t_{j+1} \times t_{j+2}$ (where t_i is the ith triangular number), which begins the $(j+1)^2$-th flock. Furthermore, the "record-tying" almost-squares (those whose F-values are equal to the F-values of their immediate predecessors in $\mathscr{A}$) are precisely the even-numbered pioneers.*

Proof. First, Lemma 5(a) tells us that the odd- and even-numbered flocks undergo their length increases in alternation, so that the pioneers alternately appear in the flocks of each parity. The first pioneer $3 = 1 \times 3$ appears in the 4th flock, and corresponds to $m = 2$ and the first appearance of $b_m = 1$ in the notation of the Main Theorem. Thus the $(2k-1)$-st pioneer will equal $m^2 - k^2$, where m corresponds to the first appearance of $b_m = k$. It is easy to see that the first appearance of $b_m = k$ occurs when $m = 2k^2$, in which case the $(2k-1)$-st pioneer is

$$m^2 - k^2 = (2k^2)^2 - k^2 = (2k^2 - k)(2k^2 + k)$$
$$= \frac{2k(2k-1)}{2} \frac{(2k+1)2k}{2} = t_{2k} t_{2k+1}.$$

Moreover, the flock in which this pioneer appears is the $2m$-th or $(2k)^2$-th flock.

Similarly, the $2k$-th pioneer will equal $(m-k-1)(m+k)$, where m corresponds to the first appearance of $a_m = k$. Again one can show that the first appearance of $a_m = k$ occurs when $m = 2k^2 + 2k + 1$, in which case the $2k$-th pioneer is

$$(m-k-1)(m+k) = (2k^2 + k)(2k^2 + 3k + 1)$$
$$= \frac{(2k+1)2k}{2} \frac{(2k+2)(2k+1)}{2} = t_{2k+1} t_{2k+2}.$$

Moreover, the flock in which this pioneer appears is the $(2m-1)$-st or $(2k+1)^2$-th flock. This establishes the first assertion of the corollary.

Since the F-values of the almost-squares form a nondecreasing sequence by the definition of almost-square, to look for almost-squares with equal F-values we only need to examine consecutive almost-squares. Furthermore, two consecutive almost-squares in the same flock never have equal F-values, since they are distinct numbers but by Lemma 4 their semiperimeters are the same. Therefore we only need to determine when the last almost-square in a flock can have the same F-value as the first almost-square in the following flock.

The relationship between the F-values of these pairs of almost-squares was determined in the proof of the Main Theorem. Specifically, the equality $F((m-1)^2) = F((m-a-1)(m+a))$ holds if and only if the right-hand inequality in (2) is actually an equality; this happens precisely when $m = 2a^2 + 2a + 1$, which corresponds to the even-numbered pioneers as was determined above. On the other hand, the equality $F(m(m-1)) = F(m^2-b^2)$ holds if and only if the inequality (3) is actually an equality; but m^2 and $2m-1$ are always relatively prime (any prime factor of m^2 must divide m and thus divides into $2m-1$ with a "remainder" of -1), implying that $m^2/(2m-1)$ is never an integer for $m \geq 2$, and so the inequality (3) can never be an equality. This establishes the second assertion of the corollary. ■

Pairs of triangles: product testing We know that all squares are almost-squares, and so t_j^2 is certainly an almost-square for any triangular number t_j; also, Corollary 4 tells us that the product $t_j t_{j+1}$ of two consecutive triangular numbers is always an almost-square. This led the author to wonder which numbers of the form $t_m t_n$ are almost-squares. If m and n differ by more than 1, it would seem that the rectangle of dimensions $t_m \times t_n$ is not the most cost-effective rectangle of area $t_m t_n$, and so the author expected that these products of two triangular numbers would behave randomly with respect to being almost-squares—that is, a few of them might be but most of them wouldn't. After some computations, however, FIGURE 3 emerged, where a point has been plotted in the (m, n) position if and only if $t_m t_n$ is an almost-square; and the table exhibited a totally unexpected regularity.

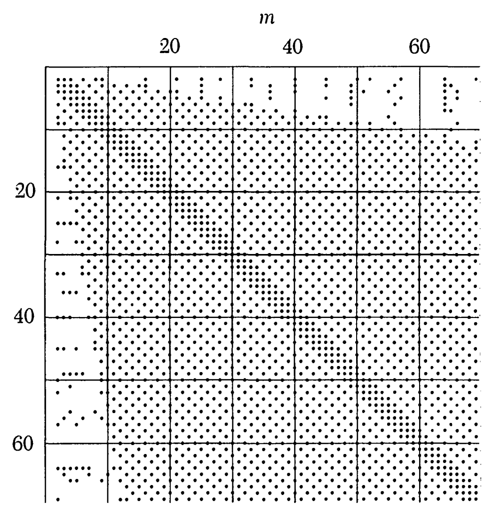


FIGURE 3
Amazing almost-square patterns in products of two triangles.

Of course the symmetry of the table across the main diagonal is to be expected since $t_m t_n = t_n t_m$. The main diagonal and the first off-diagonals are filled with plotted points, corresponding to the almost-squares t_m^2 and $t_m t_{m+1}$; and, in hindsight, the second off-diagonals correspond to

$$t_m t_{m+2} = \frac{m(m-1)}{2} \frac{(m+2)(m+1)}{2} = \frac{m^2+m-2}{2} \frac{m^2+m}{2},$$

which is the product of two consecutive integers (since m^2+m is always even) and is thus an almost-square as well. But apart from these central diagonals and some garbage along the edges of the table where m and n are quite different in size, the checkerboard-like pattern in the kite-shaped region of the table seems to be telling us that the only thing that matters in determining whether $t_m t_n$ is an almost-square is whether m and n have the same parity!

Once this phenomenon had been discovered, it turned out that the following corollary could be derived from the prior results in this paper. We leave the proof as a challenge to the reader.

COROLLARY 5. *Let m and n be positive integers with $n-1>m>3n-\sqrt{8n(n-1)}-1$. Then $t_m t_n$ is an almost-square if and only if $n-m$ is even.*

We remark that the function $3n-\sqrt{8n(n-1)}-1$ is asymptotic to $(3-2\sqrt{2})n+(\sqrt{2}-1)$, which explains the straight lines of slope $-(3-2\sqrt{2})\approx -0.17$ and $-1/(3-2\sqrt{2})\approx -5.83$ that seem to the eye to separate the orderly central region in FIGURE 3 from the garbage along the edges.

6. Counting and computing

In this section we establish Corollaries 2 and 3. We begin by defining a function $B(x)$ that will serve as the backbone of our investigation of the almost-square counting function $A(x)$. Let $\{x\}=x-\lfloor x\rfloor$ denote the fractional part of x, and define the quantities $\gamma=\gamma(x)=\{\sqrt{2}x^{1/4}\}$ and $\delta=\delta(x)=\{x^{1/4}/\sqrt{2}\}$. Let $B(x)=B_0(x)+B_1(x)$, where

$$B_0(x)=\frac{2\sqrt{2}}{3}x^{3/4}+\frac{1}{2}x^{1/2}+\left(\frac{2\sqrt{2}}{3}+\frac{\gamma(1-\gamma)}{\sqrt{2}}\right)x^{1/4} \qquad (6)$$

and

$$B_1(x)=\frac{\gamma^3}{6}-\frac{\gamma^2}{4}-\frac{5\gamma}{12}-\frac{\delta}{2}-1.$$

We remark that $\gamma=\{2\delta\}$ and that $B_1(x^4)$ is a periodic function of x with period $\sqrt{2}$, and so it is easy to check that the inequalities $-2\le B_1(x)\le -1$ always hold. The following lemma shows how the strange function $B(x)$ arises in connection with the almost-squares.

LEMMA 7. *For any integer $M\ge 1$, we have $A(M^2)=B(M^2)$.*

Proof. As remarked at the end of Section 4, the number of almost-squares between $(m-1)^2+1$ and m^2 is $\lfloor\sqrt{2m}\rfloor+1$ for $m\ge 2$. Therefore

$$A(M^2)=1+\sum_{m=2}^{M}\left(\lfloor\sqrt{2m}\rfloor+1\right)=M-1+\sum_{m=1}^{M}\lfloor\sqrt{2m}\rfloor.$$

It's almost always a good idea to interchange orders of summation whenever possible —and if there aren't enough summation signs, find a way to create some more! In this case, we convert the greatest-integer function into a sum of ones over the appropriate range of integers:

$$\begin{aligned}A(M^2)&=M-1+\sum_{m=1}^{M}\sum_{1\le k\le\sqrt{2m}}1\\&=M-1+\sum_{1\le k\le\sqrt{2M}}\ \sum_{k^2/2\le m\le M}1\\&=M-1+\sum_{\substack{1\le k\le\sqrt{2M}\\ k\text{ odd}}}\left(M-\frac{k^2-1}{2}\right)+\sum_{\substack{1\le k\le\sqrt{2M}\\ k\text{ even}}}\left(M-\left(\frac{k^2}{2}-1\right)\right).\end{aligned}$$

If we temporarily write μ for $\lfloor\sqrt{2M}\rfloor$, then

$$A(M^2) = M - 1 + \mu\left(M + \frac{1}{2}\right) - \frac{1}{2}\sum_{k=1}^{\mu} k^2 + \sum_{\substack{k=1 \\ k \text{ even}}}^{\mu} \frac{1}{2}$$

$$= M(\mu+1) + \frac{\mu}{2} - 1 - \frac{1}{2}\frac{\mu(\mu+1)(2\mu+1)}{6} - \frac{1}{2}\left\lfloor\frac{\mu}{2}\right\rfloor, \tag{7}$$

using the well-known formula for the sum of the first μ squares. Since $\lfloor\lfloor x\rfloor/n\rfloor = \lfloor x/n\rfloor$ for any real number $x \geq 0$ and any positive integer n, the last term can be written as

$$\frac{1}{2}\left\lfloor\frac{\mu}{2}\right\rfloor = \frac{1}{2}\left\lfloor\frac{\lfloor\sqrt{2M}\rfloor}{2}\right\rfloor = \frac{1}{2}\left\lfloor\sqrt{\frac{M}{2}}\right\rfloor = \frac{1}{2}\sqrt{\frac{M}{2}} - \frac{1}{2}\left\{\sqrt{\frac{M}{2}}\right\} = \frac{1}{2}\sqrt{\frac{M}{2}} - \frac{\delta(M^2)}{2},$$

while we can replace the other occurrences of μ in equation (7) by $\sqrt{2M} - \{\sqrt{2M}\} = \sqrt{2M} - \gamma(M^2)$. Writing γ for $\gamma(M^2)$ and δ for $\delta(M^2)$, we see that

$$\begin{aligned}
A(M^2) &= M\left(\sqrt{2M} - \gamma + 1\right) + \sqrt{2M} - \gamma - 1 \\
&\quad - \frac{1}{2}\frac{(\sqrt{2M} - \gamma)(\sqrt{2M} - \gamma + 1)\left(2(\sqrt{2M} - \gamma) + 1\right)}{6} - \frac{1}{2}\sqrt{\frac{M}{2}} + \frac{\delta}{2} \\
&= \frac{2\sqrt{2}}{3}M^{3/2} + \frac{M}{2} + \left(\frac{2\sqrt{2}}{3} - \frac{\gamma(1-\gamma)}{\sqrt{2}}\right)\sqrt{M} + \frac{\gamma^3}{6} - \frac{\gamma^2}{4} - \frac{5\gamma}{12} - 1 - \frac{\delta}{2} \\
&= B(M^2)
\end{aligned}$$

after much algebraic simplification. This establishes the lemma. ∎

Now $B(x)$ is a rather complicated function of x, but the next lemma gives us a couple of ways to predict the behavior of $B(x)$. First, it tells us how to predict $B(x+y)$ from $B(x)$ if y is small compared to x (roughly speaking, their difference will be $y/\sqrt{2}x^{1/4}$); second, it tells us how to predict approximately when $B(x)$ assumes a given integer value.

LEMMA 8. *There is a positive constant C such that*

(a) *for all real numbers $x \geq 1$ and $0 \leq y \leq \min\{x/2, 3\sqrt{x}\}$, we have*

$$\left|B(x+y) - B(x) - \frac{y}{\sqrt{2}x^{1/4}}\right| < C; \tag{8}$$

(b) *if we define $z_j = \frac{1}{2}(3j)^{2/3} - \frac{1}{4}(3j)^{1/3}$ for any positive integer j, then for all $j > C$ we have $z_j > 2$ and $B((z_j - 1)^2) < j < B(z_j^2)$.*

If the proof of Lemma 5 was omitted due to its tediousness, the proof of this lemma should be omitted and then buried.... The idea of the proof is to rewrite $B_0(x)x^{-3/4}$ using the new variable $t = x^{-1/4}$, and then expand in a Taylor series in t (a slight but easily overcome difficulty being that the term $\gamma(x)(1-\gamma(x))$ is not differentiable when $\sqrt{2}x^{1/4}$ is an integer). For the proof of part (b), we also need to rewrite $z_j j^{-2/3}$ using the new variable $u = j^{-1/3}$ and expand in a Taylor series in u. We remark that the constant C in Lemma 8 can be taken to be quite small—in fact, $C = 5$ will suffice.

With these last lemmas in hand, we can dispatch Corollaries 2 and 3 in quick succession.

Proof of Corollary 2. Let $x > 1$ be a real number and define $R(x) = A(x) - 2\sqrt{2}x^{3/4}/3 - \sqrt{x}/2$, as in the statement of the corollary. We will describe how to prove the following more precise statement:

$$R(x) = \left(\frac{2\sqrt{2}}{3} + g(\sqrt{2}x^{1/4}) - h(2\sqrt{x})\right)x^{1/4} + R_1(x), \tag{9}$$

where

$$g(t) = \frac{\{t\}(1-\{t\})}{\sqrt{2}} \quad \text{and} \quad h(t) = \begin{cases} \dfrac{\{t\}}{\sqrt{2}} & \text{if } 0 \le \{t\} \le \dfrac{1}{2}; \\ \sqrt{1-\{t\}} - \dfrac{1-\{t\}}{\sqrt{2}} & \text{if } \dfrac{1}{2} \le \{t\} \le 1 \end{cases}$$

and $|R_1(x)| < D$ for some constant D. The functions g and h are continuous and periodic with period 1, and are the causes of the oscillations in the error term $R(x)$. The expression $g(\sqrt{2}x^{1/4})$ goes through a complete cycle when x increases by about $2\sqrt{2}x^{3/4}$ (one can show this using Taylor expansions yet again!), which causes the large-scale bounces in the normalized error term $R(x)x^{-1/4}$ shown in FIGURE 4 below. Similarly, the expression $h(2\sqrt{x})$ goes through a complete cycle when x increases by about $\sqrt{x}$, which causes the smaller-scale stutters shown in the (horizontally magnified) right-hand graph in FIGURE 4.

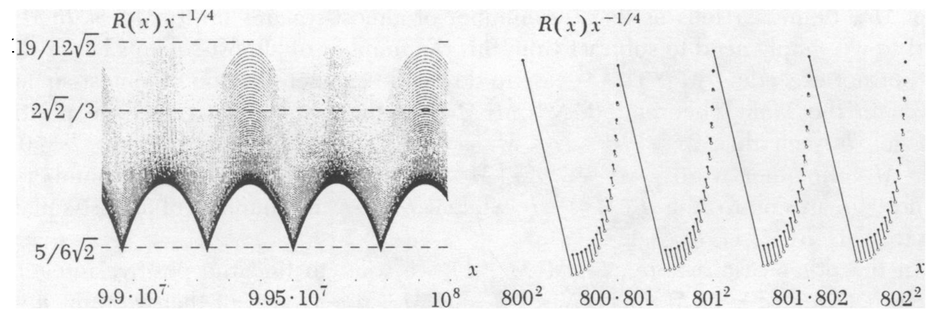


FIGURE 4
Big bounces and small stutters for $R(x)x^{-1/4}$.

To establish the formula (9), we shrewdly add $B(x) - B(x)$ to the expression defining $R(x)$, which yields

$$R(x) = (A(x) - B(x)) + \left(\frac{2\sqrt{2}}{3} + \frac{\gamma(x)(1-\gamma(x))}{\sqrt{2}}\right)x^{1/4} + B_1(x)$$

from the definition (6) of $B_0(x)$. Now $B_1(x)$ is a bounded function; and since $\gamma(x) = \{\sqrt{2}x^{1/4}\}$, the expression $\gamma(x)(1-\gamma(x))/\sqrt{2}$ is precisely $g(\sqrt{2}x^{1/4})$. So what we need to show is that $B(x) - A(x) = h(2\sqrt{x})x^{1/4} + R_2(x)$, where $R_2(x)$ is another bounded function.

While we won't give all the details, the outline of showing this last fact is as follows: Suppose first that $x \ge m^2$ but that x is less than the first almost-square $(m+1+a_{m+1})(m-a_{m+1})$ in the $(2m+1)$-st flock, so that $A(x) = A(m^2)$. Since $A(m^2) = B(m^2)$ by Lemma 7, we only need to show that $B(x) - B(m^2)$ is approximately $h(2\sqrt{x})x^{1/4}$; this we can accomplish with the help of Lemma 8(a).

Similarly, if $x < m^2$ but x is at least as large as the first almost-square $(m+b_m)(m-b_m)$ in the $2m$-th flock, the same method works as long as we take into account the difference between $A(m^2)$ and $A(x)$, which is $\lfloor\sqrt{m^2-x}\rfloor$ by the Main Theorem. And if x is close to an almost-square of the form $m(m-1)$ rather than m^2, the same method applies; even though $A(m(m-1))$ and $B(m(m-1))$ are not exactly equal, they differ by a bounded amount.

Notice that the functions $g(t)$ and $h(t)$ take values in $[0, 1/4\sqrt{2}]$ and $[0, 1/2\sqrt{2}]$, respectively. From this and the formula (9) we can conclude that

$$\liminf_{x\to\infty}\frac{R(x)}{x^{1/4}}=\frac{5}{6\sqrt{2}}\quad\text{and}\quad\limsup_{x\to\infty}\frac{R(x)}{x^{1/4}}=\frac{19}{12\sqrt{2}}.$$

The interested reader can check, for example, that the sequences $y_j=4j^4+j^2$ and $z_j=(2j^2+j)^2$ satisfy $\lim_{j\to\infty}R(y_j)/y_j^{1/4}=5/6\sqrt{2}$ and $\lim_{j\to\infty}R(z_j)/z_j^{1/4}=19/12\sqrt{2}$. ■

Proof of Corollary 3. The algorithms we describe will involve only the following types of operations: performing ordinary arithmetical calculations $+$, $-$, $\times$, $\div$; computing the greatest-integer $\lfloor\cdot\rfloor$, least-integer $\lceil\cdot\rceil$, and fractional-part $\{\cdot\}$ functions; taking square roots, cube roots, and fourth roots; and comparing two numbers to see which is bigger. All of these operations can easily be performed in polynomial time. To get the ball rolling, we remark that the functions a_m, b_m, $B(x)$, and z_j can all be computed in polynomial time, since their definitions only involve the types of operations just stated.

We first describe a polynomial-time algorithm for computing the number of almost-squares up to a given positive integer N. Let $M=\lceil\sqrt{N}\rceil$, so that $(M-1)^2<N\le M^2$. Lemma 7 tells us that the number of almost-squares up to M^2 is $B(M^2)$, and so we simply need to subtract from this the number of almost-squares larger than N but not exceeding M^2. This is easy to do by the characterization of almost-squares given in the Main Theorem. If $N>M(M-1)$, then we want to find the positive integer b such that $M^2-b^2\le N<M^2-(b-1)^2$, except that we want $b=0$ if $N=M^2$. In other words, we set $b=\lceil M^2-N\rceil$. Then, if $b\le b_M$, the number of almost-squares up to N is $B(M^2)-b$, while if $b>b_M$, the number of almost-squares up to N is $B(M^2)-b_M-1$.

In the other case, where $N\le M(M-1)$, we want to find the positive integer a such that $(M-a-1)(M+a)\le N<(M-a)(M+a-1)$, except that we want $a=0$ if $N=M(M-1)$. In other words, we set $a=\lceil\sqrt{(M-1/2)^2-N}+1/2\rceil$. Then, if $a\le a_M$, the number of almost-squares up to N is $B(M^2)-b_m-1-a$, while if $a>a_M$, the number of almost-squares up to N is $B((M-1)^2)$. This shows that $A(N)$ can be computed in polynomial time, which establishes part (c) of the corollary.

Suppose now that we want to compute the Nth almost-square. We compute in any way we like the first C almost-squares, where C is as in Lemma 8; this only takes a constant amount of time (it doesn't change as N grows) which certainly qualifies as polynomial time. If $N\le C$ then we are done, so assume that $N>C$. Let $M=\lceil z_N\rceil$, where z_N is defined as in Lemma 8(b), so that M is at least 3 by the definition of C. By Lemma 7,

$$A(M^2)=B(M^2)\ge B(z_N^2)>N\text{ and }A((M-2)^2)=B((M-2)^2)<B((z_N-1)^2)<N,$$

where the last inequality in each case follows from Lemma 8(b). Therefore the Nth almost-square lies between $(M-2)^2$ and M^2, and so is either in the $2M$-th flock or one of the preceding three flocks. If $0\le B(M^2)-N\le b_M$, then the Nth almost-square is in the $2M$-th flock, and by setting $b=B(M^2)-N$ we conclude that the Nth almost-square is M^2-b^2 and the dimensions of the optimal rectangle are $(M-b)\times(M+b)$. If $1+b_M\le B(M^2)-N\le b_M+1+a_M$, then the Nth almost-square is in the $(2M-1)$-st flock, and so on. This establishes part (d) of the corollary.

Finally, we can determine the greatest almost-square not exceeding N by computing $J = A(N)$ and then computing the Jth almost-square, both of which can be done in polynomial time by parts (c) and (d); and we can determine whether N is an almost-square simply by checking whether this result equals N. This establishes the corollary in its entirety. ■

7. Final filibuster

We have toured some very pretty and precise properties of the almost-squares, and there are surely other natural questions that can be asked about them, some of which have already been noted. When Grantham posed this problem, he recalled the common variation on the original calculus problem where the fence for one of the sides of the rectangle is more expensive for some reason (that side borders a road or something), and suggested the more general problem of finding the most cost-effective rectangle with integer side lengths and area at most N, where one of the sides must be fenced at a higher cost. This corresponds to replacing $s(n)$ with the more general function $s_\alpha(n) = \min_{d \mid n}(d + \alpha n/d)$, where α is some constant bigger than 1. While the elegance of the characterization of such "α-almost-squares" might not match that of Corollary 1, it seems reasonable to hope that an enumeration every bit as precise as the Main Theorem would be possible to establish.

How about generalizing this problem to higher dimensions? For example, given a positive integer N, find the dimensions of the rectangular box with integer side lengths and volume at most N whose volume-to-surface area ratio is largest among all such boxes. (It seems a little more natural to consider surface area rather than the sum of the box's length, width, and height, but who knows which problem has a more elegant solution?) Perhaps these "almost-cubes" have an attractive characterization analogous to Corollary 1; almost certainly a result like the Main Theorem, listing the almost-cubes in order, would be very complicated. And of course there is no reason to stop at dimension 3.

In another direction, intuitively it seems that numbers with many divisors are more likely to be almost-squares, and the author thought to test this theory with integers of the form $n!$. However, computations reveal that the only values of $n \le 500$ for which $n!$ is an almost-square are $n = 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 15$. Is it the case that these are the only factorial almost-squares? This seems like quite a hard question to resolve. Perhaps a better intuition about the almost-squares is that only those numbers that lie at the right distance from a number of the form m^2 or $m(m-1)$ are almost-squares —more an issue of good fortune than of having enough divisors.

Readers are welcome to try for themselves the *Mathematica* code used to calculate the functions related to almost-squares described in this paper; see `http://www.maa.org/pubs/mm_supplements/index.html`. With this code, for instance, one can verify that with 8,675,309 square feet of land at his disposal, it is most cost-effective for Farmer Ted to build a $2,919' \times 2,972'$ supercoop... speaking of which, we almost forgot to finish the Farmer Ted story:

> After learning the ways of the almost-squares, Farmer Ted went back to Builders Square, where the salesman viewed the arrival of his $\mathbb{R}$-rival with trepidation. But Farmer Ted reassured him, "Don't worry—I no longer think it's inane to measure fences in $\mathbb{N}$." From that day onward, the two developed a flourishing business relationship, as Farmer Ted became an integral customer of the store.

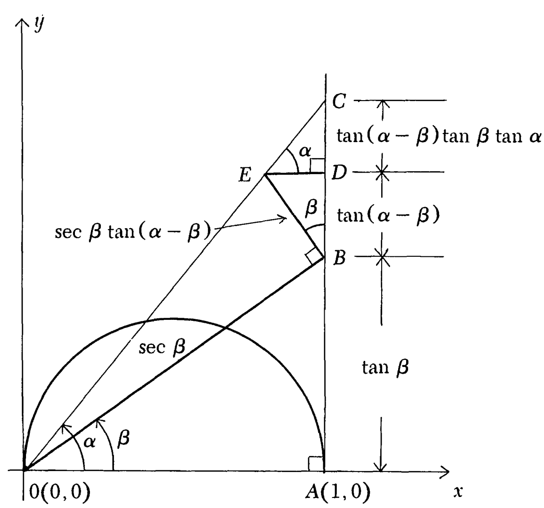
And that, according to this paper, is that.

Acknowledgment. The author would like to thank Andrew Granville and the anonymous referees for their valuable comments, which improved the presentation of this paper. The author would also like to acknowledge the support of National Science Foundation grant number DMS 9304580... the NSF may or may not wish to acknowledge this paper.

REFERENCES

1. J. H. Conway and R. K. Guy, *The Book of Numbers*, Copernicus, New York, NY, 1996.
2. A. K. Dewdney, *The (new) Turing Omnibus*, Computer Science Press, New York, NY, 1993.
3. S. W. Golomb, Problem E2491, *Amer. Math. Monthly* 82 (1975), 854–855.
4. C. Pomerance, A tale of two sieves, *Notices Amer. Math. Soc.* 43 (1996), 1473–1485.
5. N. J. A. Sloane, An on-line version of the encyclopedia of integer sequences, *Electron. J. Combin.* 1 (1994), feature 1 (electronic).
6. N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press Inc., San Diego, CA, 1995.

Proof Without Words: The Difference Identity for Tangents



$$AC - AB = BD + DC$$
$$\therefore\ \tan\alpha - \tan\beta = \tan(\alpha - \beta) + \tan\alpha\tan\beta\tan(\alpha - \beta)$$
$$\tan(\alpha - \beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha\tan\beta}$$

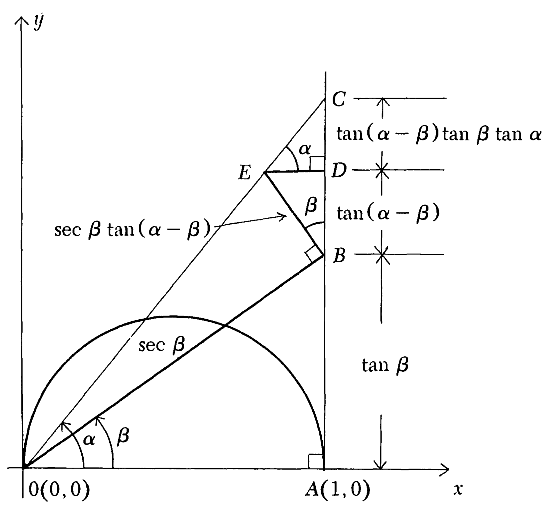
—Fukuzo Suzuki
Gunma College of Technology
Gunma 371-0845
Japan

Acknowledgment. The author would like to thank Andrew Granville and the anonymous referees for their valuable comments, which improved the presentation of this paper. The author would also like to acknowledge the support of National Science Foundation grant number DMS 9304580... the NSF may or may not wish to acknowledge this paper.

REFERENCES

1. J. H. Conway and R. K. Guy, *The Book of Numbers*, Copernicus, New York, NY, 1996.
2. A. K. Dewdney, *The (new) Turing Omnibus*, Computer Science Press, New York, NY, 1993.
3. S. W. Golomb, Problem E2491, *Amer. Math. Monthly* 82 (1975), 854–855.
4. C. Pomerance, A tale of two sieves, *Notices Amer. Math. Soc.* 43 (1996), 1473–1485.
5. N. J. A. Sloane, An on-line version of the encyclopedia of integer sequences, *Electron. J. Combin.* 1 (1994), feature 1 (electronic).
6. N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press Inc., San Diego, CA, 1995.

Proof Without Words: The Difference Identity for Tangents



$$AC - AB = BD + DC$$

$$\therefore\ \tan\alpha - \tan\beta = \tan(\alpha - \beta) + \tan\alpha\tan\beta\tan(\alpha - \beta)$$

$$\tan(\alpha - \beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha\tan\beta}$$

—Fukuzo Suzuki
Gunma College of Technology
Gunma 371-0845
Japan

Integer Antiprisms and Integer Octahedra

BLAKE E. PETERSON
Brigham Young University
Provo, UT 84602-6563

JAMES H. JORDAN
Washington State University
Pullman, WA 99164-3113

1. Introduction

Polyhedra Is it possible to find a rectangular box with integer dimensions and all diagonals of integer length as well? Such a box is referred to as a *perfect box* or a *perfect rational cuboid* and its existence or non-existence is an unsolved problem [**4**], [**7**]. An example of a box that is "almost" perfect has edges of length 44, 117, and 240, but one of the diagonals is irrational (see FIGURE 1). The goal of this paper is to introduce the reader to other types of three-dimensional figures with integer edges and diagonals. In particular, we will look at triangular antiprisms and a certain class of octahedra (see FIGURE 2). Accomplishing this goal, however, requires we first become familiar with prisms to understand antiprisms, with polyhedra to understand octahedra, and with all four of these terms to understand integer versions of any of them. For the purposes of this paper a *polyhedron* is a simple closed surface that can be expressed as the union of a finite set of polygonal regions, where any two polygons share at most one edge. The word polyhedron basically means many (*poly*) faces or planes (*hedra*).

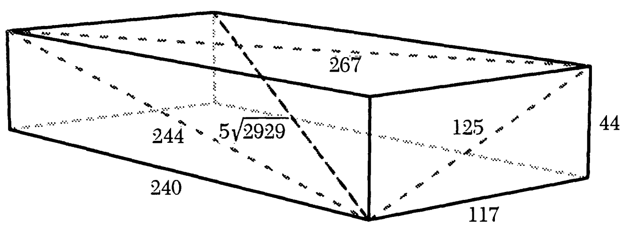


FIGURE 1
"Almost perfect box"

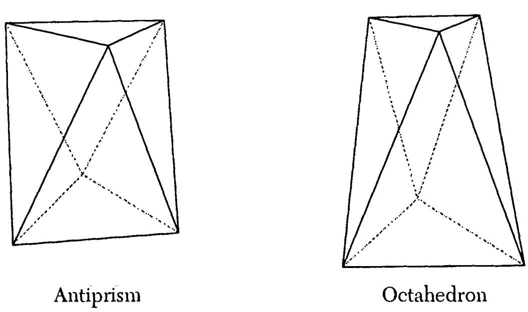
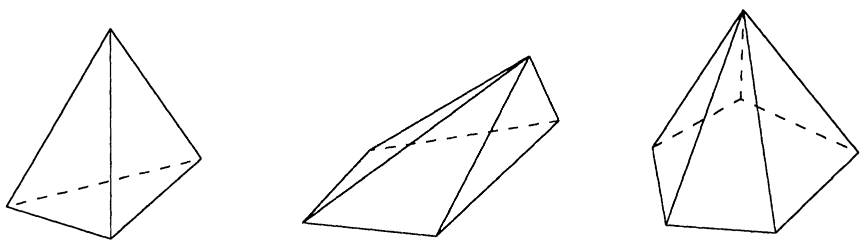


FIGURE 2

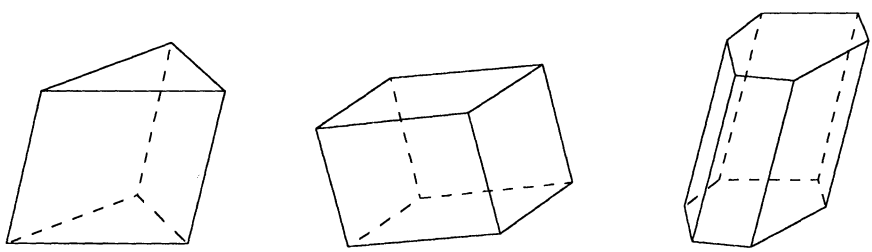
Polyhedra are usually classified either by their combinatorial type or by the number of faces. Two of the more common types of polyhedra are pyramids and prisms. A *pyramid* is constructed by joining each vertex of a polygon (the base) to a point not in the plane of the polygon. The pyramids in FIGURE 3 can be referred to as triangular, quadrilateral, and pentagonal pyramids depending on the base. They are sometimes described according to the total number of faces. Thus the triangular pyramid may also be called a tetrahedron because it has four faces. The quadrilateral pyramid is also called a pentahedron, and the pentagonal pyramid can be called a hexahedron because it has six faces.



Triangular pyramid (tetrahedron) | Quadrilateral pyramid (pentahedron) | Pentagonal pyramid (hexahedron)

FIGURE 3
Pyramid examples

The second common type of polyhedron is the *prism*, described as two congruent polygons (called the bases) in parallel planes joined in such a way that all of the other faces (the lateral faces) are parallelograms. As with pyramids, prisms may be described either according to the type of polygon that form the bases, or according to the total number of faces. FIGURE 4 gives examples of some prisms.

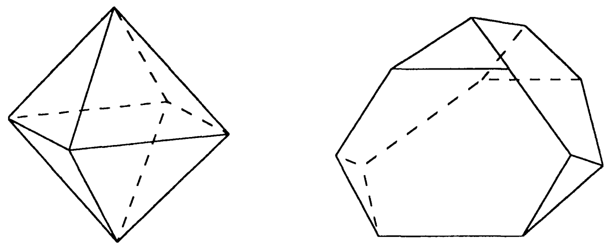


Triangular prism (pentahedron) | Rectangular prism (hexahedron) | Hexagonal prism (octahedron)

FIGURE 4
Prism examples

You may have noticed that the quadrilateral pyramid in FIGURE 3 and the triangular prism in FIGURE 4 are both pentahedra, and yet are quite different in their combinatorial type. The pyramid has 1 quadrilateral face and 4 triangular faces, while the prism has 3 quadrilateral faces and 2 triangular faces. In fact all pentahedra will have one or the other of these face combinations. Similarly, there are many different combinatorial types of octahedra **[9]**.

The hexagonal prism in FIGURE 4 is an octahedron but not the most common of the different octahedra. The octahedron in FIGURE 5a has eight equilateral triangular faces and is called a regular polyhedron or Platonic solid because all of the faces are the same regular polygon. The octahedron in FIGURE 5b consists of 4 equilateral triangles and 4 regular hexagons. Since all vertices consist of the same arrangement of polygons, FIGURE 5b is a semiregular polyhedron called a truncated tetrahedron. This polyhedron is also referred to as an Archimedean solid.



(a) Platonic solid

(b) Archimedean solid (truncated tetrahedron)

FIGURE 5
Some common octahedra

Integer polygons Before constructing examples of integer antiprisms and integer octahedra, we discuss integer polygons and integer polyhedra in general. An *integer polygon* is a convex polygon such that the distance between every pair of vertices is an integer. (Some authors also require that the polygon have a rational area. We do not require this.)

Finding integer triangles is as simple as finding three integers that satisfy the triangle inequality. An equilateral triangle with a side length of 1 is the simplest and smallest such example. Integer polygons become progressively more difficult to find as the number of sides increase. It is common, however, to use integer triangles as building blocks for quadrilaterals and quadrilaterals as building blocks for pentagons, and so forth. Some common examples of integer quadrilaterals, pentagons, and hexagons are shown in FIGURE 6. Another example of an integer pentagon comparable in size to the one shown in FIGURE 6 can be constructed by removing a vertex from the hexagon shown in FIGURE 6. (A more detailed description of these figures and their origins can be found in [11].) Notice that the quadrilateral in FIGURE 6 is an isosceles

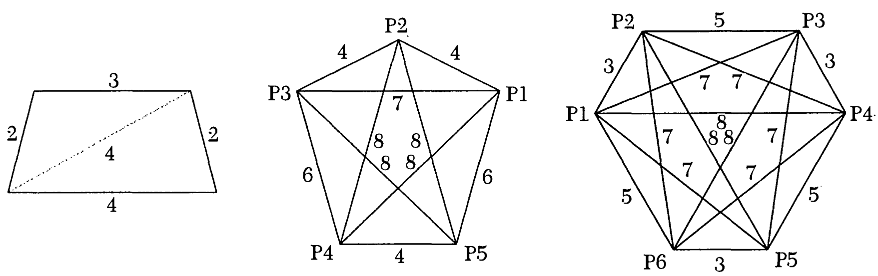
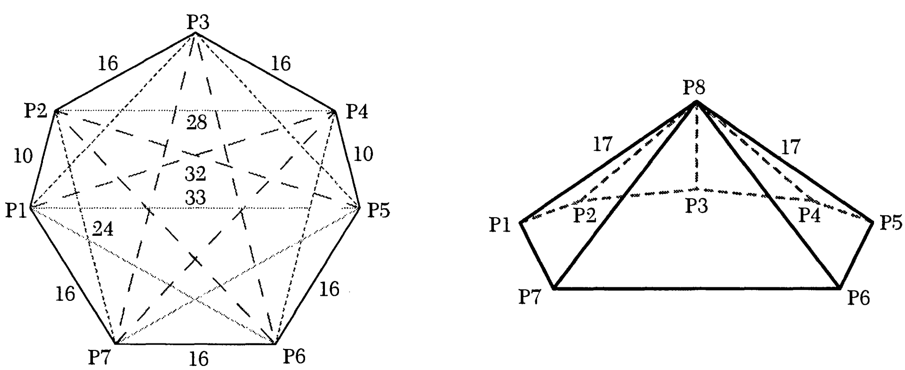


FIGURE 6
Small integer quadrilateral, pentagon, and hexagon

trapezoid, so it is cyclic, which means that it can be inscribed in a circle. This could be seen independently via Ptolemy's theorem: A quadrilateral is cyclic if and only if the sum of the products of the lengths of opposite sides is equal to the product of the lengths of the diagonals.

Integer polyhedra Analogous to an integer polygon, an *integer polyhedron* is a convex polyhedron such that the distance between every pair of vertices is an integer. Historically, rational surface area and rational volume were also sought **[1]**, but we will not require these restrictions. Since faces of an integer polyhedron are integer polygons, it is natural to use integer polygons as building blocks to construct integer polyhedra. Integer pyramids are simple to construct if a cyclic integer polygon can be found and used as the base according to Möller's method **[8]**, which follows.

The center of the circumscribed circle of a cyclic polygon is equidistant from all of the vertices; thus all of the points on the line through the center perpendicular to the plane of the polygon are equidistant from the vertices. It follows that there are infinitely many points on that line an integer distance from the vertices. Therefore, an integer octahedron may be constructed by using a cyclic integer heptagon as the base of a pyramid. Such a heptagon and the resulting octahedron are shown in FIGURE 7. In this case the radius of the circle circumscribing the heptagon is $\frac{64}{\sqrt{15}} \approx 16.52$ and thus the length of the lateral edges of the pyramid can be chosen to be 17.

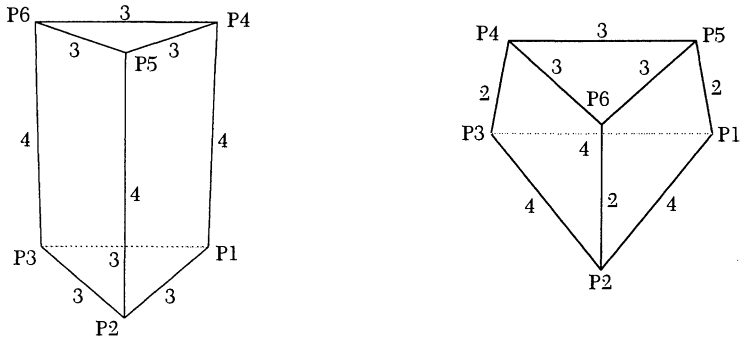


(a) Cyclic integer heptagon (b) Integer octahedron (pyramid)

FIGURE 7

Continuing with the building block idea, consider the integer rectangle with sides of length 3 and 4 and diagonals of length 5. Three copies of this rectangle can be placed together to form a triangular prism, as shown in FIGURE 8a. A similar method can be used with three copies of the integer quadrilateral from FIGURE 6 to form the integer pentahedron in FIGURE 8b. Although this second pentahedron looks like a prism, it is not because the two bases are not congruent.

Integer hexahedra are particularly interesting because of their relation to the perfect box problem mentioned earlier. In a perfect box, all six faces must be rectangles (see **[4]**, **[7]**). If we require only that the faces be quadrilaterals, then many examples can be found, and are described in **[10]**. The method used to construct these integer hexahedra with all quadrilateral faces is similar to the method used to construct the integer octahedra described later in this paper. The hexahedron is constructed by first placing the 2 copies of the rectangle in FIGURE 9 on top of one another and rotating the upper one 90° about the center of the rectangle. Next, raise



(a) Prism (b) Tetrahedron with one truncated vertex

FIGURE 8
Integer pentahedra constructed from integer quadrilaterals

the top rectangle vertically above the lower rectangle. The remaining four faces of this hexahedron, formed by joining vertices of these two rectangles, are isosceles trapezoids congruent to that shown in Figure 9. If this construction is begun with integer rectangles and integer isosceles trapezoids, then only the main diagonal needs to be verified for integer length. It has been proven that the smallest hexahedron constructed with this method is the one shown in Figure 10 **[10]**.

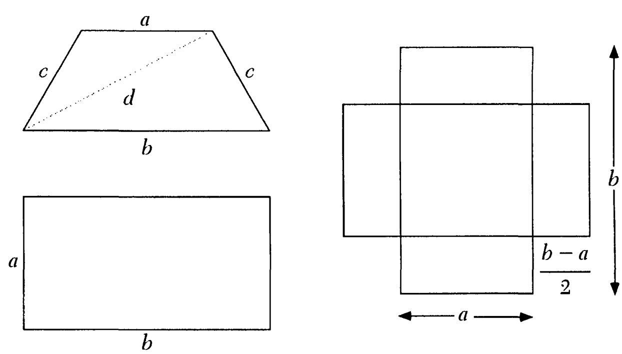


FIGURE 9
Pieces of a box

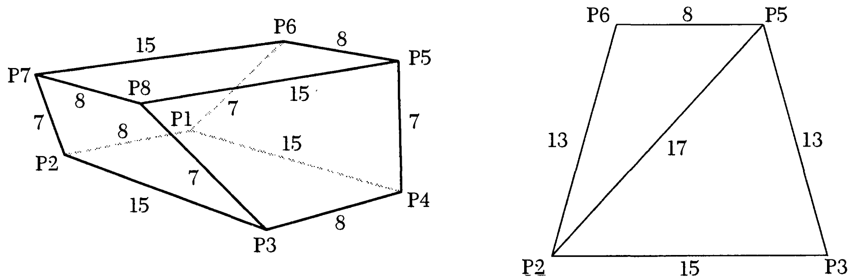


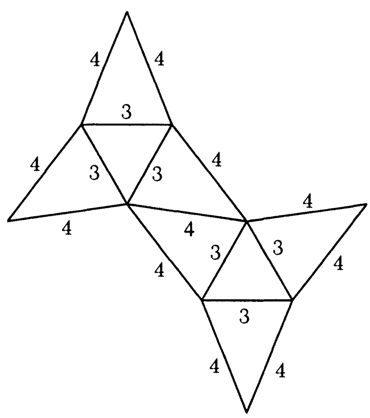
FIGURE 10
Hexahedron with six quadrilateral faces and its cross-section

The diagonals of the rectangular faces are of length 17 while the isosceles trapezoids that serve as the lateral faces have diagonals of length 13. The polygon with vertices P2, P3, P5, P6 is seen to be a cross section of the polyhedron. It is the isosceles trapezoid shown in FIGURE 10. Since it is an isosceles trapezoid, the lengths of the sides and diagonals satisfy Ptolemy's theorem; this yields 17 as the length of the main diagonal of the polyhedron.

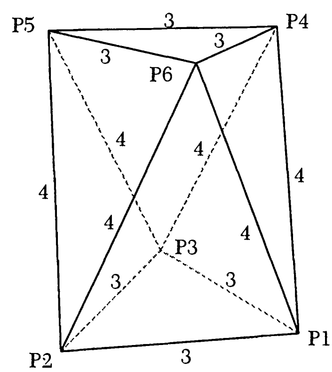
2. Integer antiprisms

An *antiprism* is defined as two congruent polygons (called bases) in parallel planes joined in such a way that all of the other faces (the lateral faces) are isosceles triangles. An antiprism resembles a prism in that its bases are congruent polygons in parallel planes; it differs from a prism in that the bases are not oriented so that the remaining faces are parallelograms. Therefore a prism with a triangular base is a six-vertex pentahedron with three rectangles as lateral faces, while an antiprism with a triangular base is a six-vertex octahedron with six triangular lateral faces. An antiprism with triangular bases can be constructed in a manner similar to that of the integer hexahedron described above. If two congruent equilateral triangles are placed on top of each other and one of the triangles is rotated 60° about its center and lifted, then the remaining 6 faces of this antiprism are congruent isosceles triangles. An *integer antiprism* is an antiprism with the integer restrictions described previously. In this section, we describe an infinite class of non-similar integer triangular antiprisms.

Consider the polygon in FIGURE 11a and the polyhedron in FIGURE 11b. The polyhedron can be folded from the polygon and is an integer antiprism; since it has eight faces, it is also an integer octahedron.



(a) Planar depiction



(b) Integer antiprism (octahedron)

FIGURE 11

This integer antiprism has four pairs of parallel congruent triangular faces. Six of the triangles are congruent 4-4-3 isosceles triangles, while the other two are equilateral 3-3-3 triangles. Each pair has the triangles oriented in opposite directions, i.e., rotated 180 degrees from each other. The two equilateral triangular faces are distance $\sqrt{13}$ apart and each of the other pairs are distance $3\sqrt{\frac{39}{55}}$ apart. All three interior diagonals have the same length, namely 5. This figure has longest distance 5, surface area $\frac{9}{2}(\sqrt{3}+\sqrt{55})$, volume $3\sqrt{39}$, and the sum of the twelve edges and three diagonals is 57. The integer distances are easily verified by using the following six points in $\mathbb{R}^3$ that could serve as vertices:

$$\left(0, \sqrt{3}, \frac{\sqrt{13}}{2}\right), \left(\pm\frac{3}{2}, -\frac{\sqrt{3}}{2}, \frac{\sqrt{13}}{2}\right), \left(0, -\sqrt{3}, \frac{-\sqrt{13}}{2}\right), \left(\pm\frac{3}{2}, \frac{\sqrt{3}}{2}, \frac{-\sqrt{13}}{2}\right).$$

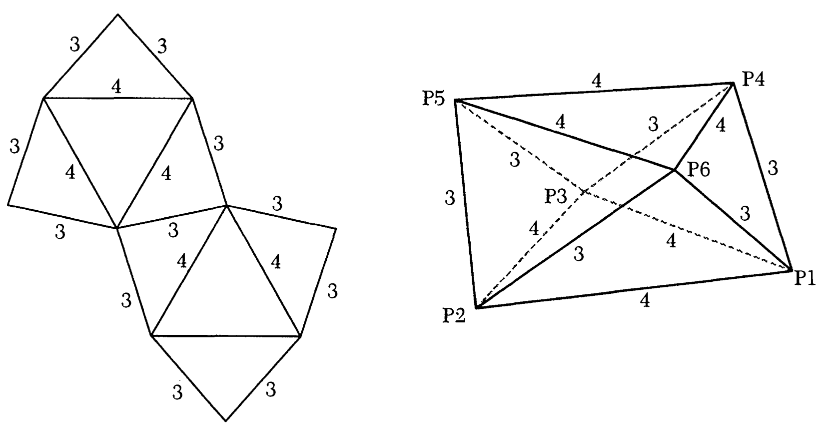
These points are on the sphere $x^2+y^2+z^2=\frac{25}{4}$.

The cross-section polygon formed by the vertices P2, P3, P4, P6 can be seen to be a rectangle and thus its legs and diagonal satisfy the Pythagorean theorem (a specific case of Ptolemy's theorem).

As a second example, consider the polygon in FIGURE 12a and the polyhedron in FIGURE 12b. The polyhedron is an integer antiprism that can be folded from the polygon. This integer antiprism has four pairs of parallel congruent triangular faces. Six of the triangles are congruent 3-3-4 isosceles triangles while the other two are equilateral 4-4-4 triangles. Each pair is comprised of triangles with opposite orientation. The two equilateral triangular faces are a distance $\sqrt{\frac{11}{3}}$ apart and the three interior diagonals all have the same length, namely 5. This figure has longest distance 5, surface area $8\sqrt{3}+12\sqrt{5}$, volume $\frac{16\sqrt{11}}{3}$, and the sum of the twelve edges and three diagonals is 57. Six points in $\mathbb{R}^3$ that could be vertices of this *integer antiprism* are

$$\left(0, \frac{4}{\sqrt{3}}, \frac{\sqrt{33}}{6}\right), \left(\pm 2, -\frac{2}{\sqrt{3}}, \frac{\sqrt{33}}{6}\right), \left(0, -\frac{4}{\sqrt{3}}, -\frac{\sqrt{33}}{6}\right), \left(\pm 2, \frac{2}{\sqrt{3}}, -\frac{\sqrt{33}}{6}\right).$$

As in the first example, these six points lie on the sphere described by $x^2+y^2+z^2 = \frac{25}{4}$, and the polygon formed by the vertices P2, P3, P4, P6 is a 3 by 4 rectangle. It is interesting to note that the 15 distances between vertices in both antiprisms are the same.



(a) Planar depiction (b) Integer antiprism (octahedron)

FIGURE 12

3. Infinitude of integer triangular antiprisms

Since the principal cross-sections (as in the previous examples) of integer triangular antiprisms are integer rectangles, the construction of a general integer triangular antiprism can begin with three lengths that satisfy the Pythagorean theorem. Consider a Pythagorean triangle with legs a and b and hypotenuse c, satisfying $c > \frac{2a}{\sqrt{3}}$. The vertices of an integer antiprism that is also an integer octahedron are described by the following six points, which lie on a sphere of radius $\frac{c}{2}$:

$$\left(0, \frac{a}{\sqrt{3}}, \frac{\sqrt{c^2 - \frac{4a^2}{3}}}{2}\right), \left(\frac{\pm a}{2}, \frac{-a}{2\sqrt{3}}, \frac{\sqrt{c^2 - \frac{4a^2}{3}}}{2}\right),$$

$$\left(0, \frac{-a}{\sqrt{3}}, \frac{-\sqrt{c^2 - \frac{4a^2}{3}}}{2}\right), \left(\frac{\pm a}{2}, \frac{a}{2\sqrt{3}}, \frac{-\sqrt{c^2 - \frac{4a^2}{3}}}{2}\right).$$

Since there are infinitely many non-similar Pythagorean triangles, and each triangle produces one, or possibly two (when the two legs can be interchanged and the inequality maintained) integer antiprisms, there are infinitely many non-similar integer antiprisms. The examples in FIGURE 11b and FIGURE 12b are both generated from the 3-4-5 right triangle.

The integer antiprism described by the six vertices above has three interior diagonals, all of length c, which also is the longest distance. The figure has surface area $\frac{\sqrt{3}a}{2}(\sqrt{12b^2 - 3a^2} + \mathrm{a})$, volume $\frac{a^2}{3}\sqrt{3b^2 - a^2}$, and the sum of the twelve edges and three diagonals is $6a + 6b + 3c$. This completely describes all *integer triangular antiprisms*.

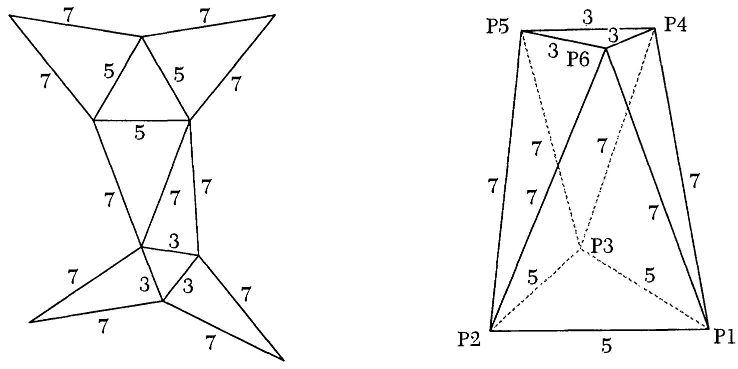
4. A generalization

A different class of *integer octahedra* can be constructed by placing two noncongruent equilateral triangles on top of each other with their centers concurrent, rotating one of the triangles 60° about the common center, and lifting. (The construction is similar to that of the antiprisms, but these polyhedra are not technically antiprisms, because their bases are not congruent.)

One small example of an integer octahedron that is not an integer antiprism is shown in FIGURE 13b. The planar depiction of this octahedron is in FIGURE 13a. The polygon can be folded into the *integer octahedron* with the side lengths of the two equilateral triangles being 5 and 3. All six lateral edges have length 7, the six lateral faces are 7-7-3 and 7-7-5 isosceles triangles, and the three interior diagonals have equal length 8. The solid has volume $\frac{128}{3}\sqrt{2}$, surface area $\frac{17}{2}\sqrt{3} + \frac{45}{4}\sqrt{19} + \frac{9}{4}\sqrt{187}$, and the sum of the lengths of the twelve edges and three diagonals is 90. The vertices could be the following six points in $\mathbb{R}^3$:

$$\left(0, \frac{5}{\sqrt{3}}, \frac{4\sqrt{6}}{3}\right), \left(\pm\frac{5}{2}, -\frac{5}{2\sqrt{3}}, \frac{4\sqrt{6}}{3}\right), \left(0, -\sqrt{3}, -\frac{4\sqrt{6}}{3}\right), \left(\pm\frac{3}{2}, \frac{\sqrt{3}}{2}, -\frac{4\sqrt{6}}{3}\right).$$

These six vertices lie on the sphere $x^2 + y^2 + (z - \frac{1}{\sqrt{6}})^2 = \frac{33}{2}$.



(a) Planar depiction (b) Integer octahedron

FIGURE 13

As with integer hexahedra and integer antiprisms, it is helpful to consider the principal cross-section polygon formed by the vertices P2, P3, P4, P6, which can be seen to be an isosceles trapezoid (and thus satisfies Ptolemy's theorem).

5. Infinitude of integer octahedra that aren't integer antiprisms

Since the principal cross sections of the integer octahedra considered in the last section satisfy Ptolemy's theorem, we can begin to construct a general integer octahedron by finding integers that satisfy this theorem. In other words, we find lengths a and b to serve as bases of an isosceles trapezoid, length c to serve as the slant side of the trapezoid and length d to serve as a diagonal, where a, b, c, and d are integers satisfying the equation $d^2 - c^2 = ab$. Such integers can be found as follows. Consider positive integers a and b (both larger than 2) whose product ab is not of the form $4k+2$ and is larger than 12. Then there exist integers d and c such that $d > \frac{a+b}{3}$ and $d^2 - ab = c^2$ (Ptolemy's Theorem), for some integer c. If ab is odd then d and c could be $\frac{ab+1}{2}$ and $\frac{ab-1}{2}$ or if ab is a multiple of four then d and c could be $\frac{ab}{4} \pm 1$. If h is the distance between the parallel planes containing the bases, then

$$h = \sqrt{d^2 - \frac{(a+b)^2}{3}}\,.$$

The surface area is $\frac{\sqrt{3}}{4}(a^2+b^2) + \frac{3a}{4}\sqrt{4c^2-a^2} + \frac{3b}{4}\sqrt{4c^2-b^2}$, the volume is $\frac{h(a+b)^2}{4\sqrt{3}}$, and the sum of the lengths of the twelve edges and three diagonals is $3a+3b+6c+3d$. Six points in three-dimensional Euclidean space that could be vertices of such an integer octahedron that is not an integer antiprism are:

$$\left(0, \frac{a}{\sqrt{3}}, \frac{h}{2}\right), \left(\pm\frac{a}{2}, -\frac{a}{2\sqrt{3}}, \frac{h}{2}\right), \left(0, \frac{-b}{\sqrt{3}}, -\frac{h}{2}\right), \left(\pm\frac{b}{2}\frac{b}{2\sqrt{3}}, -\frac{h}{2}\right).$$

These six points lie on the sphere

$$x^2 + y^2 + \left(z + \left(\frac{b^2-a^2}{6h}\right)\right)^2 = \frac{a^2b^2 - 2abc^2 - 3c^4}{4(a^2 - ab + b^2 - 3c^2)}.$$

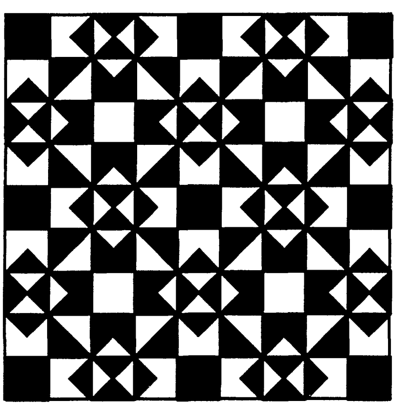
Since this can be done for an infinitude of distinct pairs of positive integers, there exists a subset of the pairs giving us an infinitude of non-similar integer octahedra that are not integer antiprisms. Indeed for $a = 3$ alone the integer b could be selected from the set $\{5, 7, 8, 9, 11, 12, 13, 15, \ldots\}$. All selections would yield non-similar examples.

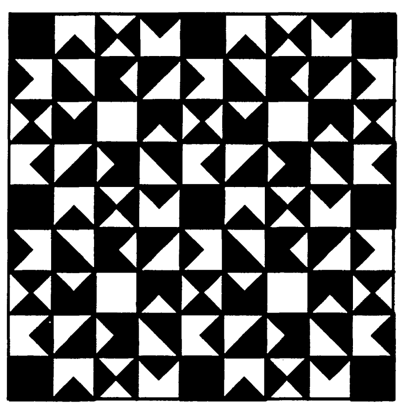
REFERENCES

1. Leonard E. Dickson, *History of the Theory of Numbers*, Vol II, Chelsea, New York, NY, 1971.
2. Richard K. Guy, *Unsolved Problems in Number Theory*, Springer-Verlag, New York, NY, 1985.
3. Heiko Harborth and Lothar Piepmeyer, Point Sets with Small Integral Distances, *Applied Geometry and Discrete Mathematics, The Victor Klee Festschrift*, Gritzmann, P. and Sturmfels, B. (eds), DIMACS 4 (1991), 319–324.
4. Victor Klee and Stan Wagon, *Old and New Unsolved Problems in Plane Geometry and Number Theory*, MAA Dolciani Mathematical Exposition No. 11, Washington, D.C., 1991.
5. Ivan Korec, Nonexistence of a small perfect rational cuboid, *Acta Mathematica Universitatis Comenianae* 42–43 (1983), 73–86.
6. Ivan Korec, Nonexistence of a small perfect rational cuboid, II, *Acta Mathematica Universitatis Comenianae* 44–45 (1984), 39–48.
7. John Leech, The rational cuboid revisited, *American Mathematical Monthly* 84 (1977) 518–533.
8. Meinhard Möller, Dissertation (unpublished), Technischen Universität Carolo-Wilhelmina zu Brauschweig 1990.
9. Jean J. Pedersen, The challenge of classifying polyhedra, *Two-Year College Mathematics Journal* 11 (1980), 162–173.
10. Blake E. Peterson and James H. Jordan, Integer hexahedra equivalent to a perfect box, *American Mathematical Monthly* 102 (1995), 41–45.
11. Blake E. Peterson and James H. Jordan, Integer geometry: some examples and constructions, *The Mathematical Gazette* Vol. 81, No. 490 (1997), 18–28.

More Repeating Arrays

As for the cover illustration, all 16 combinations of black and white triangles inside a square are arranged in a regularly repeating 4 × 4 array. The graphics were created by Barry Cipra.





Integral Triangles

RAYMOND A. BEAUREGARD
E. R. SURYANARAYAN
University of Rhode Island
Kingston, RI 02881

1. Introduction

A perusal of Dickson's monumental text [4] makes clear that integer-sided triangles have always been a fascination. Certain classic cases arise depending on the imposition of an additional condition. For example, if the area is required to be an integer we have a class of Heron triangles. If we require instead that one angle be an integer number of degrees we obtain a type of triangle that is the subject of this paper.

If a triangle has integer sides a, b, c with the angle opposite side c measuring an integer number of degrees, then the cosine of the angle must be rational. It is known (see, e.g., [8], p. 41) that the only such cosine values are 0, $\pm 1/2$, ± 1. Thus there are only three types of such triangles. In the first case we have right triangles described by Pythagorean triples (PTs); these have been written about extensively (see [3] and [5] and their references). In the remaining two cases (cosine values of $\pm 1/2$) we have triangles whose sides a, b, c satisfy the Diophantine equation $a^2 \pm ab + b^2 = c^2$, with corresponding angle opposite side c of 60° or 120° depending on the choice of algebraic sign. These cases were described separately in [6] and [9]. It is our purpose to give a unified presentation and explore the striking analogy with the nature of PTs.

We refer to triangles with integer sides and an integer degree-angle as *integral triangles*. We call a triple (a, b, c) of integers satisfying

$$a^2 + ab + b^2 = c^2 \tag{1}$$

an ET. We place no direct restriction on a and b but require c to be positive.

FIGURE 1 suggests how every integral triangle that is not a right triangle is bred from an equilateral triangle (hence the name ET). Notice that if (a, b, c) is a solution of (1)

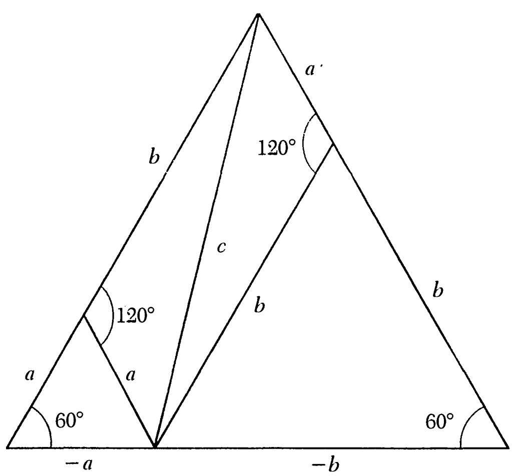


FIGURE 1
Triangles corresponding to ETs.

then so are $(a+b, -b, c)$ and $(a+b, -a, c)$. We have used negative values to label some of the sides of triangles in FIGURE 1 to illustrate how an ET with a negative middle component corresponds to the 60° case.

The analogy of ETs with PTs includes algebraic structure (described in Sections 2 and 3) as well as parametric representation (described in Section 4). We also determine which integers can serve as the components of an ET. In the final section we describe the peculiar relationship of PTs to ETs.

2. The semigroup of triples

The set of PTs has a nice algebraic structure, versions of which are described in **[3]** and **[5]**. Likewise, the set

$$\mathscr{E} = \{(a,b,c) | a,b,c \in \mathbb{Z},\ c > 0,\ a^2 + ab + b^2 = c^2\}$$

of ETs is endowed with an algebraic structure. The key step in recognizing it is to represent (1) as the determinantal equation

$$\det \begin{bmatrix} a & b \\ -b & a+b \end{bmatrix} = c^2.$$

The fact that the set of integer matrices of the form

$$\begin{bmatrix} a & b \\ -b & a+b \end{bmatrix} \tag{2}$$

is closed under multiplication, together with the multiplicative property of determinants tell us that the operation defined by

$$(a_1, b_1, c_1) * (a_2, b_2, c_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1 + b_1 b_2, c_1 c_2)$$

is an associative binary operation on $\mathscr{E}$. In fact $(\mathscr{E}, *)$ is a commutative semigroup with identity $E = (1, 0, 1)$; it is isomorphic to the semigroup of integer matrices (2) having a square determinant (i.e., satisfying (1)).

An ET (a, b, c) is said to be *primitive* if the components a, b, and c have no proper common divisor. In view of (1) it is necessary (and sufficient) that any pair of components be relatively prime. Each ET (a, b, c) can be written as $k(a_1, b_1, c_1) = (ka_1, kb_1, kc_1)$ where (a_1, b_1, c_1) is primitive and k is a positive integer (if $ab \neq 0$, $k = \gcd(a, b, c)$).

If $A = (a, b, c)$, we consider $A' = (a+b, -b, c)$ to be the *quasi-inverse* of A, since $A * A' = c^2(1, 0, 1)$. FIGURE 1 illustrates (a, b, c) which corresponds to the 120° triangle (extending to the upper right-hand portion of the largest equilateral triangle) and its quasi-inverse corresponding to the 60° triangle containing it. The left-hand side of the figure represents a similar situation for (b, a, c). FIGURE 2 illustrates the case where $(-a, a+b, c)' = (b, -a-b, c)$ and $(-b, a+b, c)' = (a, -a-b, c)$ both correspond to 60° triangles (the 60° angles at the base). The following proposition, which gives some additional properties of quasi-inverses, indicates that semigroup $\mathscr{E}$ is cancellative.

PROPOSITION 1. *Let A, B, and C be ETs.*

1. $(A * B)' = A' * B'$,
2. $A'' = A$,
3. *If* $A * B = A * C$ *then* $B = C$.

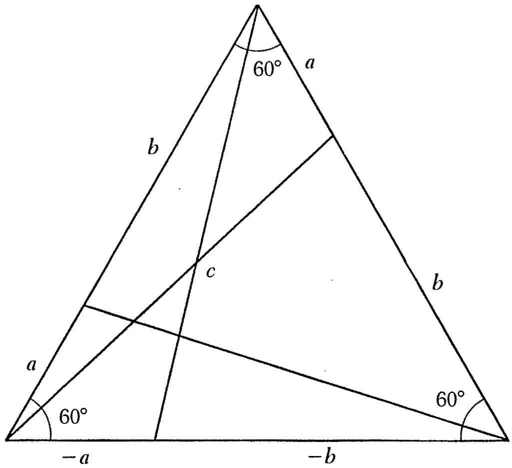


FIGURE 2
60° ETs with their 60° quasi-inverses.

We have seen that $\mathscr{E}$ is isomorphic to the semigroup of integer matrices (2) with a square determinant. We can obtain more information about $\mathscr{E}$ by computing the eigenvalues of these matrices. They are

$$\frac{2a+b(1\pm\sqrt{-3})}{2} \tag{3}$$

and have modulus equal to c^2. Diagonalization gives a similarity transformation showing that $\mathscr{E}$ is isomorphic to the semigroup of diagonal matrices with diagonal entries of the form (3), where $a,b\in\mathbb{Z}$, and having modulus a square. Equivalently the mapping

$$(a,b,c)\to\frac{2a+b(1+\sqrt{-3})}{2} \tag{4}$$

is an isomorphism of $\mathscr{E}$ with the multiplicative semigroup of algebraic integers of $\mathbb{Q}(\sqrt{-3})$ having modulus a square (see **[1]**, p. 366).

3. The group of triples

Following our analogy of ETs with the development in **[3]** we seek to convert the semigroup $\mathscr{E}$ into a group. To this end we define a relation on $\mathscr{E}$ by $A\sim B$ if there exist positive integers n and k such that $nA=kB$. Then $\sim$ is an equivalence relation on $\mathscr{E}$. Indeed $\sim$ is a congruence on $\mathscr{E}$ in the sense that if $A\sim B$ and $C\sim D$ then $A*C\sim B*D$.

We denote by $A^{\sim}$ the equivalence class containing A and by $\mathscr{E}^{\sim}$ the set of all such equivalence classes. The induced operation $A^{\sim}*B^{\sim}=(A*B)^{\sim}$ is well-defined because $\sim$ is a congruence. Thus $\mathscr{E}^{\sim}$ is an abelian group, with identity element $(1,0,1)^{\sim}$; the inverse of $(a,b,c)^{\sim}$ is $(a+b,\,-b,c)^{\sim}$.

The equivalence relation $\sim$ converts the semigroup isomorphism of $\mathscr{E}$ with the corresponding set of the matrices (2) into the group isomorphism

$$(a,b,c)^{\sim}\to\begin{bmatrix} a/c & b/c \\ -b/c & (a+b)/c\end{bmatrix} \tag{5}$$

of $\mathscr{E}^{\sim}$ with the set of matrices in (5), a subgroup of the group of all 2×2 unimodular matrices (that is, those having determinant 1). Furthermore, (4) corresponds to the group isomorphism

$$(a,b,c)^{\sim} \to \frac{2a+b(1+\sqrt{-3})}{2c} \tag{6}$$

of $\mathscr{E}^{\sim}$ with this subgroup of modulus-one complex numbers.

Multiplying (a,b,c) by $(-1,0,1)$ changes the algebraic signs of a and b. Thus the group element $(-1,0,1)^{\sim}$ is its own inverse, and has order 2; similarly, $(0,-1,1)^{\sim}$ has order 3, and the product of these two elements is $(0,1,1)^{\sim}$, which has order 6. These three elements together with their inverses and the identity element comprise the torsion subgroup of $\mathscr{E}^{\sim}$; they correspond to the six sixth roots of unity.

Why are there no other elements of finite order? If $(a,b,c)^{\sim}$ has order n then the complex number in (6) is a primitive nth root of unity. Assuming that its argument is $2\pi/n$, we have

$$\cos\left(\frac{2\pi}{n}\right) = \frac{2a+b}{2c}.$$

But for rational multiples of 2π the only rational values of cosine are 0, $\pm 1/2$, or ± 1. Thus we are led back to one of the cases just described.

4. Parametric forms

It is not difficult to verify that any triple of the form

$$(s^2-t^2,\, t^2-2st,\, s^2-st+t^2), \tag{7}$$

where s and t are integers, is an ET. Conversely, every primitive ET has this form or this form divided by 3, for relatively prime integers s and t. (This parametrization is described in **[10]**, p. 171.) Here is a simple way to find the parameters for the ET (a,b,c) (see **[3]** for an analogous approach to computing the parameters of a PT):

For $a \neq 0$, write $(c-b)/a$ in reduced form; the numerator is s and the denominator is $s-t$. Then (a,b,c) has the form (7) *unless $s+t$ is a multiple of* 3, *in which case each component of* (7) *must be divided by* 3.

To see why this works, let $c-b=sd$ and $a=(s-t)d$ where d is the greatest common divisor of $c-b$ and a. Then $(c-b)/a=s/(s-t)$ where s and t are relatively prime. With a bit of pencil pushing we find that

$$(a,b,c) = \frac{d}{s+t}(s^2-t^2,\, t^2-2st,\, s^2-st+t^2). \tag{8}$$

To prove that $d/(s+t)$ is either 1 or $1/3$, let p be a prime divisor of $s+t$. Since s and t are relatively prime, p divides neither s nor t. Suppose $p \neq 3$. Looking at the middle component in (8) we see that p divides d or t or $t-2s$; in the latter case p divides $t-2s-(s+t)=-3s$. Thus p divides d. It follows that the fraction $d/(s+t)$ in (8) is an integer if $s+t$ is not a multiple of 3, and so equals 1 since the ET (a,b,c) is primitive. Now suppose $p=3$. Since 3 divides $t-2s=(t+s)-3s$, the first two components (and hence the third component) of (8) are multiples of 3. However, if 9 divides $s+t$, say $s+t=9k$, then $t-2s=3(3k-s)$; since 3 does not divide $3k-s$, 9 cannot divide $t(t-2s)$, the middle component in (8). We conclude that when $s+t$ is a multiple of 3 the components of (8), when divided by 3, are integers and $d/(s+t)=1/3$.

For example, to find the parameters for $(8, -3, 7)$ we compute $(7+3)/8 = 5/4$, so $s = 5$ and $t = 1$, and the division of (7) by 3 is necessary.

Alternative parameters for (a, b, c) are described in **[6]** and **[9]**, namely

$$(a,b,c) = \left(2xy - x^2,\, y^2 - 2xy,\, x^2 - xy + y^2\right).$$

We can find x and y in a manner similar to that described above when $a \neq 0$: compute $(c+b)/a$ in lowest terms; the numerator is $y - x$ and the denominator is x; division by 3 is necessary when $x + y \in 3\mathbb{Z}$. These are essentially our parameters for $(-a-b, b, c)$.

5. The nature of components

Much is known about which integers can comprise a PT. For instance, if (a, b, c) is a PT then a, b, and c include a multiple of 3, a multiple of 4, and a multiple of 5. Analogous properties hold for ETs which relate to the components of A or A'. A version of the following result may be found in **[6]**.

PROPOSITION 2. *If (a, b, c) is a nontrivial ET $(c > 1)$ then one of a, b, $a+b$ is a multiple of* 8, *one is a multiple of* 3, *and one is a multiple of* 5. *Furthermore, one of a, b, $a+b$, c is a multiple of* 7.

Which integers can actually serve as components of an ET? We answer this question first for third components. Let (a, b, c) be a primitive ET. By Eq. (1), c is odd. Now $(-b, a+b, c)$ and $(a+b, -a, c)$ are also primitive ETs and one of these three has an even middle component. Thus we may work with the ET (a, b, c) and assume that b is even. Since

$$c^2 = \left(a + \frac{b}{2}\right)^2 + 3\left(\frac{b}{2}\right)^2$$

we see that -3 is a quadratic residue of any prime p that divides c. But -3 is a quadratic residue of a prime p if and only if $p \equiv 1 \pmod 6$ (see, e.g., **[1]**, p. 131).

Conversely, if p is prime and $p \equiv 1 \pmod 6$ then -3 is a quadratic residue of p, so there are integers n and m such that $n^2 + 3 = mp$. Thus p divides $(n + \sqrt{-3})(n - \sqrt{-3})$ in the unique factorization domain $\mathscr{A}$ of algebraic integers of $\mathbb{Q}(\sqrt{-3})$ (**[1]**, p. 383). Clearly, p does not divide either of these factors, so p is not a prime in the ring $\mathscr{A}$. Thus p is reducible and may be factored properly as $p = \alpha\beta$ for $\alpha, \beta \in \mathscr{A}$. Computing norms we see that $p^2 = N(\alpha)N(\beta)$ which shows that $p = N(\alpha)$. That is, there exist nonzero integers a and b like those in (4) such that $\alpha = a + b(1 + \sqrt{-3})/2$ and $p = a^2 + ab + b^2$. Thus $(a^2 - b^2, b^2 + 2ab, p)$ is a primitive ET of the form (7) with parameters $s = a$ and $t = -b$. By composing such ETs we see that every third component (greater than 1) of an ET is a product of primes of the form $6k + 1$. Since the case $c = 1$ is trivial we have established the following result (an analogous result holds for PTs).

PROPOSITION 3. *Let c be a positive integer. There exist integers a and b such that (a, b, c) is a primitive ET if and only if $c = 1$ or c is a product of primes $p \equiv 1$ (mod* 6).

Every odd integer and every multiple of 4 is the first component of a primitive PT. We shall show that the same is true for ETs, except that an even integer must be a multiple of 8. Since the first two components are interchangeable we focus on first components and assume (without loss of generality) that they are positive. Expressing (a, b, c) in parametric form (7) we have $a = pq$ where $p = s + t$, $q = s - t$ or *vice versa*, and $\gcd(s, t) = 1$. Thus (7) reduces to the following possible triples:

$$\left(pq, \frac{p^2 - 2pq - 3q^2}{4}, \frac{p^2 + 3q^2}{4}\right); \tag{9}$$

$$\left(pq, \frac{3p^2 - 2pq - q^2}{4}, \frac{3p^2 + q^2}{4}\right). \tag{10}$$

Now pq is either odd or even; we handle each case separately.

Case (i): pq is odd. Here p and q are both odd, and the components of (9) and (10) are integers. These ETs are both primitive if $pq \notin 3\mathbb{Z}$, but only one is primitive if $pq \in 3\mathbb{Z}$. Note that if pq is a power of a prime other than 3 or 2 then we must take $p = 1$ or $q = 1$ to get primitive ETs. This gives exactly four primitive ETs; for powers of 3 there are just two. For example, $(27, 533, 547)$ and $(27, -560, 547)$ are the only primitive ETs with first component 27.

Case (ii): pq is even. Both p and q are even since $p = q + 2t$; writing $p = 2p_1$ and $q = 2q_1$, (9) and (10) reduce to

$$\left(4p_1q_1, p_1^2 - 2p_1q_1 - 3q_1^2, p_1^2 + 3q_1^2\right); \tag{11}$$

$$\left(4p_1q_1, 3p_1^2 - 2p_1q_1 - q_1^2, 3p_1^2 + q_1^2\right). \tag{12}$$

For primitivity exactly one of p_1 or q_1 must be even. In this case both (11) and (12) are primitive if $p_1q_1 \notin 3Z$ but just one is primitive if $p_1q_1 \in 3Z$. If $4p_1q_1$ is a power of 2 we must take $q_1 = 1$ or $p_1 = 1$, resulting in exactly four primitive ETs. For example, if $4p_1q_1 = 8$, we obtain $(8, -3, 7)$, $(8, 7, 13)$, $(8, -15, 13)$, and $(8, -5, 7)$.

We summarize our work and then give a computational illustration.

PROPOSITION 4. *Every positive odd integer is the first component of a primitive ET; for even integers only multiples of* 8 *qualify. For each odd prime* $p \neq 3$, p^i $(i \geq 1)$ *is the first component of exactly four primitive ETs; there are also four primitive ETs with first component* 2^{i+2}, *but only two primitive ETs with first component* 3^i.

6. Illustration

An ET is called *positive* if each of its components is positive (corresponding to a 120° triangle), and *negative* if one component is negative (corresponding to a 60° triangle). We will determine all positive and negative ETs with a given first component.

Consider $385 = 5 \times 7 \times 11 = pq$. Since the middle components (9) and (10) factor as $(p - 3q)(p + q)/4$, and $(3p + q)(p - q)/4$, we find eight positive and primitive ETs with first components 385. These are listed in Table 1; four are obtained from (9) with $p > 3q$ and four from (10) with $p > q$. Each of these ETs (a, b, c) gives rise a corresponding negative ET $(a, -a - b, c)$, and this completes the list of all primitive ETs with first component 385.

In general, the number of positive and primitive ETs (for a given a) can be less than the number of negative and primitive ETs. The example with $a = 8$ above illustrates this. For another example, $(21, 320, 331)$ is the only positive primitive ET starting with 21, the negative ones are $(21, -341, 331)$, $(21, -5, 19)$, and $(21, -16, 19)$.

TABLE 1. Positive ETs with first component 385

p	q	(a,b,c) using (9), $p>3q$	(a,b,c) using (10), $p>q$
35	11	$(385,23,397)$	$(385,696,949)$
55	7	$(385,527,793)$	$(385,2064,228)$
77	5	$(385,1271,1501)$	$(385,4248,4453)$
385	1	$(385,36863,37057)$	$(385,110976,111169)$

7. Pythagorean ETs

Can an ET (a,b,c) also be a PT? Clearly the answer is no except in the trivial case that $ab=0$. But the question of when an ET (a,b,c) also corresponds to a PT (a,b,d) arises naturally as follows.

Consider the rectangle in FIGURE 3 with integer sides and diagonals. Can we squash the rectangle into a parallelogram by rotating a side of length a toward a side of length b through an integer number of degrees so that the diagonals of the parallelogram remain integers? By the law of cosines, we would need $a^2 \pm ab + b^2$ to be perfect squares, but then their product $a^4+a^2b^2+b^4$ would also be a square and the latter is known to be impossible ([**7**], p. 19). Thus only one of the diagonals of the parallelogram can remain an integer and the angle of rotation must be 30°. (See the paragraph preceding Section 4.)

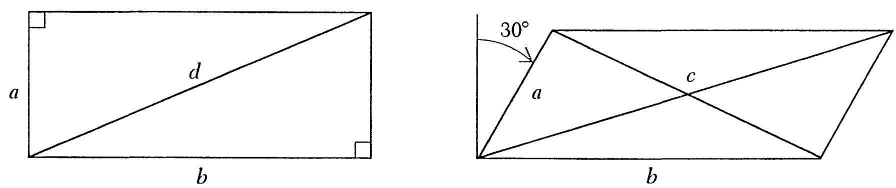


FIGURE 3
Squashing a rectangle while preserving integrity.

Finding those values of a and b that form the first two components of both a PT and an ET is an interesting number-theoretic exercise. A short computer run produces just two (up to similar triangles) primitive (and nontrivial) ETs that fill the bill. These are given in Table 2; one is negative and the other is positive. We refer to such ETs as PETs (Pythagorean ETs).

TABLE 2. Two nontrivial PETs

PET	PT
$(8,-15,13)$	$(8,-15,17)$
$(1768,2415,3637)$	$(1768,2415,2993)$

The simultaneous solutions of the Diophantine equations

$$a^2+b^2=d^2, \quad a^2+ab+b^2=c^2$$

were described by Aubry in 1911 [**2**]. The solutions are found as follows. If a is even and (a,b,c) is a (primitive) PET with (a,b,d) the corresponding PT, let $a=2nm$ and $b=n^2-m^2$ (these are the PT parameters and may be found by writing

$(d+b)/a = n/m$ in lowest terms [3]). If $(d \pm c)/b = u/v$, reduced to lowest terms, then the values

$$N = au^2 + bv^2, \quad M = bu^2 + av^2$$

give rise to two new PETs (A, B, C), where $A = 2NM$ and $B = N^2 - M^2$. Thus we obtain an infinite set of (primitive) PETs. For example, the first PET in Table 2 gives rise to the PETs

$$(1768, 2415, 3637) \quad \text{and} \quad (10130640, -8109409, 9286489).$$

The second PET in Table 2 gives rise to the PETs

$$(498993199440, 136318711969, 579309170089),$$

and

$$(422390893185635192, -101477031226926255, 381901401745295077).$$

The components grow quickly in magnitude. Computing PETs in this way we found that the $100th$ PET has components each with more than 5928 digits!

Acknowledgment. We thank our colleague Professor Dean Clark for drawing the figures.

REFERENCES

1. A. Adler and J. Coury, *The Theory of Numbers*, Jones & Bartlett Pub. Co., Boston, MA, 1995.
2. L. Aubry, Untitled, *L'intermediaire des mathematiciens* 17 (1910), 283–285.
3. R. A. Beauregard and E. R. Suryanarayan, Pythagorean triples: the hyperbolic view, *College Math. J.* 27 (1996), 170–181.
4. L. E. Dickson, *History of Theory of Numbers*, Vol. II, Chelsea, New York, NY, 1952.
5. E. Eckert, The group of primitive Pythagorean triangles, this MAGAZINE 54 (1984), 22–27.
6. J. Gilder, Integer-sided triangles with an angle of 60°, *Math. Gazette* 66 (1982), 261–266.
7. L. Mordell, *Diophantine Equations*, Academic Press, New York, NY, 1969.
8. I. Niven, *Irrational Numbers*, The Carus Mathematical Monographs, No. 11, MAA, Washington, DC, 1956.
9. K. Selkirk, Integer-sided triangles with an angle of 120°, *Math. Gazette* 67 (1983), 251–255.
10. B. M. Stewart, *The Theory of Numbers*, Macmillan, New York, NY, 1964.

NOTES

A Combinatorial Interpretation of a Catalan Numbers Identity

DAVID CALLAN
Department of Statistics
University of Wisconsin-Madison
1210 W. Dayton Street
Madison, WI 53706-1693

Introduction The Catalan numbers $C_n := \frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n-1}$ seem to be ubiquitous in combinatorics, and they have featured frequently in these pages. A list of some fifteen more or less different situations in which they occur is given in **[2]**, and they make an unexpected appearance in a problem on the expected length of a World Series **[3]**.

Their generating function $\sum_{n=0}^{\infty} C_n x^n$ is given by $C(x) := \frac{1-\sqrt{1-4x}}{2x}$ (expand the radical by the binomial theorem). Once an identity involving $C(x)$ is found—by whatever means—the question of a combinatorial interpretation arises. For example, isolating the radical and squaring leads to a fundamental identity: $xC(x)^2 = C(x) - 1$. Equating coefficients of x, this says that $C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-1}C_0$, $n \geq 1$, an identity with a well-known interpretation: the left side C_n counts Catalan paths of $2n$ steps (see below) and the right side counts these paths by "first return to the diagonal."

The following identity is obtained in **[4]** in connection with "moments of a Catalan triangle":

$$\sum_{k \geq 1} k(C(x) - 1)^k = \sum_{n \geq 1} 4^{n-1} x^n. \tag{1}$$

(Note that $C(x) - 1$ has constant term 0 and hence the left side of (1) makes sense.) The object of this note is to give a combinatorial interpretation of (1) using lattice paths; that is, we describe two sets of such paths whose cardinalities are the coefficients of x^n on the two sides of (1), and then give a bijection between these two sets.

Combinatorial setting To set the stage, let E and N denote unit steps East and North respectively. A *path* is (for present purposes) a finite sequence of an *even* number of steps and it can be represented graphically in an obvious way. For example, *ENEENE* is given in FIGURE 1 with its steps numbered.

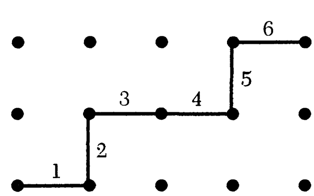


FIGURE 1

It turns out to be helpful to focus on the consecutive *pairs* of steps comprising a path (while not forgetting entirely about the underlying individual steps). Let us denote the pairs EE, NN, EN, NE by H(for horizontal), V(vertical), L(lower hook), U(upper hook) respectively, and call them *double-steps* for short. Thus the path in FIGURE 1 above is also specified by LHU. A *subpath* of a path P is any subsequence of consecutive double-steps in the sequence of double-steps that defines P. In other words, a subpath must contain an even number of steps and begin at an odd-numbered step. Thus, in FIGURE 1, steps numbered 34 and 3456 form subpaths, but 123, 1256, and 2345 do not.

If a path has $2n$ steps, and hence n double-steps, its *length* is n, and we call it an *n-path*. Let $\mathscr{P}_n$ denote the set of n-paths. Since the steps of a path can be chosen independently, $|\mathscr{P}_n| = 2^{2n} = 4^n$ and so $|\mathscr{P}_{n-1}|$ is the coefficient of x^n on the right in (1).

To bring Catalan numbers into the picture, define a *Catalan* path to be a path containing an equal number of E's and N's, whose graph lies entirely on or below the line—necessarily of slope 1—joining its initial and terminal points. We'll call this the *diagonal line* of the path. The two Catalan 2-paths are given in FIGURE 2.

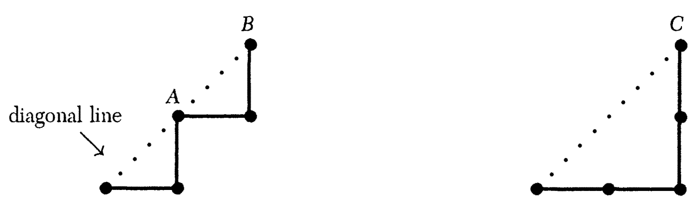


FIGURE 2

In fact the number of Catalan n-paths is the Catalan number C_n (see, for example, [1]).

Catalan paths can be strung together by concatenation, maintaining catalan-icity, so it is useful to identify the points where a Catalan path P returns to its diagonal line. Any such point (with the exception of the initial point) is called a *return* of P. Thus the Catalan paths in FIGURE 2 have two returns (at A and B) and only one return (at C) respectively. (The terminal point is always a return.) We use $[x^n]R(x)$ to denote the coefficient of x^n in $R(x)$ and our next task is to find some class of combinatorial objects counted by $[x^n]\Sigma_{k\geq 1}k(C(x)-1)^k$. This is not as hard as it may seem. First, by direct expansion, $[x^n](C(x)-1)^k$ is $\Sigma C_{i_1}C_{i_2}\cdots C_{i_k}$, where the sum runs over all k-tuples $(i_1, i_2, \dots, i_k)$ of *positive* integers whose sum is n, and hence is the number of k-tuples of *nonvanishing* Catalan paths whose total length is n. Imagine marking the terminal points of these k paths and concatenating them to form a single object, a "marked" Catalan path.

Thus we define a *k-marked Catalan path* as one in which k of its return points (arbitrary except that they must include its terminal point) have been distinguished by "marking" them. Since the original k-tuple of paths can be uniquely retrieved from the single k-marked path, we see that $[x^n](C(x)-1)^k$ is the number of k-marked Catalan n-paths. Now to account for the factor k on the left in (1), we need k copies of each k-marked path, $k \geq 1$. An obvious way to distinguish between k such copies is to "highlight" in each a different one of its k marks. Let $\mathscr{H}_n$ denote the set of all highlighted Catalan n-paths (any number $k \geq 1$ of marked returns). By construction of $\mathscr{H}_n$, therefore, the coefficient of x^n on the left side of (1) is $|\mathscr{H}_n|$.

The bijection Now we would like a bijection $\pi : \mathscr{P}_{n-1} \to \mathscr{H}_n$. Here is one. First note that if two Catalan subpaths of a given path $P \in \mathscr{P}_{n-1}$ overlap (even just at an endpoint), their union is again a Catalan subpath because, unless one path is contained in the other, their diagonal lines must coincide. Hence the maximal Catalan subpaths of P_{n-1} are disjoint. Find these subpaths and leave them intact, but replace every other double-step by an upper hook, add an E at the start and an N at the finish. Denote the resulting path $\bar{P}$. An example is given in FIGURE 3, below. (The heavy lines represent the maximal Catalan subpaths of P.)

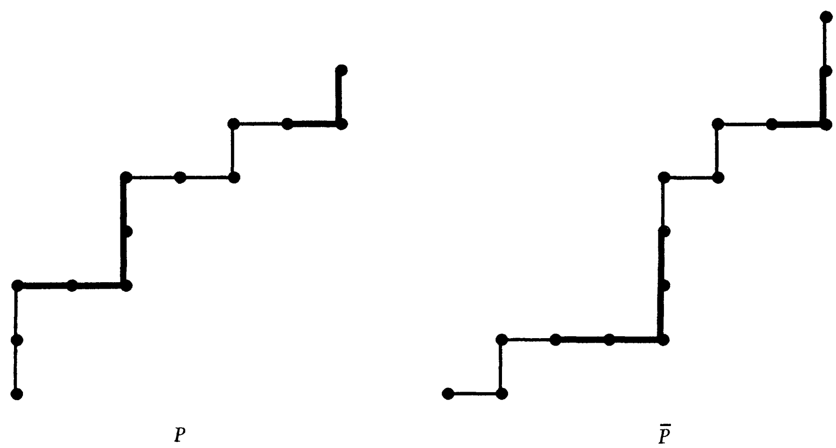


FIGURE 3

Clearly, in $\bar{P}$ the diagonal lines of all Catalan paths left intact from P coincide, and $\bar{P}$ is Catalan with diagonal line one unit higher. Next we will mark and highlight $\bar{P}$ appropriately. The terminal point of $\bar{P}$ is of course a return that must be marked. Every other return of $\bar{P}$ is the midpoint of a double-step in P that was not part of a maximal—hence of any—Catalan subpath of P (and *vice versa*). None of these double-steps was a lower hook (which would already be a Catalan path). Thus the nonterminal returns in $\bar{P}$ correspond only to H's, V's or U's in P. Now mark those returns that correspond to H's or V's and suppose $A_1, A_2, \ldots, A_s$ is a listing, in order, of these H's and V's. Can we canonically choose one of the marked returns in $\bar{P}$ to highlight? Yes: the key observation is that a V will never follow an H in the list $A_1, A_2, \ldots, A_s$. For, if it did, P would have a subpath of the form $H * * * V$, where the asterisks represent a (possibly empty) sequence of upper hooks and/or Catalan subpaths. But then $H * * * V$ would be a Catalan subpath (draw a picture to see this) whereas by definition of this H and V, they do not lie on any Catalan subpath. So the sequence $\{A_i\}$ looks like $V \cdots VH \cdots H$. This allows us to highlight the first marked return in $\bar{P}$ corresponding to an H in this sequence (or the terminal point of $\bar{P}$ if the sequence contains no H's). Thus, for $\bar{P}$ in FIGURE 3 above, returns $1, 2, 4$ get marked, the sequence $\{A_i\}$ is VH, and so return 2 gets highlighted (with a circle).

Is this mapping π reversible? Yes! Given, say, the highlighted marked Catalan path in FIGURE 4a, draw its diagonal line and a parallel line one unit below it (the dotted lines in FIGURE 4a). Discard the initial and terminal steps and leave what's below the lower dotted line intact (heavy lines in FIGURE 4). Modify each upper hook that lies between the dotted lines as follows: leave it intact unless its midpoint is marked, in which case change it to a V if it lies (strictly) to the left of the highlight and to an H otherwise. Thus the path in FIGURE 4a becomes the path in FIGURE 4b.

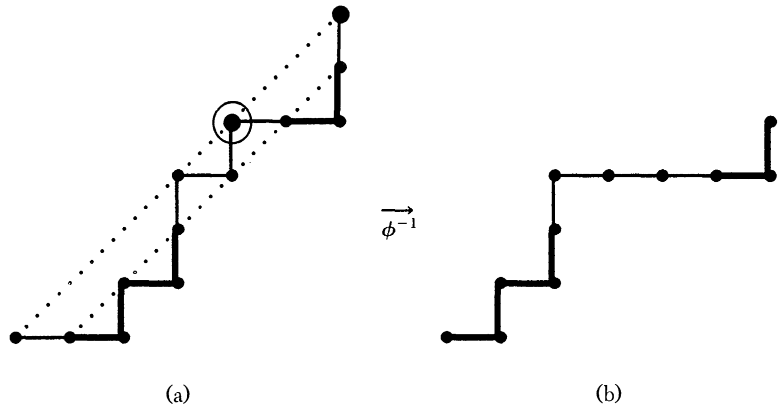


FIGURE 4

Conclusion Since π has an inverse, it is a bijection. To make this bijection fully explicit, an algorithm for identifying the maximal Catalan subpaths of a given path would be needed. As an interesting exercise, the reader might like to try to devise such an algorithm and perhaps write a computer program to implement it.

Finally, if you've read this far, you might like to try your hand at a simpler bijection: There are $\binom{2n}{n}$ paths of n East and n North steps; hence $\frac{1}{2}\binom{2n}{n}$ such paths that start with an East step. On the other hand, there are $\frac{1}{2}\binom{2n}{n}$ marked Catalan n-paths (any number $k \geq 1$ of marked returns, no highlight). See if you can find a bijection between these two sets. Hint: try flipping appropriate Catalan subpaths in the diagonal. If you succeed, you will have found a combinatorial interpretation for

$$\sum_{k=1}^{\infty} (C(x) - 1)^k = \frac{x}{2\sqrt{1-4x}},$$

an identity also given in [4].

Acknowledgment. Thanks to the referees for several helpful suggestions and reference [4].

REFERENCES

1. D. Callan, Pair them up!: A visual approach to the Chung–Feller theorem, *College Math. J.* 26 (1995), 196–198
2. R. Eggleton and R. Guy, Catalan strikes again! How likely is a function to be convex?, this MAGAZINE 61 (1988), 211–219
3. L. Shapiro and W. Hamilton, The Catalan numbers visit the World Series, this MAGAZINE 66 (1993), 20–22
4. L. Shapiro, W.-J. Woan, S. Getu, Runs, slides and moments, *Siam J. Alg. Disc. Meth.* 4 (1983) 459–466

Cones, *k*-Cycles of Linear Operators, and Problem B4 on the 1993 Putnam Competition

SELWYN HOLLIS
Armstrong Atlantic State University
Savannah, GA 31419

1. Introduction Problem B4 on the 54th William Lowell Putnam Mathematical Competition was as follows:

The function $K(x, y)$ is positive and continuous for $0 \leq x \leq 1$, $0 \leq y \leq 1$, and the functions $f(x)$ and $g(x)$ are positive and continuous for $0 \leq x \leq 1$. Suppose that for all x, $0 \leq x \leq 1$,

$$\int_0^1 f(y)K(x,y)\,dy = g(x) \quad \textit{and} \quad \int_0^1 g(y)K(x,y)\,dy = f(x).$$

Show that $f(x) = g(x)$ for $0 \leq x \leq 1$.

We will present a simple "geometric" solution to this problem which, unlike the solution presented in [**3**], lends itself naturally to generalization. We believe that the process is an interesting exercise in abstraction—discovering the true essence of a specific problem and exploiting it to obtain a far more general result. For example, in addition to solving Problem B4, the theorem that we will present also implies the following result:

Let A be a real $n \times n$ matrix with positive entries. If x is a vector with positive entries and $A^k x = x$ for some $k \in \mathbb{N}$, then $Ax = x$.

2. Solving B4 The function K in Problem B4 is the "kernel" of a linear integral operator $\mathscr{K}$, defined on the set of continuous functions φ on $[0, 1]$ by

$$\mathscr{K}: \varphi \mapsto \mathscr{K}\varphi,$$

where

$$[\mathscr{K}\varphi](x) = \int_0^1 K(x,y)\varphi(y)\,dy \quad \text{for} \quad 0 \leq x \leq 1.$$

Thus we can restate the problem as:

Show that if f and g are positive, continuous functions on $[0, 1]$ with $\mathscr{K}f = g$ and $\mathscr{K}g = f$, then $f = g$.

Now, what are the important properties of the operator $\mathscr{K}$? First of all, $\mathscr{K}$ is a linear operator, since $\mathscr{K}(f+g) = \mathscr{K}f + \mathscr{K}g$ and $\mathscr{K}(cf) = c\mathscr{K}f$ for all continuous functions f, g on $[0, 1]$ and real numbers c. Also, since the kernel K is a positive function on $[0, 1] \times [0, 1]$, the operator $\mathscr{K}$ maps nonnegative continuous functions to nonnegative continuous functions; that is, if φ is nonnegative and continuous on $[0, 1]$, then $\mathscr{K}\varphi$ is nonnegative and continuous on $[0, 1]$. But in fact, a stronger statement is true, which is nearly as easy to see, and left to the reader:

Exercise 1. Show that if φ is nonnegative, continuous, and not identically zero on $[0, 1]$, then $\mathscr{K}\varphi$ is *positive* and continuous on $[0, 1]$.

We contend that this property and linearity are the only important properties of the operator $\mathscr{K}$ at work in the problem.

Let us now proceed with our solution to Problem B4. Suppose that f and g are positive continuous functions on [0, 1] with $\mathscr{K}f = g$ and $\mathscr{K}g = f$. First observe that if $f = \lambda g$ for some positive constant λ, then $\mathscr{K}f = \lambda\mathscr{K}g$, which implies that $g = \lambda f$. Consequently $\lambda = 1$, and we are done. So assume that f is not a constant multiple of g. Since f and g are each positive on [0, 1], each of $f - \lambda g$ and $g - \lambda f$ will be positive on [0, 1] for all sufficiently small $\lambda > 0$. Also, for all sufficiently large $\lambda > 0$, neither $f - \lambda g$ nor $g - \lambda f$ is positive on [0, 1]. So let

$$\lambda^* = \sup\{\lambda > 0 : f - \lambda g \text{ and } g - \lambda f \text{ are positive on } [0, 1]\}.$$

Then $f - \lambda^* g$ and $g - \lambda^* f$ are each nonnegative on [0, 1], and at least one of these functions is not positive on [0, 1]. Suppose that $f - \lambda^* g$ is not positive, and observe that

$$\mathscr{K}(g - \lambda^* f) = \mathscr{K}g - \lambda^*\mathscr{K}f = f - \lambda^* g.$$

This contradicts the fact that $\mathscr{K}\varphi$ must be positive on [0, 1] whenever φ is nonnegative and continuous on [0, 1]. Thus we must conclude that $f = g$.

3. Generalizing the problem The framework in which we will work is that of a *normed linear space*. A *linear space* (or vector space) over the real numbers $\mathbb{R}$ is essentially a set X on which operations of addition and scalar multiplication are defined in such a way that

$$x, y \in X \Rightarrow x + y \in X, \quad \text{and} \quad x \in X, \alpha \in \mathbb{R} \Rightarrow \alpha x \in X.$$

Identity elements, additive inverses, and commutativity and associativity properties are necessary as well; we refer the reader to a book such as **[1]** or **[5]** for more details. A linear space X becomes a *normed linear space* when we define upon it a *norm*, i.e., a function $\|\cdot\|: x \mapsto \|x\|$ from V into the nonnegative real numbers that provides a measure of magnitude, or distance to the zero element of X. For example, $\mathbb{R}$ itself is a normed linear space with the usual definitions of addition and scalar multiplication, and norm defined by absolute value: $\|x\| = |x|$. Also, $\mathbb{R}^n$ is a normed linear space with the usual definitions of addition and scalar multiplication of vectors and norm defined by the Euclidean distance to the origin: $\|x\| = \sum_{i=1}^{n} x_i^2$. See **[5**, Chapter 10] for the full list of properties that define a norm. A norm on a linear space defines a *metric*, or notion of distance, on that space by

$$\operatorname{dist}(x, y) = \|x - y\|.$$

The set of continuous functions on [0, 1], denoted by $\mathscr{C}[0, 1]$, is another example of a normed linear space; addition and scalar multiplication are defined in the natural way. The usual norm on $\mathscr{C}[0, 1]$ is defined by

$$\|f\| = \max_{x \in [0,1]} |f(x)|.$$

The set of nonnegative continuous functions on [0, 1], denoted by $\mathscr{C}_+[0, 1]$, is a subset of $\mathscr{C}[0, 1]$, but is not a subspace, because it is not closed under (negative) scalar multiplication. It is, however, a *strict cone* **[4]**. A subset C of a linear space is a *cone* if C is closed under addition and closed under multiplication by *nonnegative* scalars. A cone is *strict* if φ and $-\varphi$ are in C only if $\varphi = 0$.

Any cone C induces a partial ordering $\succcurlyeq$ on the entire space defined by

$$\varphi \succcurlyeq \psi \Leftrightarrow \varphi - \psi \in C.$$

In particular, $\varphi \in C \Leftrightarrow \varphi \succcurlyeq 0$, where 0 represents the zero element of the space. The induced ordering is *partial* because, in general, arbitrary elements $x, y \in X$ need not satisfy either $x \succcurlyeq y$ or $y \succcurlyeq x$. (For more details on cones and partial orderings, see **[2]**.)

In a normed linear space, we can talk about the *interior* of a cone C, which we will denote by $\operatorname{int} C$, defined as the set of all $x \in C$ for which some ε-neighborhood $\mathscr{N}_\varepsilon = \{y \in X \mid \|x - y\| < \varepsilon\}$ of x is contained in C. In this case we write

$$\varphi \succ \psi \Leftrightarrow \varphi - \psi \in \operatorname{int} C.$$

The reader should note that $[0, \infty)$ is an example of a strict cone in $\mathbb{R}$ with interior $(0, \infty)$ and that $\{x \in \mathbb{R}^n | x_i \geq 0 \,\forall\, i\}$ is an example of a strict cone in $\mathbb{R}^n$ with interior $\{x \in \mathbb{R}^n | x_i > 0 \,\forall\, i\}$. These are the standard *nonnegative cones* in these spaces; many other cones are possible in $\mathbb{R}$ and many other strict cones are possible in $\mathbb{R}^n$, $n \geq 2$.

In the context of Problem B4, where the cone C of interest is the set of nonnegative continuous functions on $[0,1]$, it is not difficult to show that $\operatorname{int} C$ is precisely the set of *positive* continuous functions on $[0,1]$ and that the partial ordering induced by C says that, for all $x \in [0,1]$, $\varphi \succcurlyeq \psi \Leftrightarrow \varphi(x) \geq \psi(x)$ and $\varphi \succ \psi \Leftrightarrow \varphi(x) > \psi(x)$. Thus, this "nonnegative cone" in the linear space of continuous functions induces a natural partial ordering on that space.

Exercise 2. Suppose that $n > 1$ and $x_1, x_2, \ldots, x_n$ are elements of a strict cone C. Show (by induction) that if $x_1 + x_2 + \cdots + x_n = 0$, then $x_1 = x_2 = \cdots = x_n = 0$.

Now we can state an abstract theorem that encompasses Problem B4. Its proof is essentially identical to the solution to Problem B4 given above.

THEOREM 1. *Let X be a normed linear space and C a strict cone in X with nonempty interior. Let $\mathscr{L}$ be a linear operator on X with the property that*

$$\mathscr{L}: C \setminus \{0\} \to \operatorname{int} C.$$

If $x, y \in C$ are such that $\mathscr{L}x = y$ and $\mathscr{L}y = x$, then $x = y$.

Proof. Suppose that x and y are members of C with $\mathscr{L}x = y$ and $\mathscr{L}y = x$. First observe that either $x = y = 0$ or $x, y \in \operatorname{int} C$; assume the latter. If $x = \lambda y$ for some positive constant λ, then $\mathscr{L}x = \lambda \mathscr{L}y$, which implies that $y = \lambda x$. Consequently $\lambda = 1$, and we are done. So assume that x is not a constant multiple of y. Since $xy \in \operatorname{int} C$,

(A_1) *each of $x - \lambda y$ and $y - \lambda x$ is in* $\operatorname{int} C$ *for all sufficiently small* $\lambda > 0$;

(A_2) *for all sufficiently large* $\lambda > 0$, *neither $x - \lambda y$ nor $y - \lambda x$ is in* $\operatorname{int} C$.

So let $\lambda^* = \sup\{\lambda > 0 : x - \lambda y$ and $y - \lambda x$ are in $\operatorname{int} C\}$. Then

(A_3) $x - \lambda^* y$ *and* $y - \lambda^* x$ *are in* $C \setminus \{0\}$, *and at least one of these is not in* $\operatorname{int} C$.

So suppose that $x - \lambda^* y \notin \operatorname{int} C$, and observe that

$$\mathscr{L}(y - \lambda^* x) = \mathscr{L}y - \lambda^* \mathscr{L}x = x - \lambda^* y.$$

This contradicts the assumption that $\mathscr{L}: C \setminus \{0\} \to \operatorname{int} C$. Therefore $x = y$. ■

Exercise 3. Give detailed proofs of A_1, A_2, and A_3 above.

4. Further generalization: k-cycles Let $\mathscr{L}$ be a linear operator on a linear space X. By a k-cycle of $\mathscr{L}$, $k \geq 1$, we mean a collection of distinct elements $\{x_1, x_2, \dots, x_k\} \subset X$ such that

$$\mathscr{L}x_1 = x_2, \mathscr{L}x_2 = x_3, \dots, \mathscr{L}x_{k-1} = x_k, \mathscr{L}x_k = x_1.$$

Theorem 1 says that the operator $\mathscr{L}$ described there has no 2-cycle in C. Note also that a 1-cycle $\{x_1\}$ of any operator $\mathscr{L}$ is just a *fixed-point* of $\mathscr{L}$, and that a k-cycle of $\mathscr{L}$ gives rise to a set of k fixed points of the k-fold composition $\mathscr{L}^k$ of $\mathscr{L}$.

Theorem 1 can be generalized as follows:

THEOREM 2. *Let X be a normed linear space and C a strict cone in X with nonempty interior. Suppose further that $\mathscr{L}$ is a linear operator on X with the property that*

$$\mathscr{L} : C \setminus \{0\} \to \operatorname{int} C.$$

Then $\mathscr{L}$ has no k-cycle in C with $k > 1$.

The proof of Theorem 2 relies on the following lemma.

LEMMA. *Suppose that the hypotheses of Theorem 2 are satisfied and that $\{x_1, x_2, \dots, x_k\}$ is a k-cycle of $\mathscr{L}$ in C with $k > 1$. Then $x_i - \lambda\Sigma_{j \neq i} x_j \neq 0$ for all $i = 1, 2, \dots, k$ and all $\lambda > 0$.*

Proof. Assume the hypotheses stated in Theorem 2, and suppose that $\{x_1, x_2, \dots, x_k\}$ is a k-cycle of $\mathscr{L}$ in C with $k > 1$. Note first that each $x_i \in \operatorname{int} C$. Suppose for the sake of contradiction that $1 \leq i \leq k$, $\lambda > 0$, and $x_i - \lambda\Sigma_{j \neq i} x_j = 0$. Without loss of generality, we can assume that $i = 1$. Now observe that in addition to $x_1 - \lambda\Sigma_{j \neq 1} x_j = 0$, we also have

$$\mathscr{L}\left(x_1 - \lambda \sum_{j \neq 1} x_j\right) = x_2 - \lambda \sum_{j \neq 2} x_j = 0.$$

Subtracting the second of these equations from the first shows that $(1 + \lambda)x_1 - (1 + \lambda)x_2 = 0$, from which we conclude that $x_1 = x_2$, contradicting the assumption that $\{x_1, x_2, \dots, x_k\}$ is a k-cycle of $\mathscr{L}$. ∎

Proof of Theorem 2. Let $\{x_1, x_2, \dots, x_k\}$ be a k-cycle of $\mathscr{L}$ in C, with $k > 1$. Note first that each $x_i \in \operatorname{int} C$. It follows that, for all sufficiently small $\lambda > 0$,

$$x_i - \lambda \sum_{j \neq i} x_j \in \operatorname{int} C$$

for each $i = 1, 2, \dots, k$. Also, for all sufficiently large $\lambda > 0$,

$$x_i - \lambda \sum_{j \neq i} x_j \notin \operatorname{int} C$$

for each $i = 1, 2, \dots, k$. So let

$$\lambda^* = \sup\left\{\lambda > 0 : x_i - \lambda \sum_{j \neq i} x_j \in \operatorname{int} C \quad \text{for each } i = 1, 2, \dots, k\right\}.$$

The lemma implies that

$$x_i - \lambda^* \sum_{j \neq i} x_j \in C \setminus \{0\}$$

for each $i = 1, 2, \ldots, k$. Furthermore, $x_i - \lambda^* \sum_{j \neq i} x_j \notin \operatorname{int} C$ for some $i = 1, 2, \ldots, k$, since the alternative would contradict the definition of λ^*. We can assume without loss of generality that $i = 2$; that is, $x_2 - \lambda^* \sum_{j \neq 2} x_j \notin \operatorname{int} C$. But $x_1 - \lambda^* \sum_{j \neq 1} x_j \in C \setminus \{0\}$ and

$$\mathscr{L}\left(x_1 - \lambda^* \sum_{j \neq 1} x_j\right) = x_2 - \lambda^* \sum_{j \neq 2} x_j.$$

This contradicts the assumption that $\mathscr{L}: C \setminus \{0\} \to \operatorname{int} C$. ■

Finally, we will prove the corollary stated in the Introduction.

COROLLARY. *Let A be a real $n \times n$ matrix with positive entries. If v is a vector with positive entries and $A^k v = v$ for some $k \in \mathbb{N}$, then $Av = v$.*

Proof. Let $\mathscr{L}$ be the linear operator on $\mathbb{R}^n$ defined by multiplication by A; that is, $\mathscr{L}x = Ax$ for all $x \in \mathbb{R}^n$. Let C be the nonnegative cone in $\mathbb{R}^n$, that is, $C = \{(x_1, x_2, \ldots, x_n) | x_i \geq 0, \ i = 1, 2, \ldots, n\}$. Then $\operatorname{int} C = \{(x_1, x_2, \ldots, x_n) | x_i > 0, i = 1, 2, \ldots, n\}$. Let $v \in C \setminus \{0\}$. Since A has positive entries, Av will have positive entries; that is, $\mathscr{L}$ maps $C \setminus \{0\}$ into $\operatorname{int} C$. We can assume without loss of generality that k is the *least* positive integer for which $A^k v = v$. So the vectors v, Av, $A^2 v, \ldots, A^{k-1} v$ form a k-cycle of $\mathscr{L}$ if $k > 1$. By Theorem 2, no such cycle can exist. Thus $k = 1$, and $Av = v$. ■

REFERENCES

1. K. H. Hoffman and R. A. Kunze, *Linear Algebra*, Prentice Hall, Englewood Cliffs, NJ, 1971.
2. J. L. Kelley, I. Namioka, et al., *Linear Topological Spaces*, Springer-Verlag, New York, NY, 1976.
3. L. F. Klosinski, G. L. Alexanderson, and L. C. Larson, The Fifty-Fourth William Lowell Putnam Mathematical Competition, *Amer. Math. Monthly* 101 (1994), 725–734.
4. R. H. Martin, Jr., *Nonlinear Operators and Differential Equations in Banach Spaces*, R. E. Krieger, Malabar, FL, 1987.
5. H. L. Royden, *Real Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1988.

Flett's Mean Value Theorem for Holomorphic Functions

R. M. DAVITT
R. C. POWERS
T. RIEDEL
P. K. SAHOO
University of Louisville
Louisville, KY 40292

The mean value theorem is a well known result usually covered in a first semester calculus course. There are many other types of mean value theorems that are less well known. In 1958, T. M. Flett **[3]** proved one such result. Specifically, if $f:[a,b]\to\mathbb{R}$ is differentiable and satisfies $f'(a)=f'(b)$, then there exists η in the open interval (a,b) such that $f(\eta)-f(a)=(\eta-a)f'(\eta)$. FIGURE 1 shows the nice geometric interpretation of this result. If the curve $y=f(x)$ has a tangent at each point in $[a,b]$, and if the tangents at $(a,f(a))$ and $(b,f(b))$ are parallel, then there is an intermediate point η such that the tangent at $(\eta,f(\eta))$ passes through the point $(a,f(a))$. For other examples of mean value theorems we refer the reader to **[4]**–**[9]**.

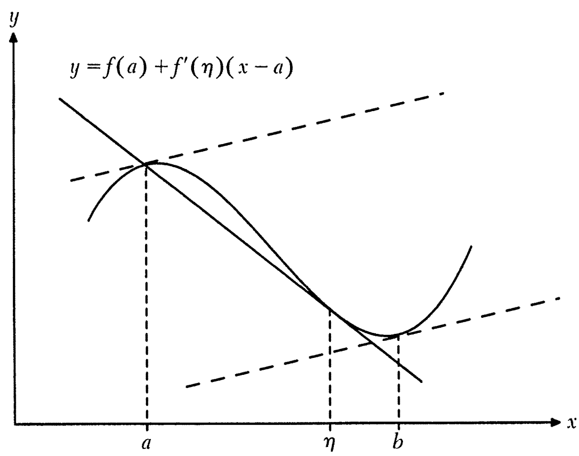


FIGURE 1
Geometric interpretation of Flett's theorem.

These results do not, in general, immediately extend to holomorphic functions of a complex variable. For the case of Rolle's theorem, the function $f(z)=e^z-1$ has value 0 at $z=0$ and $z=2\pi i$, but $f'(z)=e^z$ has no zeros in the complex plane. Evard and Jafari **[1]** get around this difficulty by working with the real and imaginary parts of a holomorphic function. Another approach is taken by Samuelsson in **[8]**. The goal of this note is to prove a version of Flett's theorem for holomorphic functions of a complex variable in the spirit of Evard and Jafari.

The first step is to extend Flett's mean value theorem for real functions to a result that does not depend on the hypothesis $f'(a)=f'(b)$, but reduces to Flett's theorem when this is the case.

THEOREM 1. *If $f:[a,b]\to\mathbb{R}$ is a differentiable function, then there exists a point $\eta\in(a,b)$ such that*

$$f(\eta)-f(a)=(\eta-a)f'(\eta)-\frac{1}{2}\frac{f'(b)-f'(a)}{b-a}(\eta-a)^2.$$

Proof. Consider the auxiliary function $\psi:[a,b]\to\mathbb{R}$ defined by

$$\psi(x)=f(x)-\frac{1}{2}\frac{f'(b)-f'(a)}{b-a}(x-a)^2.$$

Then ψ is differentiable on $[a,b]$, and

$$\psi'(x)=f'(x)-\frac{f'(b)-f'(a)}{b-a}(x-a).$$

It follows that $\psi'(a)=\psi'(b)=f'(a)$. Applying Flett's mean value theorem to ψ gives $\psi(\eta)-\psi(a)=(\eta-a)\psi'(\eta)$ for some $\eta\in(a,b)$. Rewriting ψ and ψ' in terms of f gives the asserted result.

Our next step is to introduce some notation. Let $\mathbb{C}$ denote the set of complex numbers. For distinct a and b in $\mathbb{C}$, let $[a,b]$ denote the set $\{a+t(b-a)|t\in[0,1]\}$; we will refer to $[a,b]$ as a *line segment* or a *closed interval* in $\mathbb{C}$. Similarly, (a,b) denotes the set $\{a+t(b-a)|t\in(0,1)\}$.

Flett's theorem is not valid for complex valued functions of a complex variable. To see this, consider the function $f(z)=e^z-z$. Then f is holomorphic, and $f'(z)=e^z-1$. Therefore, we have $f'(2k\pi i)=e^{2k\pi i}-1=0$ for all integers k. In particular, $f'(0)=f'(2\pi i)$, that is, the derivatives of f at the endpoints of the closed interval $[0,2\pi i]$ are equal. Nevertheless, the equation

$$f(z)-f(0)=f'(z)z$$

has no solution on the interval $(0,2\pi i)$, as we now show. The equation above gives $(1-z)=e^{-z}$ and, since $z=iy$, we get $1-iy=\cos y-i\sin y$. The comparison of the real and imaginary parts gives the system $\cos y=1$ and $\sin y=y$, which has no solution in the interval $(0,2\pi)$. Thus Flett's theorem fails in the complex domain.

We now present a generalization of Theorem 1 for holomorphic functions where, for any two complex numbers u and v,

$$\langle u,v\rangle=\operatorname{Re}(u\bar{v}).$$

THEOREM 2. *Let f be a holomorphic function defined on an open convex subset D of $\mathbb{C}$. Let a and b be two distinct points in D. Then there exist $z_1, z_2\in(a,b)$ such that*

$$\operatorname{Re}(f'(z_1))=\frac{\langle b-a,f(z_1)-f(a)\rangle}{\langle b-a,z_1-a\rangle}+\frac{1}{2}\frac{\operatorname{Re}(f'(b)-f'(a))}{b-a}(z_1-a)$$

and

$$\operatorname{Im}(f'(z_2))=\frac{\langle b-a,-i[f(z_2)-f(a)]\rangle}{\langle b-a,z_2-a\rangle}+\frac{1}{2}\frac{\operatorname{Im}(f'(b)-f'(a))}{b-a}(z_2-a).$$

Proof. Let $u(z) = \mathrm{Re}(f(z))$ and $v(z) = \mathrm{Im}(f(z))$ for $z \in D$. We now define the auxiliary function $\phi : [0,1] \to \mathbb{R}$ by

$$\phi(t) = \langle b-a, f(a+t(b-a))\rangle, \tag{1}$$

which is

$$\phi(t) = \mathrm{Re}(b-a)u(a+t(b-a)) + \mathrm{Im}(b-a)v(a+t(b-a))$$

for every $t \in [0,1]$. Therefore, using the Cauchy–Riemann equations, we get

$$\begin{aligned}\phi'(t) &= \langle b-a, (b-a)f'(a+t(b-a))\rangle \\ &= \mathrm{Re}\big((b-a)^2\big)\frac{\partial u(z)}{\partial x} + \mathrm{Im}\big((b-a)^2\big)\frac{\partial u(z)}{\partial x} \\ &= |b-a|^2\frac{\partial u(z)}{\partial x} \\ &= |b-a|^2\,\mathrm{Re}(f'(z)).\end{aligned}$$

Applying Theorem 1 to ϕ on $[0,1]$, we obtain

$$t_1\phi'(t_1) = \phi(t_1) - \phi(0) + \frac{1}{2}\frac{\phi'(1)-\phi'(0)}{1-0}(t_1-0)^2$$

for some $t_1 \in (0,1)$. Thus

$$t_1|b-a|^2\,\mathrm{Re}(f'(z_1)) = \phi(t_1) - \phi(0) + \frac{1}{2}[\phi'(1)-\phi'(0)]t_1^2,$$

where $z_1 = a + t_1(b-a)$. Further, since $z_1 = a + t_1(b-a)$ and $t_1 \in [0,1]$, we have $t_1|b-a|^2 = \langle b-a, z_1-a\rangle$. Hence the equation $t_1|b-a|^2\,\mathrm{Re}(f'(z_1)) = \phi(t_1) - \phi(0) + \frac{1}{2}[\phi'(1)-\phi'(0)]t_1^2$ reduces to

$$\mathrm{Re}(f'(z_1)) = \frac{\phi(t_1)-\phi(0)}{t_1|b-a|^2} + \frac{1}{2}\frac{\phi'(1)-\phi'(0)}{|b-a|^2}t_1.$$

Using (1) and the fact that $z_1 = a + t_1(b-a)$ in the above equation, we obtain

$$\mathrm{Re}(f'(z_1)) = \frac{\langle b-a, f(z_1)-f(a)\rangle}{\langle b-a, z_1-a\rangle} + \frac{1}{2}\frac{\mathrm{Re}(f'(b)-f'(a))}{b-a}(z_1-a). \tag{2}$$

Letting $g = -if$, we have

$$\mathrm{Re}(g'(z)) = \frac{\partial v(z)}{\partial x} = -\frac{\partial u(z)}{\partial y} = \mathrm{Im}(f'(z)).$$

Now, applying the first part to g, we obtain

$$\mathrm{Re}(g'(z_2)) = \frac{\langle b-a, g(z_2)-g(a)\rangle}{\langle b-a, z_2-a\rangle} + \frac{1}{2}\frac{\mathrm{Re}(g'(b)-g'(a))}{b-a}(z_2-a)$$

for some $z_2 \in (a,b)$. By (2) the above yields

$$\mathrm{Im}(f'(z_2)) = \frac{\langle b-a, -i[f(z_2)-f(a)]\rangle}{\langle b-a, z_2-a\rangle} + \frac{1}{2}\frac{\mathrm{Im}(f'(b)-f'(a))}{b-a}(z_2-a)$$

and the proof is complete.

The next corollary follows immediately; it is the complex version of Flett's mean value theorem.

COROLLARY. *Let f be a holomorphic function defined on an open convex subset D of $\mathbb{C}$. Let a and b be two distinct points in D, and $f'(a)=f'(b)$. Then there exist $z_1, z_2 \in (a,b)$ such that*

$$\mathrm{Re}(f'(z_1)) = \frac{\langle b-a, f(z_1)-f(a)\rangle}{\langle b-a, z_1-a\rangle}$$

and

$$\mathrm{Im}(f'(z_2)) = \frac{\langle b-a, -i[f(z_2)-f(a)]\rangle}{\langle b-a, z_2-a\rangle}.$$

Returning to our original example $f(z)=e^z-z$, z_1 and z_2 predicted by Theorem 2 have values $z_1 \approx 4.49341i$ and $z_2 \approx 2.33112i$.

There are many ways to generalize the results of this note. For example, in 1966, Trahan **[9]** extended Flett's theorem by replacing the boundary condition $f'(a)=f'(b)$ with $[f'(a)-m][f'(b)-m]>0$, where $m=\frac{f(b)-f(a)}{b-a}$. A similar modification of the Corollary gives an extension of Trahan's result to the complex plane. Specifically, replace the boundary condition $f'(a)=f'(b)$ with the conditions $[\mathrm{Re}(f'(a))-m_1][\mathrm{Re}(f'(b))-m_1]>0$ and $[\mathrm{Im}(f'(a))-m_2][\mathrm{Im}(f'(b))-m_2]>0$, where $m_1=\frac{\langle b-a, f(b)-f(a)\rangle}{\langle b-a, b-a\rangle}$ and $m_2=\frac{\langle b-a, -i[f(b)-f(a)]\rangle}{\langle b-a, b-a\rangle}$. In 1995, Evard, Jafari, and Polyakov **[2]** extended Rolle's theorem for holomorphic functions on line segments **[1]** to a result satisfied by arbitrary curves connecting a and b where $f(a)=f(b)$. For their result to hold the domain must satisfy a relatively weak *almost convexity* condition. The same method could be used here, but we wish to limit ourselves to a more classical setting.

Acknowledgment. The authors would like to thank the referees for their many helpful comments.

REFERENCES

1. J. C. Evard and F. Jafari, A complex Rolle's theorem. *Amer. Math. Monthly*, 99 (1992), 858–861.
2. J. C. Evard, F. Jafari, and P. Polyakov, Generalizations and applications of a complex Rolle's theorem, *Nieuw Archief voor Wiskunde*, 13 (1995), 173–179.
3. T. M. Flett, A mean value theorem, *Math. Gazette*, 42 (1958), 38–39.
4. M. Furi and M. Martelli, A multidimensional version of Rolle's theorem, *Amer. Math. Monthly*, 102 (1995), 243–249.
5. R. M. McLeod, Mean value theorem for vector valued functions, *Proc. Edinburgh Math. Soc.*, 14 (1964), 197–209.
6. I. Rosenholtz, A topological mean value theorem for the plane, *Amer. Math. Monthly*, 98 (1991), 149–153.
7. P. K. Sahoo and T. Riedel, *Mean Value Theorems and Functional Equations*, World Scientific Publishing, Singapore, 1998.
8. A. Samuelsson, A local mean value theorem for analytic functions, *Amer. Math. Monthly*, 80 (1973), 45–46.
9. D. H. Trahan, A new type of mean value theorem, this MAGAZINE, 39 (1966), 264–268.

Fermat's Little Theorem: A Proof by Function Iteration

LIONEL LEVINE
Harvard University
Cambridge, MA 02138

It is a beautiful property of prime numbers, first proved more than three centuries ago by Fermat, that $k^p \equiv k \pmod{p}$ for all prime numbers p and all integers k. Here we present a simple proof of Fermat's "little" theorem by considering iterates of the function $f(z) = z^k$ on the complex plane. The method of proof has the advantage of generalizing the theorem to composite exponents: for every n we find a degree-n polynomial, with coefficients ± 1, that is always divisible by n. This is different from Euler's generalization ($k^{\phi(n)} \equiv 1 \pmod{n}$ for k and n coprime). The method of proof is potentially more general still, since it is easily adapted to other functions f. Indeed, for any set S, every function $f: S \to S$ satisfying a certain property corresponds to a divisibility result similar to Fermat's little theorem.

Let k be a positive integer and p be prime. Consider the function $f(z) = z^k$ for complex z. The pth iterate of f is evidently $f^p(z) = z^{k^p}$. Let P_p be the set of those z that are fixed under f^p but not under f itself. Then $|P_p| = k^p - k$. But if $z \in P_p$, then $f^i(z) \in P_p$ for every $i = 0, 1, \dots, p-1$; and since p is prime, the p values $z, f(z), \dots, f^{p-1}(z)$ are all distinct. Hence, we can partition P_p into equivalence classes, each containing p elements, obtaining

$$p \mid k^p - k, \tag{1}$$

Fermat's little theorem! The advantage to such an unusual approach is that it allows us to see a generalization that we might have missed otherwise. In general, if $f^n(z) = z$, then there must be some least positive integer d such that $f^d(z) = z$. Then $d \mid n$. Call this d the *order* of z. Let P_n be the set of all z of order n. As before, if $z \in P_n$ then $f^i(z) \in P_n$ for all $i = 0, 1, \dots, n-1$; and the n values $z, f(z), \dots, f^{n-1}(z)$ are all distinct because n is the *least* positive integer such that $f^n(z) = z$. Hence

$$n \mid |P_n| \tag{2}$$

for all positive integers n. In the case when n is prime, (2) reduces to (1), Fermat's little theorem. But when n is composite, (2) gives a different degree-n polynomial, instead of $k^n - k$, that n must divide.

To illustrate what happens for general n, consider first the case $n = pq$, where p and q are distinct primes. There are k^{pq} values of z fixed under f^{pq}, and each such z has order d for exactly one d dividing pq. So

$$|P_{pq}| + |P_p| + |P_q| + |P_1| = k^{pq}.$$

Substituting $|P_p| = k^p - k$, $|P_q| = k^q - k$, and $|P_1| = k$, and solving for $|P_{pq}|$, we get

$$|P_{pq}| = k^{pq} - k^p - k^q + k,$$

so by (2), $pq \mid k^{pq} - k^p - k^q + k$. So the polynomial $k^{pq} - k^p - k^q + k$ in the case $n = pq$ is the counterpart of the Fermat polynomial $k^p - k$ in the case $n = p$. For general n, there are k^n values of z fixed under f^n and every such z has order d for exactly one d dividing n, so

$$\sum_{d|n} |P_d| = k^n. \tag{3}$$

In their current form, the equations (3)—there is one equation for each n = 1, 2, 3, ... —give an explicit formula for k^n in terms of the values $|P_d|$. What we'd like to do is "invert" (3) into an explicit formula for each $|P_n|$ in terms of the powers of k. By (2), this will yield for each n a polynomial in k that is always divisible by n. The technique that accomplishes this task is called *Mobius inversion*:

Given two sequences $\{a_n\}_{n \geq 1}$ and $\{b_n\}_{n \geq 1}$ such that $\sum_{d|n} a_d = b_n$, Mobius inversion says that $a_n = \sum_{d|n} \mu\left(\frac{n}{d}\right) b_d$, where the function μ is defined by $\mu(p) = -1$ for p prime, $\mu(p^m) = 0$ for $m \geq 2$, and $\mu(ab) = \mu(a)\mu(b)$ for a, b coprime. (For further explanation of the Mobius function μ and a proof of Mobius inversion, see **[2]**.) Letting $a_n = |P_n|$ and $b_n = k^n$ in (3), we get $|P_n| = \sum_{d|n} \mu\left(\frac{n}{d}\right) k^d$, so by (2), we obtain our main result:

THEOREM (generalized form of Fermat's little theorem). *For all positive integers n and k, $n | \sum_{d|n} \mu\left(\frac{n}{d}\right) k^d$.*

The method we used to prove this theorem can also be used to prove other such results. We applied the equation $n || P_n |$ to the particular function $f(z) = z^k$; but in fact, the same argument shows that $n || P_n |$ holds whenever P_n is the set of points of order n for *any* function f. Let f be any function from a set S to itself such that f^n has finitely many fixed points for every n. If $T(n)$ is the number of points fixed under f^n, then

$$n \,\Bigg|\, \sum_{d|n} \mu\left(\frac{n}{d}\right) T(d) \tag{4}$$

for all positive integers n.

A final question: We have shown that (4) is a necessary condition for the sequence $\{T(n)\}_{n \geq 1}$ to be of the form $T(n) = |\{z \in S | f^n(z) = z\}|$ for some function $f: S \to S$. Is (4) a sufficient condition as well? In other words, given any sequence of nonnegative integers $\{T(n)\}_{n \geq 1}$ satisfying (4), does there exist a function $f: S \to S$ such that f^n has $T(n)$ fixed points for every positive integer n?

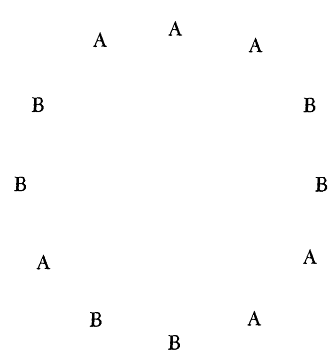
REFERENCES

1. R. Devaney, *A First Course in Chaotic Dynamical Systems*, Addison-Wesley, Reading, MA, 1992.
2. W. LeVeque, *Fundamentals of Number Theory*, Dover, Mineola, NY, 1977.

The Hexachordal Theorem: A Mathematical Look at Interval Relations in Twelve-Tone Composition

STEVEN K. BLAU
Physics Department
Ripon College
Ripon, WI 54971

Introduction Take the numbers 1 through 12 and divide them into two complementary six-member subsets A and B. Each subset has $\binom{6}{2} = 15$ pairs of members; the *interval* between any pair of numbers is the (positive) difference between them, subject to the equivalence $n \sim 12 - n$. So, for the pair $(1, 11)$ the interval may be said to be either 10 or 2. A convenient way to represent a complementary pair of interval sets, which makes visually apparent the equivalence just described, is to present a "clock face" with six A's and six B's, as follows:



In the case shown, $A = \{1, 4, 5, 8, 11, 12\}$ and $B = \{2, 3, 6, 7, 9, 10\}$. The interval between the A's at one o'clock and eleven o'clock is either 10 (counting clockwise) or 2 (counting anticlockwise). In what follows I will use the equivalence relation to describe all intervals as being between 1 and 6.

The *interval multiset* associated with a six-member set is the collection of 15 intervals determined by all possible pairs drawn from the set. For the set A shown, the interval multiset is $\{1, 1, 1, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 6\}$. As one can check directly, the set B has the *same* interval multiset. This is no accident—the equality of the interval multisets is the content of the hexachordal theorem.

THEOREM. *Let the numbers* 1 *through* 12 *be partitioned into any two complementary sets A and B, each with six elements. Then A and B have identical interval multisets.*

Before proving the theorem, we consider its musical meaning.

The musical context The hexachordal theorem was empirically discovered by composers working with Arnold Schönberg's twelve-tone method. In the musical context, the numbers 1 through 12 represent the twelve notes of the chromatic scale. Twelve-tone composers use a twelve note "row" containing all the members of the chromatic scale, arranged as they see fit, as a basis for their melodic composition. They often think of the first and second halves of the row as two complementary *hexachords*. In the musical context, "complementary" means that a note contained in one hexachord is not present in its companion. In the mathematical setting, complementary hexachords are complementary six-member subsets of $\{1, 2,, \ldots, 12\}$.

What I called an interval in the theorem would also be called an interval by composers: it represents the number of semitones (notes on a keyboard) needed to connect one note to another. Thus the hexachordal theorem is a statement about the interval structures of the two complementary halves of a twelve-tone row.

The hexachordal theorem was first proved by Milton Babbitt, a celebrated composer with a degree in mathematics, and David Lewin, then a graduate student in mathematics. In describing his proof with Lewin, Babbitt wrote, "We used topological methods. We hit this little problem with all kinds of heavy hammers, and we solved it" [**1**]. Later, Lewin (working on his own) and Ralph Fox constructed different proofs using group-theoretic methods. Lewin described his work on the hexachordal and related theorems in the *Journal of Music Theory* (see [**2**] and [**3**]). He did not include a proof of the hexachordal theorem, but sketched his work on related theorems, observing after one of his proofs [**2**]: "The mathematical reasoning by which I arrived at this result is not communicable to a reader who does not have considerable mathematical training." The proof of the hexachordal theorem that follows requires no advanced mathematics; it can be followed by musicians with little mathematical training, and it may interest mathematicians.

Challenged to give a four-word history of western music up to 1900, I would offer: "Modulations became more frequent." By the late 1800s in the works of Wagner, for example, there are sections where the modulations come so quickly that tonalities are established only for seconds before they change. Beethoven, by contrast, typically allowed tonalities to be established for a much greater time period before modulating to a new tonality. Early in his career, Arnold Schönberg wrote music following the Wagnerian line, but later decided that such music contained a sort of inconsistency. The momentum of music created since Bach was pointed toward an equality of the notes of the chromatic scale, but early twentieth-century composers still felt tied to the tonal conventions of earlier centuries. Schönberg advocated an "emancipation of dissonance," abandoning the notion that music must be conceived in terms of tonalities. A corollary was the freedom to use all twelve notes of the chromatic scale equally in composing, though Schönberg did occasionally write tonal music throughout his career.

The early pieces of Schönberg and his "Second Viennese School" were almost all short and aphoristic. The freedom enjoyed by composers was apparently so great that it was difficult to write on a grand scale music that maintained its internal logic. Schönberg sought a structure that would allow atonal composers to create coherent works on a large scale. His solution was the "method of composing with twelve tones related only to one another." One begins with a scale, or row, containing the twelve notes of the chromatic scale arranged in some fixed order. Then one manipulates the row in prescribed ways in composing. Part of the composer's skill is knowing when to break Schönberg's rules, but these rules do offer cohesion to atonal compositions. Some twelve-tone composers felt that their work would be further unified if the interval multisets implied by various subsets of their twelve-tone rows were identical.

Consequently, they empirically discovered the hexachordal theorem, as an offshoot of their search for unification. Babbitt and Lewin, and then others, proved the empirical result mathematically.

Proof of the theorem Given any pair of complementary hexachords A and B, displayed as a clock face, one can generate a new complementary pair by switching an A with a neighboring B. Indeed, *any* complementary pair of hexachords can be generated, through a sequence of switches, from the pairing with A's in positions 1–6 and B's in positions 7–12. This special pairing clearly satisfies the conclusion of the hexachordal theorem. Therefore, the hexachordal theorem is a consequence of the following lemma.

LEMMA. *Suppose that two complementary hexachords A and B have identical interval multisets. Then the hexachords obtained by switching a pair of neighboring A and B elements also have identical interval multisets.*

Proof. FIGURE 1 shows, at left, the clock face introduced earlier to represent a particular pair of hexachords. Switching the underlined entries (an A at eight o'clock and a B at seven o'clock) produces the the clock face on the right:

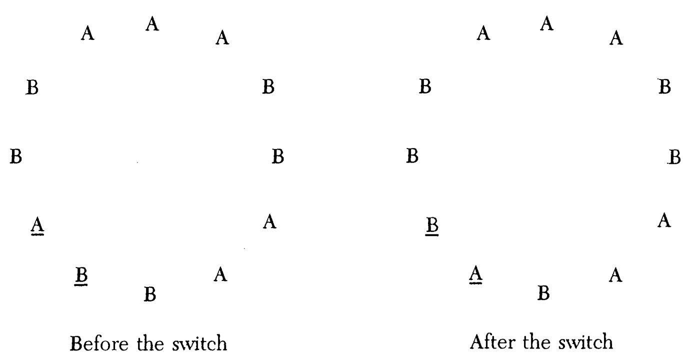


FIGURE 1
Switching an adjacent pair.

After identifying an adjacent pair of A and B elements to be switched, one may partition the remaining 10 elements into 5 pairs: for each integer n from 1 to 5, we consider the pair of elements lying n hours to either side of the switched A and B. There are four possibilities for the membership of each such pairs: two A's, two B's, or one of each (in either order). All four possibilities are illustrated in FIGURE 2 (the particular values of n correspond to the example illustrated above):

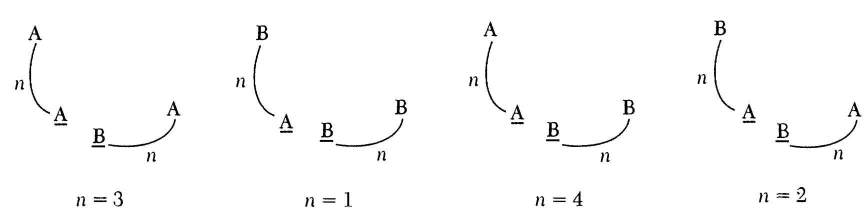


FIGURE 2
Pairs of elements lying n hours to either side of the switched A and B.

By considering the five pairs defined as above, one considers all the intervals possibly affected by the switch.

We consider four cases. If a pair contains two A's (as illustrated for $n = 3$), then the switch changes two intervals in the A hexachord: two original intervals, with lengths n and $n+1$, become two new intervals, with lengths $n+1$ and n, respectively. Hence the switch has no net effect on the A hexachord multiset (or on the B multiset, since the pair involves no B's). A similar argument shows that the switch has no effect if both elements in a pair are B's (as illustrated for $n = 1$).

If the element of the pair nearest to the switched A is an A, and the other element a B (as shown in FIGURE 2 for $n = 4$), then the switch does alter the interval multisets for the A and B hexachords, but in identical ways—in each multiset, an original interval of size n becomes a new interval of size $n+1$. The remaining case (illustrated for $n = 2$) is similar.

Thus, for all n, the switching operation has the same effect, if any, on both the A and the B interval multisets, and the proof is complete.

A generalization Nothing in the proof above relied on the (musical) facts that the chromatic scale has 12 notes and that composers considered dividing the chromatic scale into two hexachords of equal length. The following theorem can be proved by the same method, and was expressed in a different form by Lewin [3] for the case $N = 12$.

THEOREM. *Let the set* $\{1, 2, \dots, N\}$ *be partitioned into disjoint sets A and* B, *with size* a *and* $N - a$, *respectively. Define interval multisets for the A chord and the B chord as was done for the hexachordal theorem. Let* $A(i)$ *be the number of i's in the A interval multiset, and similarly for* $B(i)$. *Consider the special partition* $A_0 = \{1, 2, 3, \dots, a\}$, $B_0 = \{a+1, a+2, a+3, \dots, N\}$. *Then, for all* i, $A(i) - B(i) = A_0(i) - B_0(i)$.

Acknowledgment. It is a pleasure to acknowledge Karl Beres, Kurt Dietrich, Norm Loomer, Donald Passman, and Raymond Stahura, who carefully read this paper and offered many helpful suggestions.

REFERENCES

1. Milton Babbitt, in *Words About Music*, ed. by Dembski, Stephen and Straus, Joseph N., University of Wisconsin Press, Madison, WI, 1986, p. 105.
2. David Lewin, Intervallic relations between two collections of notes, *Journal of Music Theory* 3 (1959), pp. 298–301.
3. David Lewin, The intervallic content of a collection of notes, intervallic relations between a collection of notes and its complement: an application to Schoenberg's hexachordal pieces, *Journal of Music Theory* 4 (1960), pp. 98–101.

Euler Convergence: Probabilistic Considerations

GIORGIO METAFUNE
DIEGO PALLARA
CARLO SEMPI
Dipartimento di Matematica
"Ennio De Giorgi"
Università di Lecce
Lecce, Italy 73100

In mathematical analysis and its applications, the need sometimes arises to generalize the concept of limit of a sequence (or of sum of a series) in order to include cases in which the sequence (or the series) does not converge in the ordinary sense. One of the several methods devised bears Euler's name.

A sequence of real numbers $\{x_n\}$ is said to converge to x *in the sense of Euler* (or, usually, to be *Euler-convergent*) if there exists s in the open interval $(0, 1)$ such that

$$\lim_{n\to\infty} \sum_{j=0}^{n} \binom{n}{j} s^j (1-s)^{n-j} x_j = x.$$

The first key theorem on Euler convergence is as follows:

THEOREM 1. *If a sequence $\{x_n\}$ of real numbers converges to $x \in \mathbb{R}$ then it converges to the same limit in the sense of Euler for every $s \in (0, 1)$.*

Euler convergence is used in a natural way in the study of the asymptotic behavior of cyclic Markov chains (see [**5**], Chapter V). While teaching Markov chains using the approach of [**5**], we noticed that Theorem 1 could be proved in an elementary manner by relying only on the Chebyshev inequality, which appears in introductory courses in probability. Our proof provides a good chance to practice probabilistic reasoning. For a purely analytical proof of Theorem 1, one is often (e.g., in [**3**]) referred to [**4**]. But Hardy's book is hard reading for an undergraduate—even more so since it deals with series rather than with sequences.

A sequence of real numbers may be Euler-convergent without being convergent. Consider, for instance the sequence $\{x_n\}$ with $x_{2j} = 1$, $x_{2j+1} = 0$ $(j \geq 0)$, which does not converge. As for Euler convergence, notice that

$$\sum_{j=0}^{n} \binom{n}{j} s^j (1-s)^{n-j} x_j = \sum_{\substack{j=0 \\ j \text{ even}}}^{n} \binom{n}{j} s^j (1-s)^{n-j};$$

this latter sum represents the probability that a binomial random variable S_n takes an even value. Here $S_n = \sum_{j=1}^{n} X_j$, where the X_j's are Bernoulli random variables with $P(X_n = 1) = s$. By the binomial theorem,

$$\sum_{\substack{j=0 \\ j \text{ even}}}^{n} \binom{n}{j} s^j (1-s)^{n-j} + \sum_{\substack{j=0 \\ j \text{ odd}}}^{n} \binom{n}{j} s^j (1-s)^{n-j} = (s+1-s)^n = 1,$$

and

$$\sum_{\substack{j=0\\ j \text{ even}}}^{n} \binom{n}{j}(-s)^j(1-s)^{n-j} + \sum_{\substack{j=0\\ j \text{ odd}}}^{n} \binom{n}{j}(-s)^j(1-s)^{n-j} = (1-2s)^n.$$

Adding the last two equations yields

$$P\Bigg(\bigcup_{j \text{ even}} \{S_n = j\}\Bigg) = \sum_{\substack{j=0\\ j \text{ even}}}^{n} \binom{n}{j} s^j(1-s)^{n-j} = \frac{1+(1-2s)^n}{2};$$

this tends to $1/2$ as n tends to ∞.

Now we prove that Euler convergence is implied by ordinary convergence.

Proof of Theorem 1. Since

$$\sum_{j=0}^{n} \binom{n}{j} s^j(1-s)^{n-j} = (s+1-s)^n = 1,$$

it suffices to consider sequences that converge to zero. Let $\{x_n\}$ be such a sequence. For a fixed $\varepsilon \in (0, s)$ there exists $\nu \in \mathbb{N}$ such that $|x_n| < \varepsilon$ for every $n \geq \nu$. Now let $\lambda := \max\{|x_j| : j = 0, 1, \dots\}$ Then, for $n > \nu$, we have

$$\begin{aligned}
\left|\sum_{j=0}^{n} \binom{n}{j} s^j(1-s)^{n-j} x_j\right| &\leq \left|\sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} x_j\right| + \left|\sum_{j=\nu}^{n} \binom{n}{j} s^j(1-s)^{n-j} x_j\right| \\
&< \lambda \sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} + \varepsilon \sum_{j=\nu}^{n} \binom{n}{j} s^j(1-s)^{n-j} \\
&< \lambda \sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} + \varepsilon.
\end{aligned}$$

Notice that if $n > (\nu - 1)/(s - \varepsilon)$ then $ns - (\nu - 1) > n\varepsilon$ and, *a fortiori*, $ns - j > n\varepsilon$ for $j = 0, 1, \dots, \nu - 1$. Finally, observe that

$$\sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} = P\Bigg(\bigcup_{j=0}^{\nu-1} \{S_n = j\}\Bigg) \leq P(|S_n - ns| \geq n\varepsilon).$$

Now, since the variance of the binomial distribution is equal to $ns(1-s)$ if $s \in (0,1)$ is the probability of success, Chebyshev's inequality yields

$$\sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} \leq \frac{ns(1-s)}{n^2\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \to 0.$$

Therefore,

$$\left|\sum_{j=0}^{n} \binom{n}{j} s^j(1-s)^{n-j} x_j\right|$$

tends to zero as $n \to \infty$.

Alternatively, in a more analytical vein, one could avoid recourse to the Chebysbev inequality by considering the sum

$$\sum_{j=0}^{\nu-1} \binom{n}{j} s^j (1-s)^{n-j} \tag{1}$$

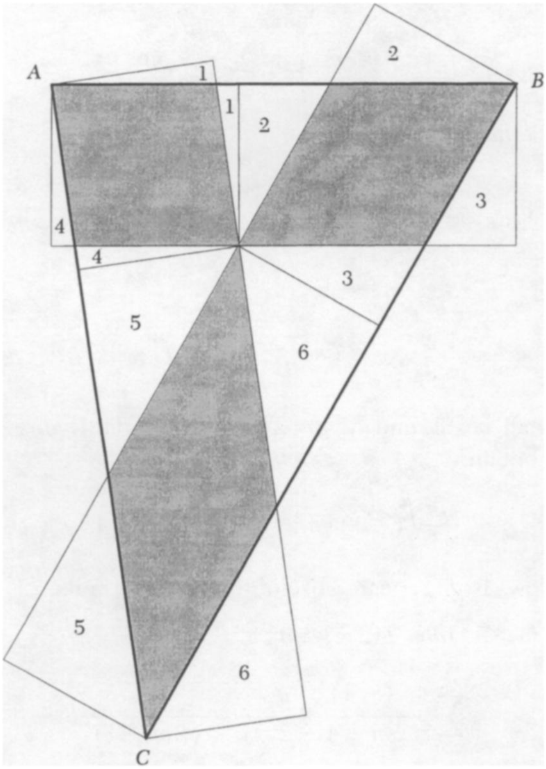
and noting that the number of terms (ν) is fixed, and that the binomial coefficient tends to infinity like n^j while the factor $(1-s)^{n-j}$ tends to zero exponentially. Thus each term in the sum (1), and hence the whole sum, tends to zero. Even this second approach has a probabilistic meaning: it is well-known (see, e.g., [**7**; Section 3.3]) that the terms of the binomial distribution increase in j from 0 to $(n+1)s$ and decrease when j runs from $(n+1)s+1$ to n. Since ν is fixed , the sum (1) represents, when n goes to infinity, the ever-decreasing probability of the tail of the binomial distribution. □

The first proof given above is an adaptation of the argument used by Bernstein [**1**] in his proof of the Weierstrass theorem on uniform approximation by polynomials of all functions that are continuous on a closed interval. A modern presentation can be found in the exercises of [**6**] or in [**2**].

REFERENCES

1. S. Bernstein, Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités, *Soob Charkov Mat. Obs.* 13 (1912), 1–2.
2. Y. S. Chow and H. Teicher, *Probability Theory, Independence, Interchangeability, Martingales*, Springer-Verlag, New York–Heidelberg–Berlin, 1978.
3. *Encylopaedia of Mathematics, Volume* 3, Kluwer, Dordrecht–Boston–London, 1989.
4. G. H. Hardy, *Divergent Series*, Clarendon, Oxford, UK, 1949.
5. J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, Van Nostrand, New York, NY, 1960 (reprinted, Springer-Verlag, New York, NY, 1976).
6. M. Loève, *Probability Theory*, Van Norstrand, New York, NY, 1963 (reprinted in two volumes, Springer-Verlag, New York–Heidelberg–Berlin, 1997).
7. E. Parzen, Modern Probability and Its Applications, Wiley, New York and London, 1960.

Proof Without Words: The Product of the Perimeter of a Triangle and Its Inradius Is Twice the Area of the Triangle



Regions bearing the same number are equal in area.

—GRACE LIN (STUDENT)
PHILLIPS EXETER ACADEMY
EXETER, NH 03833

Revisiting an Old Favorite: $\zeta(2m)$

NEVILLE ROBBINS
San Francisco State University
San Francisco, CA 94132

In this note, we present a new proof of the following result due to Chen **[6]**, namely:

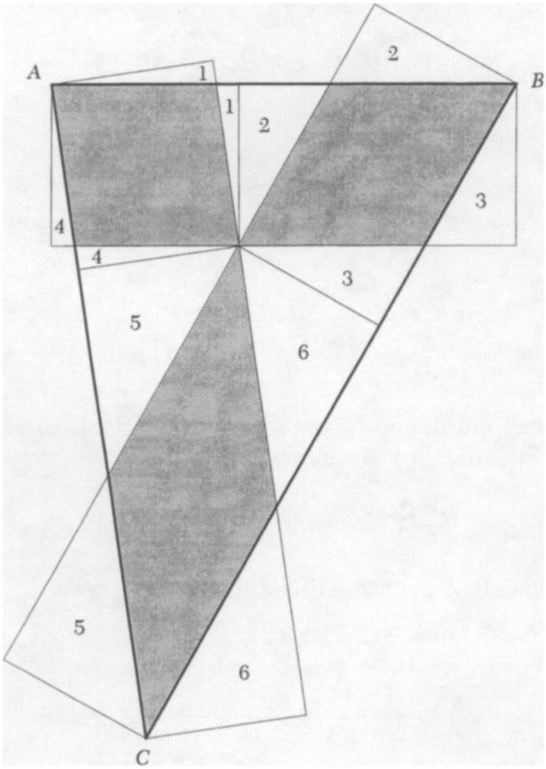
$$\zeta(2m) = a_m \pi^{2m}$$

where the a_m are rational numbers given recursively by:

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

Our method uses the Fourier expansion of x^{2m} on $[-\pi, \pi]$, integration by parts, and a fortuitous opportunity to interchange the order of a double summation. The steps are outlined below.

Proof Without Words: The Product of the Perimeter of a Triangle and Its Inradius Is Twice the Area of the Triangle



Regions bearing the same number are equal in area.

—GRACE LIN (STUDENT)
PHILLIPS EXETER ACADEMY
EXETER, NH 03833

Revisiting an Old Favorite: $\zeta(2m)$

NEVILLE ROBBINS
San Francisco State University
San Francisco, CA 94132

In this note, we present a new proof of the following result due to Chen [6], namely:

$$\zeta(2m) = a_m \pi^{2m}$$

where the a_m are rational numbers given recursively by:

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

Our method uses the Fourier expansion of x^{2m} on $[-\pi, \pi]$, integration by parts, and a fortuitous opportunity to interchange the order of a double summation. The steps are outlined below.

THEOREM 1. *If $m \geq 0$, let $f_m(x)$ be defined as follows: $f_m(x) = x^{2m}$ for $|x| \leq \pi$; $f_m(x + 2\pi) = f_m(x)$ for all x. Furthermore, let $f_m(x)$ have the Fourier cosine series expansion:*

$$f_m(x) = \tfrac{1}{2}c_{m,0} + \sum_{n=1}^{\infty} c_{m,n} \cos nx. \tag{1}$$

Then, for $m \geq 1$, we have:

$$c_{m,n} = 2(-1)^n (2m)! \sum_{j=1}^{m} \frac{(-1)^{j-1} \pi^{2(m-j)}}{(2m+1-2j)! n^{2j}}.$$

Proof. We have

$$c_{m,n} = \frac{1}{\pi} \int_{-\pi}^{\pi} x^{2m} \cos nx \, dx = \frac{2}{\pi} \int_0^{\pi} x^{2m} \cos nx \, dx$$

so that $c_{0,n} = 0$ for all $n \geq 1$, and $c_{m,0} = 2\pi^{2m}/(2m+1)$. If $m \geq 1$, then, integrating by parts twice, we obtain:

$$c_{m,n} = \frac{4m}{n^2} \left\{ (-1)^n \pi^{2m-2} - \left(\frac{2m-1}{2} \right) c_{m-1,n} \right\}.$$

The conclusion follows from repeated use of the latter formula. □

THEOREM 2. *If $m \geq 1$, then $\zeta(2m) = a_m \pi^{2m}$, where*

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

Proof. If we set $x = \pi$ in (1), divide by 2, simplify, and use the definition of $\zeta(2j)$, we obtain:

$$\sum_{j=1}^{m} \frac{(-1)^{j-1} 2(m-j)}{(2m+1-2j)!} \zeta(2j) = \frac{m\pi^{2m}}{(2m+1)!}.$$

The conclusion now follows by induction on m. □

Remark. Using Fourier series to obtain $\zeta(2m)$ in the cases $1 \leq m \leq 4$ is suggested in the textbooks [**3**] and [**8**]. Furthermore, it is known that

$$\zeta(2m) = (-1)^{m-1} 2^{2m-1} \pi^{2m} B_{2m}/(2m)!$$

where B_n denotes the nth Bernoulli number. Thus we have

$$a_m = (-1)^{m-1} 2^{2m-1} B_{2m}/(2m)!.$$

Also, as a consequence of the latter formula and Theorem 2, we obtain the following Bernoulli identity:

THEOREM 3. *If $m \geq 1$, then*

$$\sum_{j=1}^{m} 2^{2j-1} \binom{2m+1}{2j} B_{2j} = m.$$

Remarks. This note is an addition to the extensive existing literature regarding the evaluation of $\zeta(n)$, where n is an even natural number. (See the references below.) According to Ayoub [2], the evaluation of $\zeta(2)$ interested Leibniz, three of the Bernoullis, and Stirling. The first closed-form result, namely $\zeta(2) = \pi^2/6$, was given by Euler [7] in 1734.

REFERENCES

1. Tom M. Apostol, Another elementary proof of Euler's formula for $\zeta(2n)$, *Amer. Math. Monthly* 80 (1973), 425–431.
2. R. Ayoub, Euler and the zeta function, *Amer. Math. Monthly* 81 (1974), 1067–1086.
3. Robert G. Bartle, *The Elements of Real Analysis*, John Wiley & Sons, New York, NY, 1976, 341–343.
4. Bruce Berndt, Elementary evaluation of $\zeta(2n)$, this MAGAZINE 48 (1975), 148–154.
5. L. Carlitz, A recurrence formula for $\zeta(2n)$, *Proc. Amer. Math. Soc.* 12 (1961), 991–992.
6. M. P. Chen, An elementary evaluation of $\zeta(2n)$, *Chinese J. Math.* 3 (1975), 11–15.
7. L. Euler, De summis serierum reciprocarum, *Comment. Acad. Sci. Petropolit.* 7 (1734–35, 1740), 123–134; *Opera Omnia* Sr. 1, v. 14, 73–86.
8. W. Kaplan, *Advanced Calculus*, 3rd Ed. Addison-Wesley, Reading, MA, 1984, 495.
9. H.-T. Kuo, A recurrence formula for $\zeta(2n)$, *Bull. Amer. Math. Soc.* 55 (1949), 573–574.
10. E. L. Stark, The series $\Sigma_{k=1}^{\infty} k^{-s}$, $s = 2, 3, 4, \ldots$ once more, this MAGAZINE 47 (1974), 197–202.
11. G. T. Williams, A new method of evaluating $\zeta(2n)$, *Amer. Math. Monthly* 60 (1953), 19–25.
12. K. S. Williams, On $\Sigma_{k=1}^{\infty} 1/n^{2k}$, this MAGAZINE 44 (1971), 273–276.

Once More: Proportionality in the Non-Euclidean Plane

DIETER RUOFF
University of Regina
Regina, Saskatchewan
Canada S4S 0A2

In the article "Thales Meets Poincaré" (this MAGAZINE 70 (1997), pp. 185–195) David E. Dobbs demonstrated by means of Poincaré's model that an important Euclidean proportionality theorem (attributed to Thales) does not extend to hyperbolic geometry. As will be shown below, this same fact can also be established by an elementary argument which requires only the most modest background in hyperbolic geometry. Indeed the consequences of Euclid's familiar congruence theorems (which extend to hyperbolic geometry) will be used for a good portion of the argument. What is different from Euclidean geometry is related to the concept of parallelism and governed by the following axiom (see FIGURE 1):

A line l and a point B not on l determine an angle $\angle CBD$ such that precisely those rays from B that enter $\angle CBD$ intersect l.

Among the non-intersecting rays, $\overrightarrow{BC}$ and $\overrightarrow{BD}$ stand out by being asymptotic to l and are called *parallel* to l; we assign *improper points* (*or infinite points*) ω and τ to them in which they and all of their parallels meet. For F, the closest point to B on l, $|\angle FB\omega|$ is called the *angle of parallelism* for FB. Note that for FE^* with $FE^* < FB$ the angle of parallelism, $|\angle FE^*\omega|$ is larger than $|\angle FB\omega|$.

Remarks. This note is an addition to the extensive existing literature regarding the evaluation of $\zeta(n)$, where n is an even natural number. (See the references below.) According to Ayoub [**2**], the evaluation of $\zeta(2)$ interested Leibniz, three of the Bernoullis, and Stirling. The first closed-form result, namely $\zeta(2) = \pi^2/6$, was given by Euler [**7**] in 1734.

REFERENCES

1. Tom M. Apostol, Another elementary proof of Euler's formula for $\zeta(2n)$, *Amer. Math. Monthly* 80 (1973), 425–431.
2. R. Ayoub, Euler and the zeta function, *Amer. Math. Monthly* 81 (1974), 1067–1086.
3. Robert G. Bartle, *The Elements of Real Analysis*, John Wiley & Sons, New York, NY, 1976, 341–343.
4. Bruce Berndt, Elementary evaluation of $\zeta(2n)$, this MAGAZINE 48 (1975), 148–154.
5. L. Carlitz, A recurrence formula for $\zeta(2n)$, *Proc. Amer. Math. Soc.* 12 (1961), 991–992.
6. M. P. Chen, An elementary evaluation of $\zeta(2n)$, *Chinese J. Math.* 3 (1975), 11–15.
7. L. Euler, De summis serierum reciprocarum, *Comment. Acad. Sci. Petropolit.* 7 (1734–35, 1740), 123–134; *Opera Omnia* Sr. 1, v. 14, 73–86.
8. W. Kaplan, *Advanced Calculus*, 3rd Ed. Addison-Wesley, Reading, MA, 1984, 495.
9. H.-T. Kuo, A recurrence formula for $\zeta(2n)$, *Bull. Amer. Math. Soc.* 55 (1949), 573–574.
10. E. L. Stark, The series $\Sigma_{k=1}^{\infty} k^{-s}$, $s = 2, 3, 4, \ldots$ once more, this MAGAZINE 47 (1974), 197–202.
11. G. T. Williams, A new method of evaluating $\zeta(2n)$, *Amer. Math. Monthly* 60 (1953), 19–25.
12. K. S. Williams, On $\Sigma_{k=1}^{\infty} 1/n^{2k}$, this MAGAZINE 44 (1971), 273–276.

Once More: Proportionality in the Non-Euclidean Plane

DIETER RUOFF
University of Regina
Regina, Saskatchewan
Canada S4S 0A2

In the article "Thales Meets Poincaré" (this MAGAZINE 70 (1997), pp. 185–195) David E. Dobbs demonstrated by means of Poincaré's model that an important Euclidean proportionality theorem (attributed to Thales) does not extend to hyperbolic geometry. As will be shown below, this same fact can also be established by an elementary argument which requires only the most modest background in hyperbolic geometry. Indeed the consequences of Euclid's familiar congruence theorems (which extend to hyperbolic geometry) will be used for a good portion of the argument. What is different from Euclidean geometry is related to the concept of parallelism and governed by the following axiom (see FIGURE 1):

A line l and a point B not on l determine an angle $\angle CBD$ such that precisely those rays from B that enter $\angle CBD$ intersect l.

Among the non-intersecting rays, $\overrightarrow{BC}$ and $\overrightarrow{BD}$ stand out by being asymptotic to l and are called *parallel* to l; we assign *improper points* (*or infinite points*) ω and τ to them in which they and all of their parallels meet. For F, the closest point to B on l, $|\angle FB\omega|$ is called the *angle of parallelism* for FB. Note that for FE^* with $FE^* < FB$ the angle of parallelism, $|\angle FE^*\omega|$ is larger than $|\angle FB\omega|$.

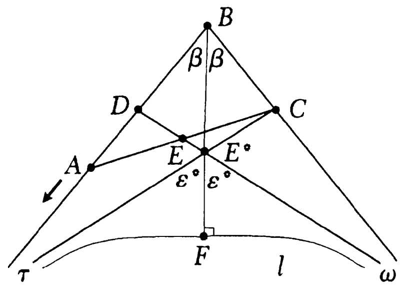


FIGURE 1

Remarks. (1) The reader who seeks somewhat more elaboration of the above can find it in most short introductions to hyperbolic geometry, such as [**2**, Ch. 9, sections 1–4] or [**1**, Ch. 16, sections 1–3].

(2) Dobbs uses a wider notion of parallelism than we, calling any two non-intersecting lines "parallel"; consequently he constructs his counterexample under less stringent conditions than are observed in the following theorem.

THEOREM. *In a hyperbolic plane it is possible to find an angle* $\angle BAC$ *together with parallel transversals* $\overrightarrow{BC}$, $\overrightarrow{DE}$ *such that* $\frac{AB}{AD} \neq \frac{BC}{DE}$.

Proof. (See FIGURE 1.) Consider a situation in which $BC = BD$. Call the improper common point of $\overrightarrow{BC}$ and its parallels ω, and that of $\overrightarrow{BD}$ and its parallels τ, and denote the intersecting point of $\overrightarrow{C\tau}$ and $\overrightarrow{D\omega}$ by E^*. Due to the symmetric position of C and D, and of ω and τ on $\angle CBD$, $CE^* = DE^*$ and the lines BE^*, $\omega\tau$ intersect at right angles in a point F. Because $FE^* < FB$, the angles of parallelism for these segments, $\varepsilon^* = |\angle FE^*\omega| = |\angle FE^*\tau|$ and $\beta = |\angle FB\omega| = |\angle FB\tau|$ satisfy $\beta < \varepsilon^*$. Consequently $|\angle CBD| = 2\beta < |\angle CE^*D| = 2\varepsilon^*$, which implies that in the isosceles triangles $\triangle BCD$ and $\triangle E^*CD$, $BC > DE^*$. Now if B, C, and D are fixed and A moves towards τ, then

$$\lim_{A\to\tau} \frac{AB}{AD} = 1 \quad \text{and} \quad \lim_{A\to\tau} \frac{BC}{DE} = \frac{BC}{DE^*} > 1,$$

which means that, for AB sufficiently large, $\frac{AB}{AD} \neq \frac{BC}{DE}$. ■

Remark. The assumption $BC = BD$ was made only to simplify the proof. In fact, as the author has shown, in every hyperbolic parallelogram BCE^*D with $\overrightarrow{BC} \| \overrightarrow{DE^*}$ and $\overrightarrow{BD} \| \overrightarrow{CE^*}$, side $BC >$ side DE^*.

REFERENCES

1. H. S. M. Coxeter, *Introduction to Geometry*, Wiley, New York, NY, 1969.
2. James R. Smart, *Modern Geometries*, Brooks/Cole, Monterey, CA, 1973.

Stacks, Bracketings, and *CG*-Arrangements

STEPHEN G. PENRICE
6-M Dorado Dr.
Morristown, NJ 07960

Bracketings and *CG*-arrangements In this MAGAZINE, Krusemeyer [3] showed a connection between a pair of "manifestations" of the Catalan sequence

$$c_n = \frac{1}{n+1}\binom{2n}{n}.$$

It is known that c_n counts the number of ways in which a non-associative product of $n+1$ operands can be grouped. These groupings are called *bracketings*, and we will denote the set of all such bracketings by B_n. It is also known that c_n counts the number of "eligible" arrangements of n pairs of empty parentheses, where an eligible arrangement is one in which no closing parenthesis precedes the corresponding opening parenthesis. These arrangements are called *CG-arrangements*, after John Conway and Richard Guy, and we will denote the set of all such arrangements by CG_n. Krusemeyer's paper is a response to Guy's suggestion that finding a "direct combinatorial comparison" between bracketings and CG-arrangements seemed "unlikely" (see [2]). Krusemeyer's solution was to give a recursive definition of a bijection between B_n and CG_n. This note takes a different approach, relating both manifestations of c_n to the *stack*, a data structure commonly used in computer science. The result is a bijection that is different from Krusemeyer's, but we shall show at the end of our discussion that a small modification in Krusemeyer's method produces our bijection.

Since both bracketings and CG-arrangements naturally involve parentheses, some confusion is possible. So we will adopt Krusemeyer's convention of using "floor" symbols for bracketings and "ceiling" symbols for CG-arrangements. However, we will deviate slightly from Krusemeyer's symbolism when depicting bracketings. We will use an operation symbol explicitly, and we will place an outer pair of floor symbols around each bracketed expression. The reasons for these seemingly superfluous symbols will become apparent when we define our bijection. To illustrate, the following table lists B_3 and CG_3:

B_3	CG_3
$\lfloor\lfloor a*\lfloor b*c\rfloor\rfloor*d\rfloor$	$\lceil\lceil\ \rceil\rceil\lceil\ \rceil$
$\lfloor\lfloor\lfloor a*b\rfloor*c\rfloor*d\rfloor$	$\lceil\ \rceil\lceil\ \rceil\lceil\ \rceil$
$\lfloor\lfloor a*b\rfloor*\lfloor c*d\rfloor\rfloor$	$\lceil\ \rceil\lceil\lceil\ \rceil\rceil$
$\lfloor a*\lfloor b*\lfloor c*d\rfloor\rfloor\rfloor$	$\lceil\lceil\ \lceil\ \rceil\ \rceil\rceil$
$\lfloor a*\lfloor\lfloor b*c\rfloor*d\rfloor\rfloor$	$\lceil\lceil\ \rceil\lceil\ \rceil\rceil$

Making the connection with stacks The definition of our bijection is simple to state; the alert reader might even guess it from the table above. However, we want to give ample motivation and show that the stack is at the heart of the bijection, a fact that is not apparent from the definition itself.

A stack is an ordered list of elements that has the *last-in-first-out* property. That is, at any given time, the only element in the list that can be accessed is the element that was added to the list most recently. An image that is often used to illustrate this concept is a stack of cafeteria trays resting on a platform supported by a spring: as each tray is added to the top, the trays already present are pushed farther down in the stack, and when a customer takes a tray, she must take the one on top. There are two operations defined on stacks: *push* and *pop*. Push is the operation that adds a new element to the top of the stack, and pop is the operation that removes the element at the top of the stack and makes it available to the program that is using the stack. More information about stacks can be found in any elementary data structures book, such as **[1]**.

The connection between stacks and CG-arrangements is easy to see. Consider a computer program that uses a stack. Assume that the stack is empty at the beginning of the program's execution, that the program never tries to pop the stack when it is empty, and that the stack is empty when the program has finished its execution. Then for each element that was put into the stack during the computation, there was a corresponding push and a corresponding pop, and the push preceded the pop. Therefore, if one observes the entire computation and marks a "$\lceil$" at every push and a "$\rceil$" at every pop, the result will be a CG-arrangement.

To see the connection between bracketings and stacks, let us consider a computer's evaluation of a bracketed expression. We will assume that, besides a stack, our computer has a special memory location called an *accumulator* to store values. The computer will parse the expression one symbol at a time, proceeding from left to right and using the following rules. It is assumed the computer knows n, the number of binary operations that will be performed.

1. Ignore left floor symbols.
2. If the current symbol is an operand, place the corresponding value in the accumulator.
3. If the current symbol is an operation symbol, push the current contents of the accumulator onto the stack.
4. If the current symbol is a right floor symbol, pop the stack and compute the product of the popped value and the current contents of the accumulator, placing the results in the accumulator.
5. Halt after the nth right floor symbol has been processed. The desired value will be in the accumulator, and the stack will be empty.

It is not hard to see that this procedure evaluates the products in the manner indicated by the bracketing.

Note that, in the above procedure, a push occurs if and only if an operation symbol is processed and a pop occurs if and only if a right floor symbol is processed. This, together with our remarks on the connection between stacks and CG-arrangements, provides motivation for the following mapping G_n from B_n to CG_n.

Let β be an element of B_n. Then we define $G_n(\beta)$ to be the element of CG_n obtained by deleting all left floor symbols and operand symbols from β, replacing each operation symbol in β with a left ceiling symbol, and replacing each right floor symbol in β with a right ceiling symbol. The correspondence G_3 is illustrated in the table above, and can be shown by inspection to be different from Krusemeyer's correspondence.

Computing the inverse We have not yet shown that G_n is a bijection, but we shall do so by defining a mapping from CG_n to B_n that is the inverse of G_n. We will define G_n^{-1} by describing an algorithm that computes it. Appropriately, this algorithm uses a stack.

In the computation of G_n^{-1}, we shall consider the elements of B_n and CG_n as strings of characters. Specifically, these characters will be the left and right floor and ceiling symbols, the binary operation symbol, and an alphabet of operands $a_1, a_2, \dots, a_{n+1}$. We will assume that the accumulator and the stack can hold strings of arbitrary length. It will also be convenient to refer to counter variables i and j that tells us, respectively, how many left and right ceiling symbols have been processed. The algorithm is as follows:

1. Place a_1 in the accumulator, and initialize i and j to 0.
2. Process the input string, which is a CG-arrangement, one symbol at a time from left to right according to the following rules:
   A. If the current input symbol is a left ceiling symbol
      i) Place an operator symbol $(*)$ at the right end of the string that is currently in the accumulator and push the resulting string onto the stack;
      ii) Increment i and place a_{i+1} in the accumulator.
   B. If the current input symbol is a right ceiling symbol
      i) Pop the stack and concatenate the popped string with the string currently in the accumulator, putting the popped string on the left and placing the resulting string in the accumulator;
      ii) Increment j and place a pair of floor symbols around the string in the accumulator, putting the result in the accumulator.
3. Output the contents of the accumulator and halt if $j = n$; otherwise go to the next input character and repeat 2.

We will not do so here, but it can be argued fairly easily that after each execution of step 2B, the accumulator contains a bracketed product, so that, in particular, the output is an element of B_n. It is easy to see that in the above algorithm an operation symbol is produced precisely when a left ceiling symbol is processed from the input string, and a right floor symbol is produced precisely when a right ceiling symbol is processed from the input string. Therefore, once we prove the following claim, it will be apparent that any input string can be recovered from the output by applying G_n. This proves that G_n^{-1} is 1-to-1 and that G_n is a bijection between B_n and CG_n.

CLAIM. *If γ and δ are characters in an output string from the algorithm above such that γ and δ are each either "$*$" or "$\rfloor$" and if γ was produced before δ, then γ is to the left of δ in the output string.*

Proof. First consider the case where γ is in the string that is in the accumulator at the time δ is produced. Since $*$ and $\rfloor$ are both added to the right of the string in the accumulator, γ will be to the left of δ in the resulting string. Since strings are never broken up or reordered in the algorithm, γ will be to the left of δ in the output string as well.

The only other possibility is that γ is in a string, say Σ, that is in the stack at the time δ is produced. Consider the first time after δ is produced that Σ is popped from the stack. The string that is in the accumulator at that time must contain δ, because otherwise δ would be in a string that is farther down in the stack, which could only happen only if Σ had already been popped. By the algorithm, when Σ is popped it is concatenated to the left of the string in the accumulator. Therefore γ will be to the left of δ in the resulting string and, by the remarks above, in the output string as well.

Computing the inverse recursively We close with a brief explanation of how a variation on Krusemeyer's method can be used to define G_n^{-1} recursively. The key insight driving Krusemeyer's approach is that if $\alpha \in CG_n$, then α can be written uniquely as $\alpha = \lceil \beta \rceil \gamma$, where $\beta \in CG_k$ and $\gamma \in CG_{n-k-1}$, for some k with $0 \leq k \leq n-1$. (Note that CG_0 consists of the empty string.) This representation can be found by scanning α from left to right, one character at a time, until reaching the first point where the number of left ceiling symbols scanned equals the number of right ceiling symbols scanned; such a point exists since these will surely be equal when all of α has been scanned. Of course, the decision to scan from left to right is completely arbitrary. Scanning from right to left, one can just as easily see that α can be written uniquely as $\alpha = \beta \lceil \gamma \rceil$, where $\beta \in CG_k$ and $\gamma \in CG_{n-k-1}$, for some k (although β and γ may be different from those given above). We will use this fact to define G_n^{-1} recursively.

The base of the recursive definition is simply to declare that G_0^{-1} maps the empty string to a single operand. Suppose that G_k^{-1} has been defined for $0 \leq k \leq n-1$. Then if $\alpha \in CG_n$, $G_n^{-1}(\alpha) = \lfloor G_k^{-1}(\beta) * G_{n-k-1}^{-1}(\gamma) \rfloor$, where $\alpha = \beta \lceil \gamma \rceil$ is the representation mentioned above, $a_1, \ldots, a_{k+1}$ is the alphabet of operands used in $G_k^{-1}(\beta)$, and $a_{k+2}, \ldots, a_{n+1}$ is the alphabet of operands used in $G_{n-k-1}^{-1}(\gamma)$. Using a proof by induction, one can show that this agrees with the earlier definition of G_n^{-1} for all positive integers n.

Acknowledgment. The observation that the bijection in this note and the one in Krusemeyer's paper share the connection described in the last section is due to an anonymous referee.

REFERENCES

1. Alfred V. Aho, John E. Hopcroft, and Jeffrey D. Ullman, *Data Structures and Algorithms*, Addison-Wesley, Reading, MA, 1983.
2. Richard K. Guy, The second strong law of small numbers, this MAGAZINE 63 (1990), 3–20.
3. Mark Krusemeyer, A parenthetical note (to a paper of Guy), this MAGAZINE 69 (1996), 257–260.

Period Doubling Near the Feigenbaum Limit

Fourteen lines accommodate
The points I've picked to illustrate.

In some systems you will find
  Orbits moving toward a station,
Then show themselves to have a mind
  To move no more on iteration.

But tweak an additive parameter
  And where before you saw them stall,
    Now they, surprisingly, begin
      A two-step foxtrot on the floor,
Like iambs in a strict tetrameter,
  Or like an active ping pong ball.
    More tweaks, and doubling comes again;
      Yet more, redoubling as before.

—J. D. MEMORY
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, NC 27695-8021

Computing the inverse recursively We close with a brief explanation of how a variation on Krusemeyer's method can be used to define G_n^{-1} recursively. The key insight driving Krusemeyer's approach is that if $\alpha \in CG_n$, then α can be written uniquely as $\alpha = \lceil \beta \rceil \gamma$, where $\beta \in CG_k$ and $\gamma \in CG_{n-k-1}$, for some k with $0 \le k \le n-1$. (Note that CG_0 consists of the empty string.) This representation can be found by scanning α from left to right, one character at a time, until reaching the first point where the number of left ceiling symbols scanned equals the number of right ceiling symbols scanned; such a point exists since these will surely be equal when all of α has been scanned. Of course, the decision to scan from left to right is completely arbitrary. Scanning from right to left, one can just as easily see that α can be written uniquely as $\alpha = \beta \lceil \gamma \rceil$, where $\beta \in CG_k$ and $\gamma \in CG_{n-k-1}$, for some k (although β and γ may be different from those given above). We will use this fact to define G_n^{-1} recursively.

The base of the recursive definition is simply to declare that G_0^{-1} maps the empty string to a single operand. Suppose that G_k^{-1} has been defined for $0 \le k \le n-1$. Then if $\alpha \in CG_n$, $G_n^{-1}(\alpha) = \lfloor G_k^{-1}(\beta) * G_{n-k-1}^{-1}(\gamma) \rfloor$, where $\alpha = \beta \lceil \gamma \rceil$ is the representation mentioned above, $a_1, \ldots, a_{k+1}$ is the alphabet of operands used in $G_k^{-1}(\beta)$, and $a_{k+2}, \ldots, a_{n+1}$ is the alphabet of operands used in $G_{n-k-1}^{-1}(\gamma)$. Using a proof by induction, one can show that this agrees with the earlier definition of G_n^{-1} for all positive integers n.

Acknowledgment. The observation that the bijection in this note and the one in Krusemeyer's paper share the connection described in the last section is due to an anonymous referee.

REFERENCES

1. Alfred V. Aho, John E. Hopcroft, and Jeffrey D. Ullman, *Data Structures and Algorithms*, Addison-Wesley, Reading, MA, 1983.
2. Richard K. Guy, The second strong law of small numbers, this MAGAZINE 63 (1990), 3–20.
3. Mark Krusemeyer, A parenthetical note (to a paper of Guy), this MAGAZINE 69 (1996), 257–260.

Period Doubling Near the Feigenbaum Limit

Fourteen lines accommodate
The points I've picked to illustrate.

In some systems you will find
  Orbits moving toward a station,
Then show themselves to have a mind
  To move no more on iteration.

But tweak an additive parameter
  And where before you saw them stall,
    Now they, surprisingly, begin
      A two-step foxtrot on the floor,
Like iambs in a strict tetrameter,
  Or like an active ping pong ball.
    More tweaks, and doubling comes again;
      Yet more, redoubling as before.

—J. D. MEMORY
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, NC 27695-8021

PROBLEMS

GEORGE T. GILBERT, *Editor*
Texas Christian University

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

Proposals

To be considered for publication, solutions should be received by March 1, 2000.

1579. *Proposed by T. S. Michael and William P. Wardlaw, U. S. Naval Academy, Annapolis, Maryland.*

For each positive integer n, find a set of positive integers whose sum is n and whose product is as large as possible.

(If repetitions are allowed with $n = 1979$, we obtain problem A-1 from the 1979 Putnam Competition.)

1580. *Proposed by Jerrold W. Grossman, Oakland University, Rochester, Michigan.*

Consider the following variation of the game of Nim. A position consists of piles of stones, with $n_i \geq 1$ stones in pile i. Two players alternately move by choosing one of the piles, permanently removing one or more stones from that pile, and, optionally, redistributing some (or all) of the remaining stones in that pile to one or more of the other remaining piles. (Once a pile is gone, no stones can be added to it.) The player who removes the last stone wins. Find a strategy for winning this game; in particular, determine which vectors of positive integers $(n_1, n_2, \ldots, n_k)$ allow the first player to win and which vectors allow the second player to win.

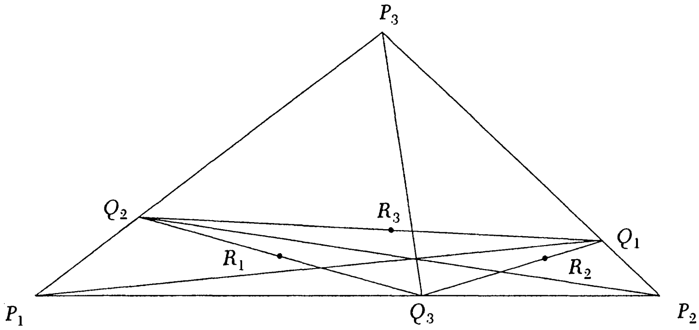
1581. *Proposed by Herbert Gülicher, Westfälische Wilhelms–Universität, Münster, Germany.*

Consider $\triangle P_1P_2P_3$ and points Q_1, Q_2, Q_3 in the interior of sides P_2P_3, P_1P_3, P_1P_2, respectively, such that P_1Q_1, P_2Q_2, and P_3Q_3 are concurrent (i.e., $\triangle Q_1Q_2Q_3$ is a cevian triangle of $\triangle P_1P_2P_3$). Let R_1, R_2, R_3 be in the interior of sides Q_2Q_3, Q_1Q_3, Q_1Q_2, respectively. Prove that the lines P_1R_1, P_2R_2, and P_3R_3 are concurrent if and only if the lines Q_1R_1, Q_2R_2, and Q_3R_3 are concurrent.

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

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

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



1582. *Proposed by Western Maryland College Problems Group, Westminster, Maryland.*

Let teams A and B play a series of games. Each game has three possible outcomes: A wins with probability p, B wins with probability q, or they tie with probability $r = 1 - p - q$. The series ends when one team has won two more games than the other, that team being declared the winner of the series.

(a) Find the probability that A wins the series.

(b) Let X be the number of games in the series. Find the probability function for X and its expected value.

1583. *Proposed by George T. Gilbert, Texas Christian University, Fort Worth, Texas.*

Classify all pairs of complex numbers a and b for which $a^2 - b^2$, $a^3 - b^3$, and $a^5 - b^5$ are rational numbers.

Quickies

Answers to the Quickies are on page 331

Q893. *Proposed by Glenn G. Chappell, Southeast Missouri State University, Cape Girardeau, Missouri, and Frank Sheeran, Speedcore, Lawrence, Kansas.*

The index of a book lists every page on which certain words appear. To save space these are listed in ranges; for example, if a word occurs on pages 1, 2, 3, 5, 8, and 9, then its index entry contains 3 ranges: 1–3, 5, 8–9.

A certain word appears on each page of an n-page book ($n \geq 1$) independently with probability p. Find the expected number of ranges in its index entry.

Q894. *Proposed by Harry Tamvakis, University of Pennsylvania, Philadelphia, Pennsylvania.*

Find the geometric locus of orthocenters of triangles inscribed in a fixed circle $\mathscr{C}$.

Solutions

Volume Cut by a Hypersurface in the Unit Cube **October 1998**

1554. *Proposed by Howard Cary Morris, Germantown, Tennessee.*

For $0 \leq r \leq 1$, find the volume $V_n(r)$ of

$$\left\{(x_1, \ldots, x_n) \in [0,1]^n : \prod_{i=1}^{n} x_i \leq r\right\}.$$

Solution by Robert A. Agnew, Deerfield, Illinois.

We will show that $V_n(r) = r\sum_{k=0}^{n-1}\frac{(-\ln r)^k}{k!}$ (with $V_n(0) = 0$).

We have $V_1(r) = r$, verifying our formula for $n = 1$, and

$$V_n(r) = \int_0^r dx + \int_r^1 V_{n-1}(r/x)\, dx = r + r\int_r^1 V_{n-1}(u)\frac{du}{u^2}$$

for $n \geq 2$. Suppose our formula is valid up through $n-1$. Then

$$V_n(r) = r + r\int_r^1 \sum_{k=0}^{n-2} \frac{(-\ln u)^k}{k!}\frac{du}{u}$$

$$= r - r\sum_{k=0}^{n-2}\frac{(-1)^k(\ln r)^{k+1}}{(k+1)!} = r\sum_{k=0}^{n-1}\frac{(-\ln r)^k}{k!}.$$

Also solved by Michael Andreoli, Nirdosh Bhatnagar, Jean Bogaert (Belgium), Larry W. Cusick, Paul Deiermann, Daniele Donini (Italy), Hans Kappus (Switzerland), Kathleen E. Lewis, Robert Patenaude, Heinz-Jürgen Seiffert (Germany), TAMUK Problem Solvers, John Towers, Tiberiu V. Trif (Romania), University of Arizona Problem Solving Group, Western Maryland College Problems Group, Michael Woltermann, Paul J. Zwier, and the proposer. There was one incorrect solution.

An Equation with Logarithmic Exponents **October 1998**

1555. *Proposed by Mihály Bencze, Braşov, Romania.*

Given a, b, and c_k, $k = 1,2,\dots,n$, all greater than 1, find all real solutions x of

$$\sum_{k=1}^{n}(x+a)^{\log_a c_k} = \sum_{k=1}^{n}(x+b)^{\log_b c_k}.$$

Solution by Matt Foss, North Hennepin Community College, Brooklyn Park, Minnesota.

If $a \neq b$ the only real solution is $x = 0$. Without loss of generality assume $1 < a < b$ and consider the functions $f(x) := \ln(x+a)/\ln a$ and $g(x) := \ln(x+b)/\ln b$. Observe that $f(0) = g(0) = 1$ and that $f'(x) > g'(x)$ for all x in their common domain. Thus $f(x) < g(x)$ for $x < 0$ and $f(x) > g(x)$ for $x > 0$. In the former case, for $c > 1$,

$$(x+a)^{\log_a c} = c^{\ln(x+a)/\ln a} < c^{\ln(x+b)/\ln b} = (x+b)^{\log_b c}.$$

It follows that

$$\sum_{k=1}^{n}(x+a)^{\log_a c_k} < \sum_{k=1}^{n}(x+b)^{\log_b c_k}$$

for $x < 0$. Similarly,

$$\sum_{k=1}^{n}(x+a)^{\log_a c_k} > \sum_{k=1}^{n}(x+b)^{\log_b c_k}$$

for $x > 0$.

Also solved by Larry W. Cusick, Richard F. Melka and Qingchuan Yao, Fary Sami, TAMUK Problem Solvers, Western Maryland College Problems Group, Michael Woltermann, and the proposer.

Inequality for a Sum of Reciprocals of Partial Sums **October 1998**

1556. *Proposed by Gregory Galperin and Hillel Gauchman, Eastern Illinois University, Charleston, Illinois.*

Let $a_1, \dots, a_n$ be positive numbers with $a_1 a_2 \cdots a_n = 1$. Set $x_i = (\sum_{k=1}^n a_k) - a_i$ for each $i = 1, \dots, n$. Prove that

$$\sum_{i=1}^{n} \frac{1}{1+x_i} \le 1.$$

I. *Solution by Mordechai Falkowitz, Hamilton, Ontario, Canada.*

Observe that

$$\frac{1}{1+1/a_1} + \frac{1}{1+a_1} = 1,$$

so that we may assume $n \ge 3$. We also have equality for $a_1 = a_2 = \cdots = a_n = 1$. In general, we have $x_i > a_n$ for $1 \le i \le n-1$. By the arithmetic mean–geometric mean inequality

$$x_n = a_1 + a_2 + \cdots + a_{n-1} \ge (n-1)(a_1 a_2 \cdots a_{n-1})^{\frac{1}{n-1}} = \frac{n-1}{\sqrt[n-1]{a_n}}.$$

For a_n sufficiently large, say $a_n \ge n^{n-1}$, we have $1 + a_n > \sqrt[n-1]{a_n} + n - 1$. In this case, it follows that

$$\sum_{i=1}^{n} \frac{1}{1+x_i} < \frac{n-1}{1+a_n} + \frac{1}{1+(n-1)/\sqrt[n-1]{a_n}} < 1.$$

Thus, it suffices to prove that the maximum value of the sum on the compact domain $\{(a_1, a_2, \dots, a_n) \in [0, n^{n-1}]^n : a_1 a_2 \cdots a_n = 1\}$ occurs when $a_1 = a_2 = \cdots = a_n = 1$. There is no loss of generality in assuming $a_1 \le a_2 \le \cdots \le a_n$.

Suppose that $a_i < a_{i+1}$ for some $i \le n-2$. For $S > 0$, the arithmetic mean–geometric mean inequality implies

$$\frac{1}{S + a_i + a_{i+1}} < \frac{1}{S + 2\sqrt{a_i a_{i+1}}}.$$

Now note that

$$1 + a_1 + \cdots + a_{i-2} + a_{i+2} + \cdots + a_n \ge 1 + a_n > \sqrt{a_i a_{i+1}}.$$

For $S > \sqrt{a_i a_{i+1}}$, straightforward algebra implies

$$\frac{1}{S+a_i} + \frac{1}{S+a_{i+1}} < \frac{2}{S + \sqrt{a_i a_{i+1}}}.$$

Therefore, $(a_1, a_2, \dots, a_n)$ cannot yield the maximum value.

Now suppose that $a_1 = a_2 = \cdots = a_{n-1} < a_n$. Set $a_1 = a$ and note that it is less than 1. Then

$$\sum_{i=1}^{n} \frac{1}{1+x_i} = \frac{n-1}{1+(n-2)a+1/a^{n-1}} + \frac{1}{1+(n-1)a} < 1$$

is equivalent to

$$1 + (n-2)a - (n-2)a^2 - 1/a^{n-2} < 0.$$

The derivative with respective to a of $h(a) = 1 + (n-2)a - (n-2)a^2 - 1/a^{n-2}$ is $(n-2)(1 - 2a + 1/a^{n-1}) > 0$ on $(0,1)$, so that $h(a) < h(1) = 0$ as desired.

II. *Solution by the proposers*. Set $a_i = b_i^n$ for $i = 1, \dots, n$. Application of the arithmetic mean–geometric mean inequality yields

$$\begin{aligned} b_1 b_2 \cdots b_{n-1}(b_1 + \cdots + b_{n-1}) &= \sum_{i=1}^{n-1} \sqrt[n]{b_1^n b_2^n \cdots b_{n-1}^n b_i^n} \\ &\le \sum_{i=1}^{n-1} \frac{b_1^n + b_2^n + \cdots + b_{n-1}^n + b_i^n}{n} \\ &= b_1^n + b_2^n + \cdots + b_{n-1}^n. \end{aligned}$$

From this inequality and $b_1 b_2 \cdots b_n = 1$, we find that

$$\begin{aligned} \frac{1}{1+x_n} &= \frac{1}{1 + b_1^n + b_2^n + \cdots + b_{n-1}^n} \\ &\le \frac{1}{1 + b_1 b_2 \cdots b_{n-1}(b_1 + \cdots + b_{n-1})} \\ &= \frac{b_n}{b_1 + \cdots + b_n}. \end{aligned}$$

Similar reasoning implies that

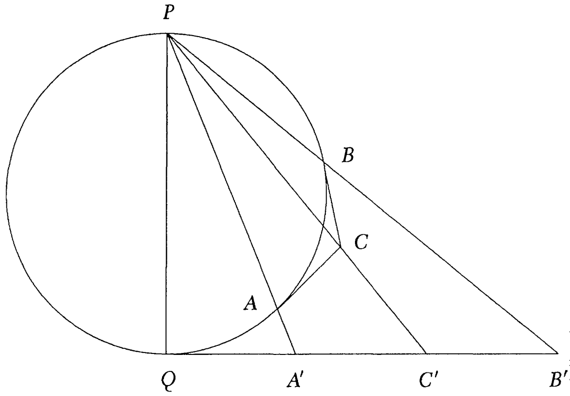
$$\sum_{i=1}^{n} \frac{1}{1+x_i} \le \sum_{i=1}^{n} \frac{b_i}{b_1 + \cdots + b_n} = 1.$$

There were two incorrect solutions.

A Circle, Three Tangents, and Three Secants **October 1998**

1557. *Proposed by Peter Y. Woo, Biola University, La Mirada, California.*

Let PQ be a diameter of a circle, with A and B two distinct points on the circle on the same side of PQ. Let C be the intersection of the tangents to the circle at A and B. Let the tangent to the circle at Q meet PA, PB, and PC at A', B', and C', respectively. Prove that C' is the midpoint of $A'B'$.



I. *Solution by Vince McGarry, Austin Community College, Austin, Texas.*

It suffices to prove the statement for the unit circle with $P = (0,1)$. Without loss of generality, we may assume $A = (\cos(\theta - \alpha), \sin(\theta - \alpha))$ and $B = (\cos(\theta + \alpha),$

$\sin(\theta+\alpha)$) with $|\theta \pm \alpha| < \pi/2$, so that $C = (\sec\alpha\cos\theta, \sec\alpha\sin\theta)$. Because the line through P and a point (r, s), $s \neq 1$, intersects $y = -1$, the tangent at Q, at the point $(2r/(1-s), -1)$, we have

$$A' = \left(\frac{2\cos(\theta-\alpha)}{1-\sin(\theta-\alpha)}, -1\right) = \left(\frac{2(1+\sin(\theta-\alpha))}{\cos(\theta-\alpha)}, -1\right),$$

$$B' = \left(\frac{2\cos(\theta+\alpha)}{1-\sin(\theta+\alpha)}, -1\right) = \left(\frac{2(1+\sin(\theta+\alpha))}{\cos(\theta+\alpha)}, -1\right),$$

$$C' = \left(\frac{2\sec\alpha\cos\theta}{1-\sec\alpha\sin\theta}, -1\right) = \left(\frac{2\cos\theta}{\cos\alpha-\sin\theta}, -1\right).$$

Beginning with the expression for the x-coordinate of the midpoint derived from A' and B',

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

which shows that C' is the midpoint of A' and B'.

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

If we invert the figure in the circle with center at P and radius PQ, we see that the circle with diameter PQ inverts to the line tangent to the circle at Q. The circle through A and B with center C, which is orthogonal to the circle with diameter PQ, inverts to a circle with center C' and radius $C'A' = C'B'$ because orthogonality must be preserved and because the center of the inverse must lie on the line PC.

Also solved by Alfredo Aguirre, Reza Akhlaghi, Michel Bataille (France), J. C. Binz (Switzerland), Mansur Boase (student, England), Jean Bogaert (Belgium), Stan Byrd, María Ascensión López Chamorro (Spain), Larry W. Cusick, Clayton W. Dodge (professor emeritus), Daniele Donini (Italy), Matt Foss, Jiro Fukuta (professor emeritus, Japan), Henok Getnet, R. Govindaraj (India), Bradley Gunsalus (student), Erhard Heil (Germany), Peter Hohler (Switzerland), Hans Kappus (Switzerland), Victor Y. Kutsenok, Ho-joo Lee (student, South Korea), Atar Sen Mittal, Jawad Sadek, Harry Sedinger, Achilleas Sinefakopoulos (Greece), TAMUK Problem Solvers, Michael Vowe (Switzerland), Thomas C. Wales, Michael Woltermann, David Zhu, Paul J. Zwier, and the proposer.

Primes in a Recursively Defined Sequence **October 1998**

1558. *Proposed by Mansur Boase, student, St. Paul's School, London, England.*

Let the sequence $(K_n)_{n\geq 1}$ be defined by $K_1 = 2$, $K_2 = 8$, and $K_{n+2} = 3K_{n+1} - K_n + 5(-1)^n$. Prove that if K_n is prime, then n must be a power of 3.

(A typographical error in a subscript in the original statement has been corrected.)

Solution by TAMUK Problem Solvers, Texas A & M University–Kingsville, Kingsville, Texas.

Let $\alpha = (3+\sqrt{5})/2$ and $\beta = (3-\sqrt{5})/2$. Then $K_n = \alpha^n + \beta^n + (-1)^n$. This can easily be checked by induction, using the facts that $\alpha\beta = 1$ and $\alpha + \beta = 3$, or by observing that $H_n := K_n - (-1)^n$ satisfies the homogeneous recurrence relation $H_{n+2} = 3H_{n+1} - H_n$, with characteristic equation $x^2 - 3x + 1 = 0$, whose zeros are α and β. Thus $H_n = c_1\alpha^n + c_2\beta^n$, and with $K_1 = 2$, $K_2 = 8$, we get K_n as given above.

Because $\alpha\beta = 1$, we see that $(\alpha^n + \beta^n - 1)(\alpha^n + \beta^n + 1) = \alpha^{2n} + \beta^{2n} + 1 = K_{2n}$, which implies that $K_n | K_{2n}$. Further observe that

$$(\alpha^m + \beta^m - 1)(\alpha^{n+m} + \beta^{n+m} + \alpha^{n+2m} + \beta^{n+2m}) = \alpha^n + \beta^n + \alpha^{n+3m} + \beta^{n+3m}.$$

This implies that, for odd m (so that the ± 1 cancel), $K_m | K_n + K_{n+3m}$. Hence it follows that if K_m divides K_n it also divides K_{n+3m}. Trodding down the sequence this way, with step size $3m$, we find that, for odd m, if $K_m | K_n$, then $K_m | K_{n+(3m)b}$ for any non-negative integer b. Hence when we take $m = 3^a = n$, we find that

$$K_{3^a} | K_{3^a+(3\cdot 3^a)b} = K_{3^a(1+3b)},$$

and because $K_{3^a} | K_{2\cdot 3^a}$, we also get that

$$K_{3^a} | K_{2\cdot 3^a+(3\cdot 3^a)b} = K_{3^a(2+3b)}.$$

From either the formula or the recursion, it follows that $(K_n)_{n\geq 1}$ is strictly increasing. Now since any positive integer n that is not a power of 3 can be written as either $3^a(1+3b)$ or $3^a(2+3b)$, we are done, because we have shown that K_{3^a} is a proper divisor of those K_n. Hence for K_n to be prime n has to be a power of 3.

Comment. John Robertson reports that K_{3^i} is composite for $i = 2, 3, \ldots, 9$.

Also solved by Michel Bataille (France), John Christopher, Charles K. Cook, Daniele Donini (Italy), Matt Foss, Natalio H. Guersenzvaig (Argentina), Laurel and Hardy Problem Group, John Robertson, Michael Vowe (Switzerland), and the proposer.

Answers

Solutions to the Quickies on page 326

A893. Let $r_n(p)$ be the expected number of ranges when there are n pages and the word occurs on each page with probability p. We show that $r_n(p) = p + (n-1)p(1-p)$ by induction on n.

When $n = 1$, we have one range with probability p and zero ranges otherwise. Thus,

$$r_n(p) = p = p + (n-1)p(1-p).$$

Suppose $n > 1$ and assume the formula holds for r_{n-1}. Before we add the last page, the expected number of ranges is $r_{n-1}(p)$. When page n is added, the number of ranges increases by one if the term occurs on page n and does not occur on page $n-1$; this happens with probability $p(1-p)$. Otherwise, the number of ranges does not change. Thus,

$$\begin{aligned} r_n(p) &= p(1-p)\cdot[r_{n-1}(p) + 1] + [1 - p(1-p)]\cdot r_{n-1}(p) \\ &= r_{n-1}(p) + p(1-p) = p + (n-1)p(1-p). \end{aligned}$$

A894. The required locus is the open circular disk concentric with $\mathscr{C}$ and having three times the radius. To see this, note that any point inside $\mathscr{C}$ is the centroid of a triangle T inscribed in $\mathscr{C}$; in fact we can take T to be isosceles. Therefore the locus of *centroids* of triangles inscribed in $\mathscr{C}$ is the open disk bounded by $\mathscr{C}$. Now use Euler's theorem: in any triangle the three points O (circumcenter), G (centroid) and H (orthocenter) are collinear in this order, and $OH = 3OG$. The result is immediate.

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.

Mackenzie, Dana, Fermat's Last Theorem extended, *Science* 285 (9 July 1999) 178.

The Taniyama-Shimura conjecture, that all elliptic curves are modular, has been proved by Christophe Breuil (Université de Paris-Sud), Brian Conrad and Richard Taylor (Harvard University), and Fred Diamond (Rutgers University). This means that modular functions can be used to help find the rational points on any elliptic curve. Andrew Wiles had proved the conjecture for semistable elliptic curves, a class that includes any elliptic curves that could arise from counterexamples to FLT.

Peterson, Ivars, The honeycomb conjecture: Proving mathematically that honeybee constructors are on the right track. *Science News* 156 (24 July 1999) 60–61. Morgan, Frank, Frank Morgan's Math Chat: Hales proves hexagonal honeycomb conjecture, `http://www.maa.org/features/mathchat/mathchat_6_17_99.html` . Morgan, Frank, The hexagonal honeycomb conjecture, *Transactions of the American Mathematical Society* 351 (1999) 1753–1763. Hales, Thomas, The honeycomb conjecture, `http://xxx.lanl.gov/abs/math.MG/9906042`.

Thomas Hales (University of Michigan), who last year proved the Kepler conjecture on the densest arrangement of spheres in 3-space, has announced a proof that paving with congruent regular hexagons is the most perimeter-efficient way to enclose infinitely many unit areas in the plane. The result had already been proven for polygons by L. Fejes-Tóth; Hales's proof covers figures with curvilinear sides.

Dawson, John W., Jr., Gödel and the limits of logic, *Scientific American* (June 1999) 76–81. Alpert, Mark, Profile: Not just fun and games, *Scientific American* (April 1999) 40–42.

The biographical account of Gödel, by the author of his biography (*Logical Dilemmas: The Life and Work of Kurt Gödel*, A K Peters, 1997), briefly but deftly surveys Gödel's work, "pursued against a background of recurrent mental instability." Unfortunately, the books and articles of the last few years about Gödel, John Nash, and Paul Erdős may leave the public with the generalization that all great mathematicians are crazy (plus doubts about you and me). Fortunately, articles like the delightful profile of John H. Conway provide some counterbalance (he admits to bouts of depression, but that's a common malady).

Gazalé, Midhat J., *Gnomon: From Pharaohs to Fractals*, Princeton University Press, 1999; xiii + 259 pp, $29.95. ISBN 0-691-00514-1.

The theme of this unusual and original book is self-similarity, and the author takes the gnomon as the prototypical self-repeating shape. Beautifully illustrated, the book moves quickly through Fibonacci sequences and continued fractions to whorled figures, spirals, Kronecker products of matrices, and fractals. Readers whose appetites are whetted, however, will wish there were more references to further sources.

Peterson, Ivars, Fibonacci at random: Uncovering a new mathematical constant, *Science News* 155 (12 June 1999) 376–377. Hayes, Brian, Computing Science: The Vibonacci numbers, *American Scientist* (July-August 1999) 296–301.

What happens if you add some randomness to the defining recurrence $F_{n+1} := F_n + F_{n-1}$ for the Fibonacci numbers? In his thesis at Cornell University, Divakar Viswanath (Mathematical Sciences Research Institute) investigated $F_{n+1} := F_n \pm F_{n-1}$, where addition and subtraction are equally likely random events. Viswanath found that $|F_n| \approx C^n$, where $C = 1.13198824\ldots$. Hillel Furstenberg (Hebrew University) and Harry Kesten (Cornell University) had already shown that a class of sequences produced by such processes features this kind of exponential behavior but had not identified the constant for this random Fibonacci sequence. Determining the value of C involved recasting the recurrence in terms of random matrices, interpreting two consecutive terms of the sequence as a point in the plane, and realizing that the process generates a random walk through the space of slopes generated by consecutive pairs of such points. Viswanath then realized that there is a correspondence between the random walks and the paths through the branches of the Stern-Brocot tree. The "Vibonacci" sequence is an example of how "repeated random movements can lead to orderly behavior," an underappreciated phenomenon that also explains the transparency of glass and why random impurities in a semiconductor do not scramble the electric current passing through it.

Banks, Robert B., *Towing Icebergs, Falling Dominoes, and Other Adventures in Applied Mathematics*, Princeton University Press, 1998; xv + 328 pp, $29.95. ISBN 0-691-05948-9.

This is a delightful book about the application of mathematics to interesting problems in economics, sports, and other topics, such as meteor craters, the feasibility of towing icebergs to alleviate water shortages, the "economic energy" of nations, and the cyclical consequences of the player draft for the National Football League. (The dust jacket says that the book "requires of its readers only a basic understanding of high school or college math"; in fact, the author freely uses the appropriate mathematics, i.e., calculus.)

Graphica 1: The World of Mathematica® Graphics. The Imaginary Made Real: The Images of Michael Trott. A K Peters, 1999; xiii + 89 pp. ISBN 1-56881-106-3. *Graphica 2: The World of Mathematica® The Pattern of Beauty: The Art of Igor Bakshee.* A K Peters, 1999; xii + 86 pp. ISBN 1-56881-107-1.

Mathematicians have enjoyed the program Mathematica as a computational tool; now they can show their nonmathematical friends how Mathematica can also be used as a tool for producing art. These two volumes, one by a mathematician and the other by a physicist who designs for textile manufacturers, are absolutely beautiful. Each contains an introduction by the author and a postscript in which the author describes the steps in the Mathematica code for building one of the works.

Chance, Don M., and Pamela P. Peterson, The new science of finance, *American Scientist* (May-June 1999) 256–263. Kestenbaum, David, Death by the numbers, *Science* 283 (26 February 1999) 1244–1247.

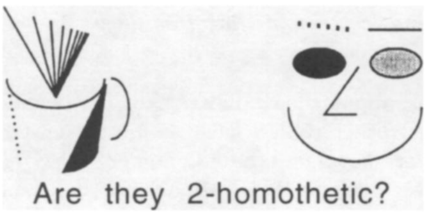
The article by Chance and Peterson is a broad account of valuation in finance without any mathematical details that tries to be comforting with assurances about the efficiency of markets that trade in financial derivatives. Kestenbaum's article, on the other hand, centers on the consequences of the August 1998 Russian financial crisis, which triggered "a global panic" about hedge funds that some blamed on "mathematical hubris." Mathematical analysis, of course, should not be a scapegoat for poor judgment and greed.

NEWS AND LETTERS

Carl B. Allendoerfer Awards – 1999

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of expository articles published in *Mathematics Magazine*. Carl B. Allendoerfer, a distinguished mathematician at the University of Washington, served as President of the Mathematical Association of America, 1959–60. This year's awards were presented at the July 1999 Mathfest, in Providence. The citations follow.

Victor Klee and John R. Reay, "A Surprising but Easily Proved Geometric Decomposition Theorem," Mathematics Magazine 71 (February 1998). The problem addressed by the article is illustrated whimsically by a great drawing:



Two planar sets A and B are *homothetic* if they are similar and similarly oriented. Two sets A and B are *2-homothetic* if each can be partitioned into two disjoint sets ($A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ with $A_1 \cap A_2 = \emptyset = B_1 \cap B_2$) in such a way that A_1 and B_1 are homothetic and A_2 and B_2 are homothetic. The authors prove a surprising result: Two sets in the plane are 2-homothetic provided each of them is bounded and has nonempty interior. The proof follows easily from a strengthened form of the Cantor–Bernstein theorem. Though accessible to undergraduates, the article ends with a set of open problems, and draws upon a wide variety of important mathematical work by such mathematicians as Banach, Bernstein, Cantor, Fraenkel, and Tarski. It's fun, surprising, crystal clear, and, somehow, both straightforward and non-trivial at the same time.

Donald G. Saari and Fabrice Valognes, "Geometry, Voting, and Paradoxes," Mathematics Magazine 71 (October 1998). Combinatorics is the traditional technique to compare and analyze voting procedures. What makes this paper so inviting is the use of *geometry* to compare voting procedures. Accessible, easily visualized mathematics leads to some surprising results. After seizing the reader's interest with simple examples, the authors present some historic voting procedures (the Borda Count and the Condorcet winner). Then, in the authors' words, they "demonstrate how geometry dramatically reduces these previously complicated issues into forms simple enough to be presented to students who can graph elementary algebraic equations." The article reveals a novel view of a classic problem.

Biographical Notes **Victor Klee's** 1949 Ph.D. was from the University of Virginia, which attracted him because of an initial interest in point-set topology. While there, he became interested also in functional analysis and convex geometry. After the move to Seattle in 1953, his interests broadened to include combinatorics, optimization, and computational complexity. These days, he says he likes to work in a variety of fields in order to spread his mistakes more thinly. He is a co-author, with Stan Wagon, of the MAA book *Old and New Unsolved Problems in Plane Geometry and Number Theory.* Professor Klee was MAA President in 1971–73.

John R. Reay studied music at Pacific Lutheran University and mathematics at the University of Washington, where Victor Klee directed his 1963 Ph.D. thesis. He now teaches at Western Washington University and plays in the Whatcom Symphony Orchestra. The joint paper on 2-homothetic sets grew out of a talk he prepared for the Visiting Lecturer Program of the MAA, and the goading of friends who wanted a written version. The talk was based on earlier lectures of Klee.

Don Saari received his Ph.D. from Purdue University and his B.S. from Michigan Technological University. He is currently the Pancoe Professor of Mathematics and a Professor of Economics at Northwestern University. His research interests center around dynamical systems and their applications—primarily as applied to the Newtonian N-body problem and to issues coming from the social sciences. His most recent book is *Basic Geometry of Voting*, Springer-Verlag, 1995. Because of Saari's current interest in voting procedures, he makes frequent research trips to the French Institute of Social Choice and Welfare, University of Caen, where he met his co-author Fabrice Valognes.

Fabrice Valognes was born in 1969 in the city of Caen (Normandy) and received all of his degrees from the University of Caen. In 1998, he received his Ph.D. in field mathematical economics from the same university. Valognes' supervisors were Professors Maurice Salles and Dominique Lepelley (both of the University of Caen). However, Donald G. Saari and William V. Gehrlein (University of Delaware) also influenced him greatly, and quickly became his "part-time" supervisors. Valognes' dissertation is entitled "Essays in Social Choice Theory." He is an Associate Professor of Economics at the University of Namur (Belgium).

Letters to the Editor

Dear Editor:

The Fernandez and Piron article (*Mathematics Magazine*, June 1999 issue) added a new twist to the car-and-goats soap opera. It is unfortunate that the article did not mention the extensive references that were published in Ed Barbeau's columns in the *College Mathematics Journal* (March 1993, pp.149-154; March 1995, pp. 132-134).

Domenico Rosa
DRosa@teikyopost.edu

Dear Editor:

S.P. Glasby's article *Extended Euclid's Algorithm*, (June issue, 1999) describes the backward method of computing the extended gcd coefficients as "pedagogically preferable to the forward recurrence method"; the forward method is described as "good for computers." But we should teach students what is good for computers.

The backward method, though indeed still presented in many texts, is inferior to the forward method, for several reasons. One is the memory requirement, and I disagree with the assertion that "there is no advantage for hand calculation, as the student will invariably have recorded each q_i on paper." This argument might have been valid 30 years ago, but today we assume that the student knows about computers (or, for hand calculations, that paper is a resource to be conserved). Computing modular inverses for large numbers is very important; a method that stores only one quotient at a time is (other things being equal) clearly superior to one that requires retaining all quotients. Moreover, in the forward method, the recurrence for the extended gcd coefficients is essentially identical to the recurrence for the remainder sequence. This allows the computation to be done in matrix form, as follows:

$$\begin{pmatrix} r_{\text{old}} & s_{\text{old}} \\ r_{\text{new}} & s_{\text{new}} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q \end{pmatrix} \begin{pmatrix} r_{\text{old}} & s_{\text{old}} \\ r_{\text{new}} & s_{\text{new}} \end{pmatrix}$$

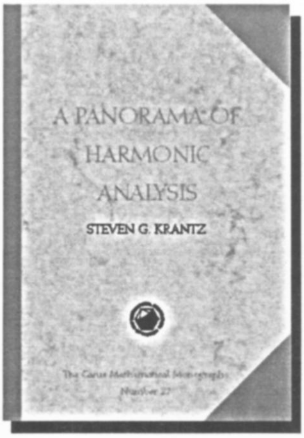
where r_{old} and r_{new} are the two most recent remainders, and s_{old} and s_{new} are the two most recent values of s, the extended gcd coefficient. Such parallelization is generally efficient. My experiments using *Mathematica* on 30-digit numbers confirm that the forward method gets a modular inverse in half the time of the backward method.

Stan Wagon
Macalester College
St. Paul, Minnesota 55105
wagon@macalester.edu

The Mathematical Association of America

A Panorama of Harmonic Analysis

Steven G. Krantz



Series: Carus Mathematical Monographs

This book treats the subject of harmonic analysis, from its earliest beginnings to the latest research. Following both an historical and a conceptual genesis, the book discusses Fourier series of one and several variables, the Fourier transform, spherical harmonics, fractional integrals, and singular integrals on Euclidean space.

The climax of the book is a consideration of the earlier ideas from the point of view of spaces of homogeneous type. The book culminates with a discussion of wavelets - one of the newest ideas in the subject.

The book is intended for graduate students, advanced undergraduates, mathematicians, and anyone wanting to get a quick overview of the subject of cummutative harmonic analysis. Applications are to mathematical physics, engineering and other parts of hard science. Required background is calculus, set theory, integration theory, and the theory of sequences and series.

Contents: Overview of Measure Theory and Functional Analysis; Fourier Series Basics; The Fourier Transform; Multiple Fourier Series; Spherical Harmonics; Fractional Integrals, Singular Integrals, and Hardy Spaces; Modern Theories of Integral Operators; Wavelets; A Retrospective. **Appendices:** The Existence of Testing Functions and Their Density in L^p; Schwartz Functions and the Fourier Transorm; The Interpolation Theorems of Marchinkiewicz and Riesz-Thorin; Haudsorff Measure and Surface Measure; Green's Theorem; The Ganach-Alaoglu Theorem; Expressing an Integral in Terms of the Distribution Function; The Stone-Weierstrass Theorem; Landau's *O* and *o* Notation.

Catalog Code: CAM-27/JR 368 pp., Hardbound, 1999 ISBN 0-88385-031-3 List: $41.50 MAA Member: $33.00

Name__________ Credit Card No.__________

Address__________ Signature__________ Exp. Date___/___

City__________ Qty_____ Price $_____ Amount $_____

State_____ Zip _____ Shipping and Handling $_____

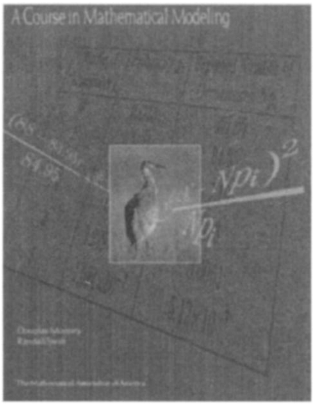
Phone __________ Catalog Code: CAM-27/JR Total $_____

Shipping and Handling: Postage and handling are charged as follows: USA orders (shipped via UPS): $2.95 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 10 days 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: $3.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $7.00 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

THE MATHEMATICAL ASSOCIATION OF AMERICA



A Course in Mathematical Modeling

Series: Classroom Resource Materials

Douglas Mooney and Randall Swift

Textbook

This book is intended as a text for a modeling course accessible to students who have mastered a one year course in calculus. It balances a variety of opposing modeling methodologies including theoretical models versus empirical models, analytical models versus simulation, deterministic models versus stochastic models, and discrete models versus continuous models. Most of the examples are drawn from real-world data or from models that have been used in various applied fields. The use of computers in both simulation and in mathematical analysis is an integral part of the presentation.

The authors emphasize the teaching of the modeling process as opposed to merely presenting models. They begin their book with the simple discrete exponential growth model, and successively refine it to include variable growth rates, multiple variables, growth rates fitted to data, and the effects of random factors. The last part of the book moves into continuous-time models. Issues of model validity and purpose are emphasized throughout.

Students taking a course based on this book should have some mathematical maturity, but will need little advanced knowledge. The book presents more advanced topics on an as-needed basis and serves to show how the different topics of undergraduate mathematics can be used together to solve problems. This perspective is valuable as either a road map for beginning students or as a capstone for more advanced students. The course presents elements of discrete dynamical systems, basic probability theory, differential equations, matrix algebra, stochastic processes, curve fitting, statistical testing, and regression analysis. Computer analysis is extensively used in conjunction with these topics.

You can also use this book if you are seeking applications to supplement a course in linear algebra, differential equations, difference equations, probability theory, or statistics.

Catalog Code: MML/JR
400 pp., Paperbound, 1999
ISBN 0-88385-712-X
List: $41.95 MAA Member: $32.95

Visit <www.wku.edu/~swiftrj/Modeling/ modeling.html> where you can visit the authors' website and download data sets, *Mathematica* files, and other modeling resources that execute the models described in the text.

Phone in Your Order Now!* ☎ *1-800-331-1622

Monday – Friday 8:30 am – 5:00 pm FAX (301) 206-9789

or mail to: The Mathematical Association of America, PO Box 91112, Washington, DC 20090-1112

Shipping and Handling: Postage and handling are charged as follows: **USA orders (shipped via UPS):** $2.95 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 10 days 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:** $3.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $7.00 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.

Name ____________________

Address ____________________

City ____________ State ____ Zip ______

Phone ____________________

QTY.	CATALOG CODE	PRICE	AMOUNT
______	MML/JR	______	______

All orders must be prepaid with the exception of books purchased for resale by bookstores and wholesalers.

Shipping & handling ______

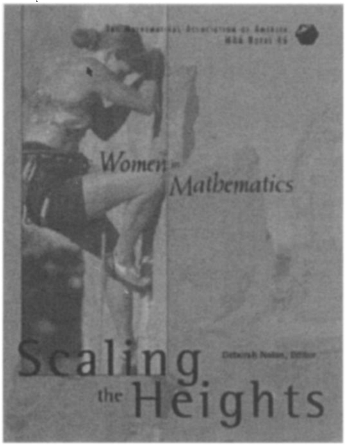
TOTAL ______

Payment ☐ Check ☐ VISA ☐ MasterCard

Credit Card No. ____________ Expires ___/___

Signature ____________________

THE MATHEMATICAL ASSOCIATION OF AMERICA



Women in Mathematics

Scaling the Heights

Deborah Nolan, Editor

Series: MAA Notes

Motivate your students to study advanced mathematics

Women in Mathematics: Scaling the Heights will provide you with a wealth of ideas and examples you can use to develop programs and courses that will motivate undergraduates who want to study advanced mathematics. The articles in this collection make a unique and valuable contribution to upper-division undergraduate mathematics education. They show us what talented undergraduates can do.

The heart of this book presents the insights of eight individuals who have taught at the Summer Mathematics Institute at Mills College. They share their course materials and give pedagogical tips on how to teach topics in mathematics that are not ordinarily part of the undergraduate curriculum, and in ways not often found in the undergraduate classroom. Although the courses described here were designed to encourage talented undergraduate women to pursue advanced degrees in mathematics, the good ideas found in them are gender free and can be used equally well with male as well as female students.

Exercises, class handouts, lists of research projects, and references are included. Topics covered are algebraic coding theory, hyperplane arrangements, p-adic numbers, quadratic reciprocity, stochastic processes, and linear optimization.

The book rounds out the material presented by the Summer Mathematics Institute instructors, with perspectives from mathematicians who have been active in the promotion of women in the field. Results from a survey of undergraduate mathematics majors in which they tell us what they think about the major and their future in mathematics complements these essays.

Catalog Code: NTE-46/JR
146 pp., Paperbound, 1997
ISBN 0-88385-156-3
List: $29.95 MAA Member: $23.95

Phone in Your Order Now! ☎ *1-800-331-1622*

Monday – Friday 8:30 am – 5:00 pm FAX (301) 206-9789

or mail to: The Mathematical Association of America, PO Box 91112, Washington, DC 20090-1112

Shipping and Handling: Postage and handling are charged as follows: **USA orders (shipped via UPS)**: $2.95 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 10 days 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**: $3.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $7.00 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.

Name ____________

Address ____________

City ____________ State _____ Zip ________

Phone ____________

QTY.	CATALOG CODE	PRICE	AMOUNT
_______	NTE-46/JR	_______	_______

All orders must be prepaid with the exception of books purchased for resale by bookstores and wholesalers.

Shipping & handling ________

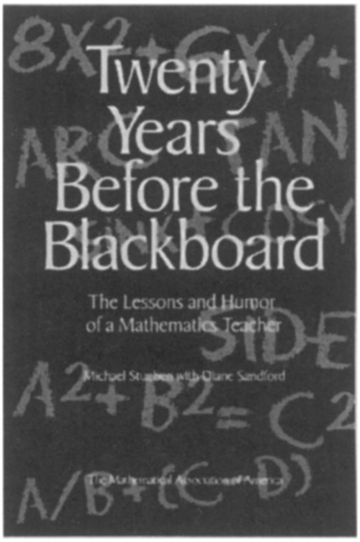
TOTAL ________

Payment ☐ Check ☐ VISA ☐ MasterCard

Credit Card No. ____________ Expires ___/___

Signature ____________

THE MATHEMATICAL ASSOCIATION OF AMERICA



Twenty Years Before the Blackboard

The Lessons and Humor of a Mathematics Teacher

Michael Stueben with Diane Sandford

Series: Spectrum

A perfect gift for the new teacher. . . or for anyone interested in the teaching of mathematics.

This book is the legacy of twenty years of mathematics teaching. During this time, the author searched for motivation techniques, mnemonics, insightful proofs, and serious applications of humor to aid his teaching. The result is this book: part philosophy, part humor, and part biography. Readers will be amused and enlightened on every page.

Mr. Stueben shows how he has used humor and wordplay to motivate his students. The book is filled with wonderful problems and proofs, as well as the author's insights about how to approach teaching problem solving to high school students. Sections of the book also treat the use of calculators and computers in the classroom. A section on mnemonics shows how teachers can use memory aids to help their students learn and retain material.

All in all, *Twenty Years Before the Blackboard* provides a goldmine of ideas for the classroom teacher. Although Mr. Stueben taught at the high school level, his book is an excellent "methods" book for mathematics teachers at all levels.

Read what Martin Gardner has to say about this fascinating book:

It's been decades since I read so entertaining a book about mathematics. The book is a treasure-trove of mathematical jokes, rhymes, anecdotes, word play, mnemonics, and beautiful proofs. For teachers there is an abundance of wise advice based on the author's twenty years in high school teaching. Mathematicians at all levels, from amateurs to college professors will not only chuckle over its gems, but learn much they did not know before. —Martin Gardner

Catalog Code: TYB/JR98
174 pp., Paperbound, 1998, ISBN 0-88385-525-9
List: $29.50 MAA Member: $23.50

Phone in Your Order Now! ☎ 1-800-331-1622

Monday – Friday 8:30 am – 5:00 pm FAX (301) 206-9789

or mail to: The Mathematical Association of America, PO Box 91112, Washington, DC 20090-1112

Shipping and Handling: Postage and handling are charged as follows: **USA orders (shipped via UPS):** $2.95 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 10 days 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:** $3.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $7.00 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.

Name ____________________

Address ____________________

City ____________ State _____ Zip ______

Phone ____________________

Qty.	Catalog Code	Price	Amount
_____	TYB/JR98	_____	_____

All orders must be prepaid with the exception of books purchased for resale by bookstores and wholesalers.

Shipping & handling _____

TOTAL _____

Payment ☐ Check ☐ VISA ☐ MasterCard

Credit Card No. ____________ Expires ___/___

Signature ____________________

CONTENTS

ARTICLES

259 Farmer Ted Goes Natural, *by Greg Martin*

277 Integer Antiprisms and Integer Octahedra, *by Blake E. Peterson and James H. Jordan*

287 Integral Triangles, *by Raymond A. Beauregard and E. R. Suryanarayan*

NOTES

276 Proof Without Words: The Difference Identity for Tangents, *by Fukuzo Suzuki*

295 A Combinatorial Interpretation of a Catalan Numbers Identity, *by David Callan*

299 Cones, k-Cycles of Linear Operators, and Problem B4 on the 1993 Putnam Competition, *by Selwyn Hollis*

304 Flett's Mean Value Theorem for Holomorphic Functions, *by R. M. Davitt, R. C. Powers, T. Riedel, and P. K. Sahoo*

308 Fermat's Little Theorem: A Proof by Function Iteration, *by Lionel Levine*

310 The Hexachordal Theorem: A Mathematical Look at Interval Relations in Twelve-Tone Composition, *by Steven K. Blau*

314 Euler Convergence: Probabilistic Considerations, *by Giorgio Metafune, Diego Pallara, and Carlo Sempi*

317 Proof Without Words: The Product of the Perimeter of a Triangle and its Inradius Is Twice the Area of the Triangle, *by Grace Lin*

317 Revisiting an Old Favorite: $\zeta(2m)$, *by Neville Robbins*

319 Once More: Proportionality in the Non-Euclidean Plane, *by Dieter Ruoff*

321 Stacks, Bracketings, and *CG*-Arrangements, *by Stephen G. Penrice*

324 Period Doubling Near the Feigenbaum Limit, *by J. D. Memory*

PROBLEMS

325 Proposals 1579–1583

326 Quickies 893–894

326 Solutions 1554–1558

331 Answers 893–894

REVIEWS

332

NEWS AND LETTERS

334 Carl B. Allendoerfer Awards 1999

335 Letters to the Editor

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