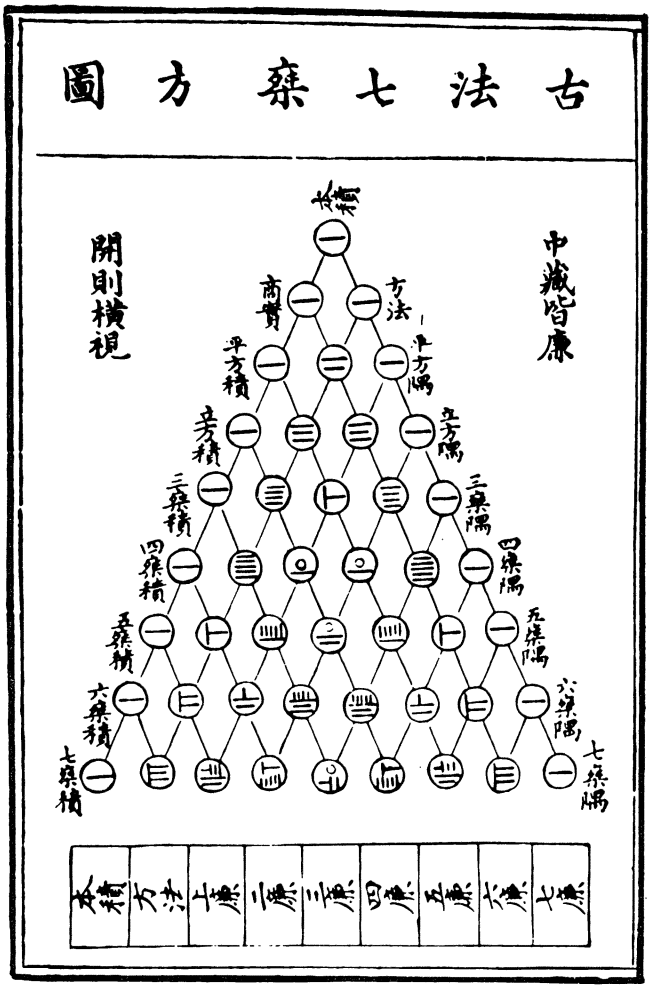
MATHEMATICS MAGAZINE



Vol. 52, No. 1
January, 1979

CHINESE MATHEMATICS • INEQUALITIES

A NEW KIND OF PRIME • $(X+Y)^n = X^n + Y^n$

MAA STUDIES IN MATHEMATICS

This series is intended to bring to the mathematical community expository articles at the collegiate and graduate level on recent developments in mathematics.

These numbers are currently available:

1. *Studies in Modern Analysis,* edited by R. C. Buck.
2. *Studies in Modern Algebra,* edited by A. A. Albert.
3. *Studies in Real and Complex Analysis,* edited by I. I Hirschman, Jr.
4. *Studies in Global Geometry and Analysis,* edited by S. S. Chern.
5. *Studies in Modern Topology,* edited by P. J. Hilton.
6. *Studies in Number Theory,* edited by W. J. LeVeque.
7. *Studies in Applied Mathematics,* edited by A. H. Taub.
8. *Studies in Model Theory,* edited by M. D. Morley.
9. *Studies in Algebraic Logic,* edited by Aubert Daigneault.
10. *Studies in Optimization,* edited by G. B. Dantzig and B. C. Eaves.
11. *Studies in Graph Theory, Part I,* edited by D. R. Fulkerson.
12. *Studies in Graph Theory, Part II,* edited by D. R. Fulkerson.
13. *Studies in Harmonic Analysis,* edited by J. M. Ash.
14. *Studies in Ordinary Differential Equations,* edited by Jack Hale.
15. *Studies in Mathematical Biology, Part I,* edited by S. A. Levin.
16. *Studies in Mathematical Biology, Part II,* edited by S. A. Levin.
17. *Studies in Combinatorics,* edited by Gian-Carlo Rota.

List price per volume: Volumes 1-12, $11.00; volumes 13-14, $15.00; volumes 15-16, $16.00; volume 17, $14.00.

Member's price per volume: Volumes 1-12, $6.50; volumes 13-14, $7.50; volumes 15-16, $12.00; volume 17, $10.00.

Special package prices: Volumes 11 and 12, list price $20.00, member's price $11.50; volumes 15 and 16, list price $27.00, member's price $20.00.

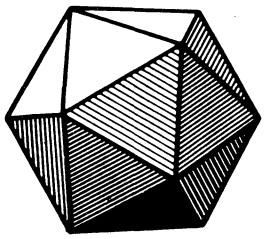
MAA members may purchase one copy of each volume in this series at the special member's price; additional copies and copies for nonmembers may be purchased at the list price. Payment must be received in advance for orders under $10.00. Postage and handling fee will be added to nonprepaid orders.

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, N.W.
Washington, D.C. 20036

MATHEMATICS MAGAZINE

Volume 52, Number 1, January 1979
(ISSN 0025-570X)



COVER: This illustration from a 13th century Chinese text shows an awareness of Pascal's (much later) triangle sufficient to enable use of the binomial theorem, a result partially understood in China as early as the third century B. C. See pp. 10–19.

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

EDITORIAL POLICY

Mathematics Magazine is a journal of collegiate mathematics designed to enrich undergraduate study of the mathematical sciences. The *Magazine* should be an inviting, informal journal emphasizing good mathematical exposition of interest to undergraduate students. Manuscripts accepted for publication in the Magazine should be written in a clear and lively expository style. The *Magazine* is not a research journal, so papers written in the terse "theorem-proof-corollary-remark" style will ordinarily be unsuitable for publication. Articles printed in the *Magazine* should be of a quality and level that makes it realistic for teachers to use them to supplement their regular courses. The editors especially invite manuscripts that provide insight into applications and history of mathematics. We welcome other informal contributions, for example, brief notes, mathematical games, graphics and humor.

Editorial correspondence should be sent to: Mathematics Magazine, Department of Mathematics, St. Olaf College, Northfield, Minnesota 55057. Manuscripts should be prepared in a style consistent with the format of *Mathematics Magazine.* They should be typewritten and double spaced on $8\frac{1}{2}$ by 11 paper. Authors should submit the original and one copy and keep one copy as protection against possible loss. Illustrations should be carefully prepared on separate sheets of paper in black ink, the original without lettering and two copies with lettering added; the printers will insert printed letters on the illustration in the appropriate locations.

Authors planning to submit manuscripts may find it helpful to obtain the more detailed statement of guidelines available from the editorial office.

ABOUT OUR AUTHORS

John O. Kiltinen ("Linearity of Exponentiation") holds a Ph.D. from Duke University where he specialized in topological algebra. After holding a research instructorship at the University of Minnesota, he moved to Northern Michigan University where he has been since 1971. He became interested in the problem treated in this article as a result of teaching an undergraduate abstract algebra class where he and the students became interested in extensions of the fact that $(a+b)^p=a^p+b^p$ for a ring of prime characteristic p.

Frank Swetz ("The Evolution of Mathematics in Ancient China") earned a doctorate in 1972 from Teachers College, Columbia University, with a study of mathematics education in China. As a result of this project, he became interested in the more general topic of societal influences on the development of mathematics and mathematical thought. This interest has also involved him in work concerning problems of mathematics education in developing countries. He is presently Associate Professor of Mathematics and Education at The Capitol Campus of The Pennsylvania State University. His *Was Pythagoras Chinese?* was recently released as a joint publication by Penn State Press and the N.C.T.M.

BUSINESS INFORMATION. Mathematics Magazine is published by the Mathematical Association of America at Washington, D.C., five times a year in January, March, May, September, and November. Ordinary subscriptions are $12 per year. Members of the Mathematical Association of America or of Mu Alpha Theta may subscribe at special reduced rates. Colleges and university mathematics departments may purchase bulk subscriptions (5 or more copies to a single address) for distribution to undergraduate students.

Subscription correspondence and notice of change of address should be sent to A. B. Willcox, Executive Director, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Back issues may be purchased, when in print, from P. and H. Bliss Co., Middletown, Connecticut 06457.

Advertising correspondence should be addressed to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, New York 14214.

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

Linearity of Exponentiation

An investigation of when ignorance is bliss: Yes, Virginia, sometimes $(a+b)^m=a^m+b^m$.

JOHN O. KILTINEN
Northern Michigan University
Marquette, MI 49855

One of the driving forces of mathematics is the search for order and simplicity. Even the youngest students of mathematics seem to implicitly sense this. Unfortunately, they often are tempted to assume more simplicity than actually holds. A case in point is the rule which has been facetiously called the "universal law of linearity." This "law" holds that for any function f and any a and b in its domain, $f(a+b)=f(a)+f(b)$. Some of the guises in which this law most frequently appears in the work of students (especially in stressful situations such as tests) are: $\sqrt{a+b} = \sqrt{a}+\sqrt{b}$, $1/(a+b) = 1/a + 1/b$, and $(a+b)^2 = a^2+b^2$. Indeed, one of the major hurdles of coming to grips with mathematical abstraction at the earliest levels has to do with recognizing when "laws" like this hold, and when they do not.

It comes as something of a shock, therefore, that after years of training which has taught him *not* to write $(a+b)^m=a^m+b^m$, the beginning student of abstract algebra learns that under some circumstances, this *linearity of exponentiation* rule holds. In particular, if m is a prime power p^k, then $(a+b)^m=a^m+b^m$ for any a and b in a ring of characteristic p [**8**, Cor. 38.2.1, p. 363]. A common response to this discovery is a sense of vindication that this simple, elegant rule to which one instinctively gravitated at a more naive level of development does indeed sometimes hold. One can follow upon this response with an effort to explore more deeply into those situations in which exponentiation is linear. That is the purpose of this paper.

Having seen that an assumption about the characteristic of a ring can yield the rule $(a+b)^m=a^m+b^m$ for certain values of m, we are led to consider the converse question: if $(a+b)^m=a^m+b^m$ holds identically in a ring R, then what does this imply about the characteristic of R? This question arose in one of the author's abstract algebra classes, and an investigation of it, in which the students participated, led to the results in this paper.

We will first show that the identity $(a+b)^m=a^m+b^m$ does not necessarily affect the characteristic of a ring, but that when the ring has an identity, the effect is significant. Specifically, we shall identify certain "weak conditions" relating integers m and n which are necessary if $(a+b)^m=a^m+b^m$ is to hold in a ring with identity of characteristic n. We then prove that these weak conditions are also sufficient for this rule to hold in a certain class of rings. Finally, we will present some "strong conditions" which are always sufficient to assure that $(a+b)^m=a^m+b^m$ holds in a ring of characteristic n, and demonstrate that these strong conditions are also necessary in a certain class of rings.

All rings in this paper are commutative. We denote the integers by Z, the natural numbers by N and the integers modulo n by Z_n. If R has an identity, we shall denote it by 1, even though that might risk confusion with the integer 1. Similarly, for any n in Z, we shall also denote by n the ring element obtained by adding n 1's in R if n is positive or by adding $-n$ -1's if n is negative. For emphasis and

clarity, we shall occasionally remind the reader when certain computations are in R or in Z by saying explicitly "in R" or "in Z." Recall that if R has an identity and if $\text{char}(R)=n$ for $n\neq 0$, then n is the least positive integer such that $n=0$ in R [**8**, p. 192]. Recall also that if $m=0$ in R for some other integer m, then n divides m in Z.

A first negative result

Throughout this paper, we say that a ring R is ***m*-linear** (for $m \geqslant 1$) provided that the equality $(a+b)^m = a^m + b^m$ holds for every a and b in R. We first observe that in general, m-linearity in a ring R need have no effect on $\text{char}(R)$. To see this, note that R is trivially m-linear if every element of R is nilpotent of order at most m, i.e., if $a^m=0$ for every a in R. Rings having this property can have arbitrary characteristic. Indeed, let S be any ring with identity of characteristic n, and let $S[X]$ be the ring of polynomials in an indeterminate X with coefficients in S. Let (X) and (X^m) denote, respectively, the ideals in $S[X]$ generated by X and X^m. Finally, let $R=(X)/(X^m)$, the quotient ring of (X) modulo (X^m) [**8**, pp. 165-66].

R can be thought of as consisting of polynomials of degree less than m with zero constant terms. These are added and multiplied in the usual manner, except that after a multiplication, all of the terms of degree m or greater are truncated from the product. Note that a product of two elements of R has a zero coefficient for X, a product of three elements also has a zero coefficient for X^2, etc. From this, it follows that $a^m=0$ for any a in R. Also, $\text{char}(R)=\text{char}(S)=n$. Hence *for any pair of positive integers m and n, there is a commutative m-linear ring R of characteristic n.*

The consequences of an identity element

Suppose that R is an m-linear ring with identity 1, and that $\text{char}(R)=n$. We will see that having an identity in R places severe restrictions on n.

LEMMA 1. *If R is an m-linear commutative ring with identity, and if $\text{char}(R)=n$, then n divides k^m-k for all k in N.*

Proof. Take k in N and consider k^m in R. By m-linearity, $k^m=(1+\cdots+1)^m=1^m+\cdots+1^m=1+\cdots+1=k$ in R. Thus, $k^m-k=0$ in R, which implies, since $n=\text{char}(R)$, that n divides k^m-k in Z.

We now examine the number-theoretic consequences of the relationship between m and n that is identified in Lemma 1. The next lemma, which has recently been published in [**3**] in a slightly different form, contains useful equivalent conditions. The reader is referred to [**3**] for some other ramifications of this result. Our first fundamental result about m-linearity follows immediately from Lemmas 1 and 2.

LEMMA 2. *Let m and n be positive integers with $m>1$. The following two relationships between m and n are equivalent:*

(1) $n \mid k^m-k$ *for all k in N,*

(2) (a) *n is square free, and*

(b) *if p is prime and $p \mid n$, then $(p-1) \mid (m-1)$.*

Proof. First suppose that (1) is true. If n had a square factor p^2, where $p>1$, then by (1) applied to $k=p$, we would have $p^2 \mid p^m-p$, whence, $p \mid p^{m-1}-1$, which is impossible, since $p>1$ and $m-1 \geqslant 1$. Thus, n is square free.

Suppose next that p is a prime and $p \mid n$. Let k be an integer which serves as a generator for the cyclic multiplicative group of the field of integers modulo p. That is, $k^{p-1} \equiv 1 \pmod{p}$, but $k^i \not\equiv 1 \pmod{p}$ for $1 \leqslant i \leqslant p-2$ [**8**, Thm. 38.7, p. 366]. Now by (1) and the fact that $p \mid n$, we have that $p \mid k^m-k$, so $k(k^{m-1}-1) \equiv 0 \pmod{p}$. Now clearly $k \not\equiv 0 \pmod{p}$, so, since Z_p is a field, $k^{m-1}-1 \equiv 0 \pmod{p}$, and hence $k^{m-1} \equiv 1 \pmod{p}$. But since $p-1$ is the multiplicative order of k modulo p, this implies that $(p-1) \mid (m-1)$, and (b) is proven.

Now let us suppose that (a) and (b) of (2) hold. Then since n is square free, it factors in the form $n=p_1\cdot\cdots\cdot p_r$, where the p_i's are distinct primes. By (b), $(p_i-1)|(m-1)$ for $1\leqslant i\leqslant r$. Now by Fermat's theorem [**8**, Cor. 38.3.1, p. 364], if $k\in N$ and $k\not\equiv 0 \pmod{p_i}$, then $k^{p_i-1}\equiv 1 \pmod{p_i}$. Since $(p_i-1)|(m-1)$, it follows that $k^{m-1}\equiv 1 \pmod{p_i}$ if $k\not\equiv 0 \pmod{p_i}$. From this, we have that $k^m\equiv k \pmod{p_i}$. This last congruence also clearly holds if $k\equiv 0 \pmod{p_i}$, so we have $k^m-k \equiv 0 \pmod{p_i}$ for all $k\in N$. It follows, since the p_i's are relatively prime, that $k^m-k\equiv 0 \pmod{n}$, and so $n|k^m-k$ for all k in N.

THEOREM 1. *If R is an m-linear commutative ring with identity, with $m>1$ and* $\text{char}(R)=n$, *then n is square free, and if p is a prime factor of n, then $(p-1)|(m-1)$.*

The following interesting corollary demonstrates how severely the characteristic of R can be restricted by m-linearity in R: *if m is even and if R is an m-linear commutative ring with identity, then* $\text{char}(R)=2$. This is an easy consequence of Theorem 1. Since m is even, $m-1$ is odd, so $p-1$ is odd for every prime divisor p of $n=\text{char}(R)$. This means that 2 is the only prime divisor of n. Since n is square free, n must be 2.

m-linearity in Z_n

We have seen that m and n must be related in a special way if m-linearity holds in a commutative ring R with identity of characteristic n. However, if m and n are integers satisfying these conditions, then is it necessarily the case that there exists a ring of characteristic n which is m-linear? We will answer this question by showing that if m and n satisfy these conditions, then Z_n is m-linear. Indeed, Z_n is m-linear for a very simple reason.

To get at this reason, we look at m-linearity in another way. In what follows we will denote by ϕ_m the mth power function on any commutative ring R. In other words, $\phi_m: x\to x^m$ for all x in R. Since R is commutative, $\phi_m(a\cdot b)=\phi_m(a)\cdot\phi_m(b)$; moreover, m-linearity holds in R if and only if $\phi_m(a+b)=\phi_m(a)+\phi_m(b)$. Thus R is m-linear if and only if ϕ_m is a ring homomorphism. Now the identity function $\text{id}_R: x\to x$ on R is a ring homomorphism. Thus, we can make the trivial observation that R is m-linear if $\phi_m=\text{id}_R$. In other words, if $x^m=x$ for all x in R, then $(a+b)^m=a^m+b^m$.

This is in fact what happens in Z_n if m and n satisfy the condition $n|k^m-k$ (for all k in N) of Lemma 1. Indeed, the condition can be restated as $k^m-k\equiv 0 \pmod{n}$, or $k^m\equiv k \pmod{n}$, which is equivalent to saying that $x^m=x$ for all x in Z_n. Thus, ϕ_m is the identity on Z_n. This of course implies that Z_n is m-linear, so we have a partial converse for Theorem 1. We summarize now a number of equivalent facts within this circle of ideas.

THEOREM 2. *Let m and n be positive integers with $m>1$, and let ϕ_m be the power function $x\to x^m$ on Z_n. Then the following are equivalent:*

(1) *n is square-free and if p is a prime factor of n, then $(p-1)|(m-1)$.*
(2) *$n|k^m-k$ for all k in N.*
(3) *Z_n is m-linear.*
(4) *ϕ_m is a ring homomorphism on Z_n.*
(5) *ϕ_m is a ring isomorphism on Z_n.*
(6) *ϕ_m is the identity function on Z_n.*

We now pause for a moment to give examples to illustrate Theorem 2. Suppose that we want to know for exactly what positive integers n the identity $(a+b)^7=a^7+b^7$ holds in Z_n. By Theorem 2, n must be a product of distinct primes p such that $p-1$ divides 6. So p can be only 2, 3, or 7. This means that the $2^3-1=7$ possible products of distinct primes from the set {2, 3, 7}, namely 2, 3, 6, 7, 14, 21, and 42 are the only possible values of n for which Z_n is 7-linear.

Conversely, we ask for what values of m is Z_{115} m-linear? Since $115=5\cdot 23$ is square free, such values of m are possible. Since 5 and 23 are the prime divisors of 115, we must have both 4 and 22 as

divisors of $m-1$. Thus, $44|(m-1)$, whence $m=44k+1$. By Theorem 2, for Z_{115} to be m-linear, it is sufficient for m to be of the form $44k+1$, where k is any positive integer. Thus, Z_{115} is 45-linear, 89-linear, 133-linear, etc.

A generalization

Theorem 2 generalizes to rings of characteristic n which are direct products or sums of rings of the form Z_k. For background on direct products and sums, the reader is referred to [**8**, Section 24, p. 220 ff.] for the finite case, and to [**2**, pp. 68-72] for the general case. The following lemma shows how the factors affect the ring.

LEMMA 3. (a) *The rings R_α are m-linear for all α in an index set Λ if and only if their direct product or sum is m-linear.*

(b) *If the characteristic of the direct product (sum) of a set $\{R_\alpha|\alpha\in\Lambda\}$ of rings is nonzero, then* $\text{char}(R_\alpha)$ *is nonzero for each α in Λ, and the characteristic of the product (or sum) is the least common multiple of the characteristics of the factors.*

Proof. For the sake of notational simplicity, we give the proof for the case of the product $R_1\times R_2$ of two rings. The ideas of the proof generalize to arbitrary (possibly infinite) products and sums.

The proof of (a) reflects the component-wise nature of the operations on a product ring. If, (x_1,y_1) and (x_2,y_2) are in $R_1\times R_2$, then $((x_1,y_1)+(x_2,y_2))^m=(x_1+x_2,y_1+y_2)^m=((x_1+x_2)^m,(y_1+y_2)^m)=(x_1^m+x_2^m,y_1^m+y_2^m)=(x_1^m,y_1^m)+(x_2^m,y_2^m)=(x_1,y_1)^m+(x_2,y_2)^m$. Thus, $R_1\times R_2$ is m-linear given that both R_1 and R_2 are m-linear. Conversely, if $R_1\times R_2$ is m-linear, then for $x_1,x_2\in R_1$, $((x_1+x_2)^m,\ 0)=((x_1+x_2),0)^m=((x_1,0)+(x_2,0))^m=(x_1,0)^m+(x_2,0)^m=(x_1^m,0)+(x_2^m,0)=(x_1^m+x_2^m,0)$. So $(x_1+x_2)^m=x_1^m+x_2^m$, thus R_1 is m-linear. Similarly, R_2 is m-linear.

To prove (b), let $n=\text{char}(R_1\times R_2)$, $n_1=\text{char}(R_1)$, and $n_2=\text{char}(R_2)$. Since we are not assuming that there are identity elements, we must define the characteristic of a ring R to be the least positive integer m such that $m\cdot x=0$ for all x in R, if such an m exists, and to be zero otherwise. A proof using the division algorithm in N assures that any k in N such that $kx=0$ for all x in R is a multiple of $\text{char}(R)$.

Note first that $(0,0)=n(x,y)=(nx,ny)$ for all (x,y) in $R_1\times R_2$, so $nx=0$ for all x in R_1 and $ny=0$ for all y in R_2. Thus $n_1|n$ and $n_2|n$, so we have that $n_1\neq 0$ and $n_2\neq 0$, since $n\neq 0$. Now suppose that $n_1|m$ and $n_2|m$ for some positive integer m. Then $mx=0$ for all x in R_1 and $my=0$ for all y in R_2, so $m(x,y)=(mx,my)=(0,0)$ for all (x,y) in $R_1\times R_2$; thus $n|m$, so $n=\text{l.c.m.}(n_1,n_2)$.

THEOREM 3. *Let m and n be positive integers with $m>1$. Let R be a direct product (or sum) of rings Z_k and let* $\text{char}(R)=n$. *Then the following are equivalent.*

(1) *n is square-free and if p is a prime factor of n, then $(p-1)|(m-1)$.*

(2) *$n|k^m-k$ for all k in N.*

(3) *R is m-linear.*

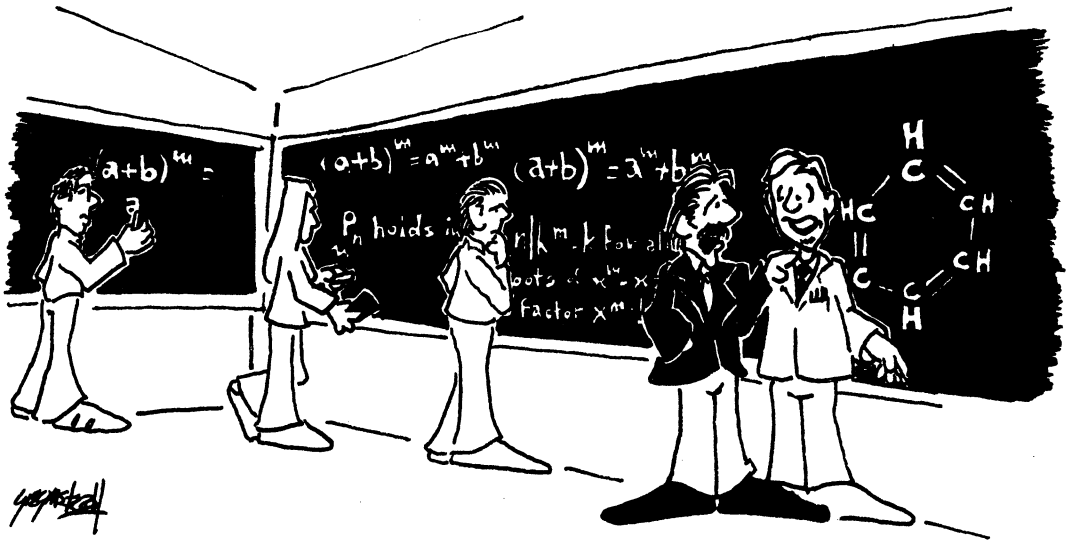
(4) *ϕ_m is a ring homomorphism on R.*

(5) *ϕ_m is a ring isomorphism on R.*

(6) *ϕ_m is the identity function on R.*

Proof. The implications (6)→(5), (5)→(4), and (4)→(3) are as easily observed here as in Theorem 2. The implication (3)→(2) and the equivalence of (2) and (1) follows from Lemmas 1 and 2 for direct products, since in this case R has an identity. However, since an infinite direct sum of rings with identity cannot have an identity, a separate proof is needed for this case.

Suppose that R is the direct sum of a set of rings $\{Z_{k_\alpha}|\alpha\in\Lambda\}$ and that R is m-linear. Then by Lemma 3(a), Z_{k_α} is m-linear for each α in Λ. Then by Theorem 2, k_α is square free and if p is prime and $p|k_\alpha$, then $(p-1)|(m-1)$ for every $\alpha\in\Lambda$. Now since by Lemma 3(b), $n=\text{l.c.m.}\{k_\alpha|\alpha\in\Lambda\}$, n will be square free. Also, if p is prime and $p|n$, then $p|k_\alpha$ for some α, and so $(p-1)|(m-1)$. Thus, m and n satisfy (1) (and hence (2) also by Lemma 2).



"That's very good Hopkins, but I'm afraid it's not quite the kind of ring we're talking about..."

Finally, suppose that (2) holds. We will show that then (6) must follow where R is the direct product or sum of rings $\{Z_{k_\alpha} | \alpha \in \Lambda\}$. Since $n | k^m - k$ for all k in N and since n is a multiple of k_α, we have that $k_\alpha | k^m - k$ for all k in N. Thus, by Theorem 2, ϕ_m is the identity function on Z_{k_α} for each α in Λ. Now because of the component-wise nature of the multiplication on R, ϕ_m as a function from R to R will also be the identity function. Thus (6) is proven.

Sufficient conditions

We have already seen some conditions relating m and n which are necessary if R is to be m-linear, where $\text{char}(R) = n$. We now present some sufficient conditions.

LEMMA 4. *Let R be a commutative ring (not necessarily with identity) such that* $\text{char}(R) = n$. *If n is a prime and m is a power of n or if m is an integer such that n divides $\binom{m}{k}$ for $1 \leqslant k \leqslant m-1$, then R is m-linear.*

Proof. The first condition, as observed earlier, is well known; the sufficiency of the second is an easy consequence of the binomial theorem. Indeed, for any a and b in R, $(a+b)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} b^k$. Since $n | \binom{m}{k}$ for $1 \leqslant k \leqslant m-1$, $\binom{m}{k} a^{m-k} b^k = 0$ for $1 \leqslant k \leqslant m-1$, and so the expansion for $(a+b)^m$ reduces to $\binom{m}{0} a^{m-0} b^0 + \binom{m}{m} a^{m-m} b^m = a^m + b^m$.

Elements of high algebraic degree

Scrutiny of the ideas leading up to Theorem 2 indicates that this theorem is made possible by the fact that the rings Z_n are generated by their identity elements. This suggests that if we want to try to find a ring R of characteristic n which is not m-linear, yet for which m and n satisfy condition (1) of Theorem 2, then we must look for an R which has elements which are "sufficiently remote" in some sense from the copy of Z_n in R generated by the identity element of R. An appropriate sense of remoteness turns out to be high algebraic degree.

For any a in R, the *degree* of a over Z_n is the smallest nonnegative integer k such that a is a root of a polynomial of degree k with coefficients in Z_n. If a is the root of no such polynomial, then we will say that a has *infinite degree* over Z_n, or that a is *transcendental* over Z_n. Our next theorem shows that the second sufficient condition of Lemma 4 is also necessary if R has elements of high enough degree over Z_n.

THEOREM 4. *Suppose that R is a commutative ring with identity such that* $\text{char}(R)=n$. *Let m be an integer greater than* 1, *and suppose that there is some a in R such that a has degree at least m over Z_n. Then R is m-linear if and only if n divides $\binom{m}{k}$ for $1 \leqslant k \leqslant m-1$.*

Proof. Suppose that R is m-linear. Then $(a+1)^m = a^m + 1^m = a^m + 1$. On the other hand, by the binomial theorem in R,

$$(a+1)^m = a^m + \binom{m}{1}a^{m-1} + \cdots + \binom{m}{m-1}a + 1.$$

Hence

$$\binom{m}{1}\cdot a^{m-1} + \cdots + \binom{m}{m-1}\cdot a = 0.$$

Since a has degree at least m over Z_n, the only way this can happen is if $\binom{m}{k}=0$ in R for $1 \leqslant k \leqslant m-1$, which is equivalent to $n \mid \binom{m}{k}$ in Z for $1 \leqslant k \leqslant m-1$.

This last theorem makes it worthwhile to determine the greatest common divisor of the binomial coefficients $\binom{m}{k}$ for $1 \leqslant k \leqslant m-1$. We do so in the following three lemmas. L. Dickson in his *History of the theory of numbers* attributes the final result (Lemma 7) to B. Ram, *Jour. of the Indian Math. Club*, Madras, 1 (1909) 39-43. Certainly others have noted it as well. But since no readily available proof was found, one is included here.

LEMMA 5. *If p is a prime and x and k are positive integers, then* $\binom{p^k x}{p^k}$ *is divisible by p exactly as many times as x is.*

Proof. Recall that $\binom{p^k x}{p^k}$ can be expressed as

$$\binom{p^k x}{p^k} = \frac{(p^k x - p^k + 1)\cdot(p^k x - p^k + 2)\cdots(p^k x - 1)\cdot p^k x}{1\cdot 2\cdots(p^k - 1)\cdot p^k}.$$

Note that for j such that $1 \leqslant j \leqslant p^k - 1$, $p^i \mid j$ if and only if $p^i \mid (p^k x - p^k + j)$. This means that all factors of p will cancel out of the expression on the right of the equality above, except for those which come from x. This yields the desired result.

LEMMA 6. *If p is a prime and i is a positive integer, then $p \mid \binom{p^i}{k}$ for $1 \leqslant k \leqslant p^i - 1$.*

Proof. Consider the ring $Z_p[X]$ of polynomials in X over Z_p. Since the characteristic of $Z_p[X]$ is p, we have that $(a+b)^{p^i} = a^{p^i} + b^{p^i}$ for all a and b in $Z_p[X]$. Now X has degree higher than p^i over Z_p, so by Theorem 4, we can conclude that $p \mid \binom{p^i}{k}$ for $1 \leqslant k \leqslant p^i - 1$.

LEMMA 7. *For any positive integer m greater than* 1,

$$gcd\left\{\binom{m}{k} \mid 1 \leqslant k \leqslant m-1\right\} = \begin{cases} 1, & \text{if } m \text{ has at least two distinct prime factors,} \\ p, & \text{if } m \text{ is a power of the prime number } p. \end{cases}$$

Proof. Let $m = p_1^{k_1}\cdot p_2^{k_2}\cdot \cdots \cdot p_r^{k_r}$ be the prime factorization of m. Let d denote the desired greatest common divisor. Note that since $d \mid \binom{m}{1} = m$, the only possible prime divisors of d are $p_1, p_2, \ldots, p_r$. Now if m has at least two distinct prime divisors, then $2 \leqslant p_i^{k_i} \leqslant m-1$ for all i such that $1 \leqslant i \leqslant r$. From Lemma 5, it follows that p_i does not divide $\binom{m_{k_i}}{p_i}$ for $1 \leqslant i \leqslant r$. Thus, $p_i \nmid d$ for $1 \leqslant i \leqslant r$, and d has no

prime divisors. This means that $d=1$. Finally, if $m=p^i$, where $i \geqslant 1$, by Lemma 6 we have $p|\binom{p^i}{k}$ for $1 \leqslant k \leqslant p^i-1$, and thus $p|d$. Also, we have that $d|\binom{p^i}{1}=p^i$, so d is a power of p. However, by Lemma 5, $\binom{p^i}{p^{i-1}}=\binom{p^{i-1}p}{p^{i-1}}$ is divisible by p exactly once. This assures that $d=p$.

As a consequence of Lemma 7, we can now restate Theorem 4 in a more insightful form:

THEOREM 5. *Suppose that R is a commutative ring with identity, that* $\operatorname{char}(R)=n$ *where* $n>1$, *and that R contains an element of degree at least m over Z_n, where $m>1$. Then R is m-linear if and only if n is a prime and m is a power of n.*

Proof. The sufficiency follows from Lemma 4. Conversely, suppose that R is m-linear. By Theorem 4, $n|\binom{m}{k}$ for $1 \leqslant k \leqslant m-1$, so n must divide d, the greatest common divisor of these integers. By Lemma 7, if m were not a prime power, then $d=1$, and $n|1$, which is a contradiction, since $n>1$. Thus m must be a prime power p^i and $d=p$. Since $n|d$ and $n>1$, it follows that $n=p$.

Open questions

We have identified two sets of relationships between positive integers m and n which are pertinent to whether m-linearity holds in a ring of characteristic n. Let us call them the *strong conditions*: n is a prime, m is a power of n; and the *weak conditions*: n is square free, and if p is a prime factor of n, then $(p-1)|(m-1)$. What we have shown is that for m-linearity to hold in a commutative ring R with identity of nonzero characteristic n, the weak conditions are always necessary, and are sufficient when R is Z_n, the ring of integers modulo n, or when R is a direct product or sum (possibly infinite) of rings of the form Z_k. The strong conditions are always sufficient, and are necessary when R contains an element of degree at least m over Z_n.

This yields two classes of commutative rings with identity of non-zero characteristic. The first, call it class A, consists of such rings R with the property that if $\operatorname{char}(R)=n$, then the weak conditions relating m and n are sufficient to assure that R is m-linear. The second, class B, consists of rings R such that if $\operatorname{char}(R)=n$, then if R is m-linear, it is necessary that the strong conditions relating m and n must hold. We have shown some types of rings to be in A and some types to be in B. Can one find other types of rings which can be shown to be in one or the other of these classes?

Could it be the case that every commutative ring with identity and nonzero characteristic must be in either class A or class B? If not, what can be said about the rings which are in neither? Are there conditions relating m and n that are somewhere between the weak and strong conditions which would be relevant?

One might also wish to look at the literature on the related questions of the consequences of $\phi_m: x \to x^m$ being a ring homomorphism in a noncommutative ring [4], [5], and of ϕ_m being a homomorphism on a nonabelian, multiplicatively presented group [1], [7]. In [6] the author delves deeper into the structure of commutative, m-linear rings.

References

[1] J. Alperin, A classification of n-abelian groups, Can. J. Math., 21(1969) 1238-1244.
[2] J. Goldhaber and G. Ehrlich, Algebra, London, Macmillan, 1970.
[3] E. Hewitt, Certain congruences that hold identically, Amer. Math. Monthly, 83(April, 1976) 270-271.
[4] I. Herstein, Power maps in rings, Michigan Math. J., 8(1961) 29-32.
[5] ———, A remark on rings and algebras, Michigan Math. J., 10(1963) 269-272.
[6] J. Kiltinen, Commutative rings with homomorphic mth power functions, to appear.
[7] H. Trotter, Groups in which raising to a power is an automorphism, Canad. Math. Bull., 8(1965) 825-827.
[8] S. Warner, Classical Modern Algebra, Englewood Cliffs, Prentice-Hall, 1971.

The Evolution of Mathematics in Ancient China

Early Chinese mathematical accomplishments reveal arithmetic and algebraic approaches based on sophisticated inductive knowledge.

FRANK SWETZ
The Pennsylvania State University
Middletown, PA 17057

A popular survey book on the development of mathematics has its text prefaced by the following remarks:

> Only a few ancient civilizations, Egypt, Babylonia, India and China, possessed what may be called the rudiments of mathematics. The history of mathematics and indeed the history of western civilization begins with what occurred in the first of these civilizations. The role of India will emerge later, whereas that of China may be ignored because it was not extensive and moreover has no influence on the subsequent development of mathematics.[1]

Even most contemporary works on the history of mathematics reinforce this impression, either by neglecting or depreciating Chinese contributions to the development of mathematics.[2] Whether by ignorance or design, such omissions limit the perspective one might obtain concerning both the evolution of mathematical ideas and the place of mathematics in early societies. In remedying this situation, western historians of mathematics may well take heed of Whittier's admonition:

> We lack but open eye and ear
> To find the Orient's marvels here.[3]

Language barriers may limit this quest for information; however, a search of English language sources will reveal that there are many "marvels" in Chinese mathematics to be considered.

Legend and Fact

The origins of mathematical activity in early China are clouded by mysticism and legend. Mythological Emperor Yü is credited with receiving a divine gift from a Lo river tortoise. The gift in the form of a diagram called the *Lo shu* is believed to contain the principles of Chinese mathematics, and pictures of Yü's reception of the *Lo shu* have adorned Chinese mathematics books for centuries. This fantasy in itself provides some valuable impressions about early Chinese science and mathematics. Yü was the patron of hydraulic engineers; his mission was to control the flood-prone waters of China and provide a safe setting in which a water-dependent civilization could flourish. The users of science and mathematics in China were initially involved with hydraulic engineering projects, the construction of dikes, canals, etc., and with the mundane tasks of logistically supporting such projects. A close inspection of the contents of the *Lo shu* reveals a number configuration (FIGURE 1) which would be known later in the West as a magic square. For Chinese soothsayers and geomancers from the Warring State period of Chinese history (403–221 B.C.) onward, this square, comprised of numbers, possessed real magical qualities because in it they saw a plan of universal harmony based on a cosmology predicated on the dualistic theory of the *Yin* and the *Yang*.[4]

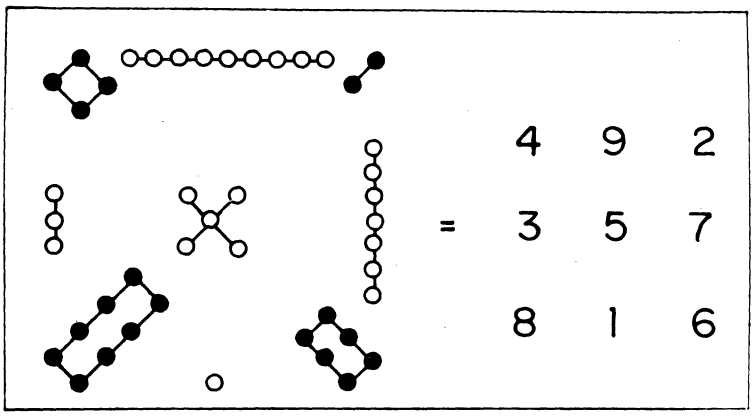
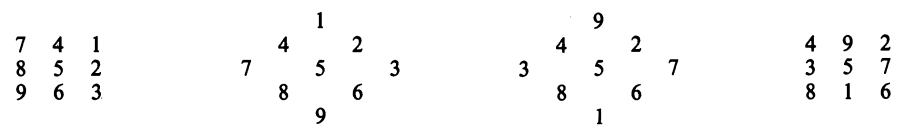


FIGURE 1.

When stripped of ritualistic significance, the principles used in constructing this first known magic square are quite simple and can best be described by use of diagrams as shown in FIGURE 2. The construction and manipulation of magic squares became an art in China even before the concept was known in the West.[5] Variations of the *Lo shu* technique were used in constructing magic squares of higher order with perhaps the most impressive square being that of order nine; see FIGURE 3.



Construct a natural square. Distort it into a diamond. Exchange corner elements. Compress back into a square.

FIGURE 2.

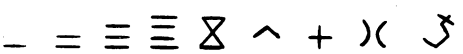
1	10	19	28	37	46	55	64	73		55	28	1		31	76	13	36	81	18	29	74	11
2	11							·		64	37	10		22	40	58	27	45	63	20	38	56
3	12							·		73	46	19		67	④	49	72	⑨	54	65	②	47
4	13							·			(b)			30	75	12	32	77	14	34	79	16
5	14							·						21	39	57	23	41	59	25	43	61
6	15							·						66	③	48	68	⑤	50	70	⑦	52
7	16							·		28	73	10		35	80	17	28	73	10	33	78	15
8	17							·		19	37	55		26	44	62	19	37	55	24	42	60
9	18							81		64	1	46		71	⑧	53	64	①	46	69	⑥	51
				(a)							(c)						(d)					

Start with a natural square (a) then fold each row into a square (b) of order 3 (example using row 1) and apply the *Lo shu* technique (c). The nine resulting magic squares of order 3 (d) are then positionally ordered according to the correspondence of the central element in their bottom rows with the numbers of the *Lo shu*, i.e., 4,9,2; 3,5,7; 8,1,6.

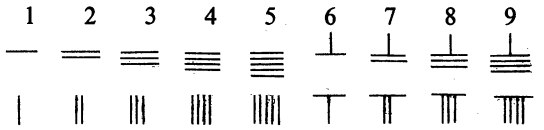
FIGURE 3.

While the *Lo shu* provides some intriguing insights into early mathematical thinking, its significance in terms of potential scientific or technological achievement is negligible. Historically, the first true evidence of mathematical activity can be found in numeration symbols on oracle bones dated from the Shang dynasty (14th century B.C.). Their numerical inscriptions contain both tally and code

symbols, are clearly decimal in their conception, and employ a positional value system. The Shang numerals for the numbers one through nine were:



By the time of the Han Dynasty (2nd century B.C.–4th century A.D.), the system had evolved into a codified notation that lent itself to computational algorithms carried out with a counting board and set of rods. The numerals and their computing-rod configurations are



Thus in this system 4716 would be represented as  .[6] (Occasionally the symbol × was used as an alternative to  .)

Counting boards were divided into columns designating positional groupings by 10. The resulting facility with which the ancient computers could carry out algorithms attests to their full understanding of decimal numeration and computation. As an example, consider the counting board method of multiplying 2 three-digit numbers, as illustrated in FIGURE 4. The continual indexing of partial products to the right as one multiplies by smaller powers of ten testifies to a thorough understanding of decimal notation. In light of such evidence, it would seem that the Chinese were the first society to understand and efficiently utilize a decimal numeration system.[7] If one views a popular schematic of the evolution of our modern system of numeration (FIGURE 5) and places the Chinese system in the

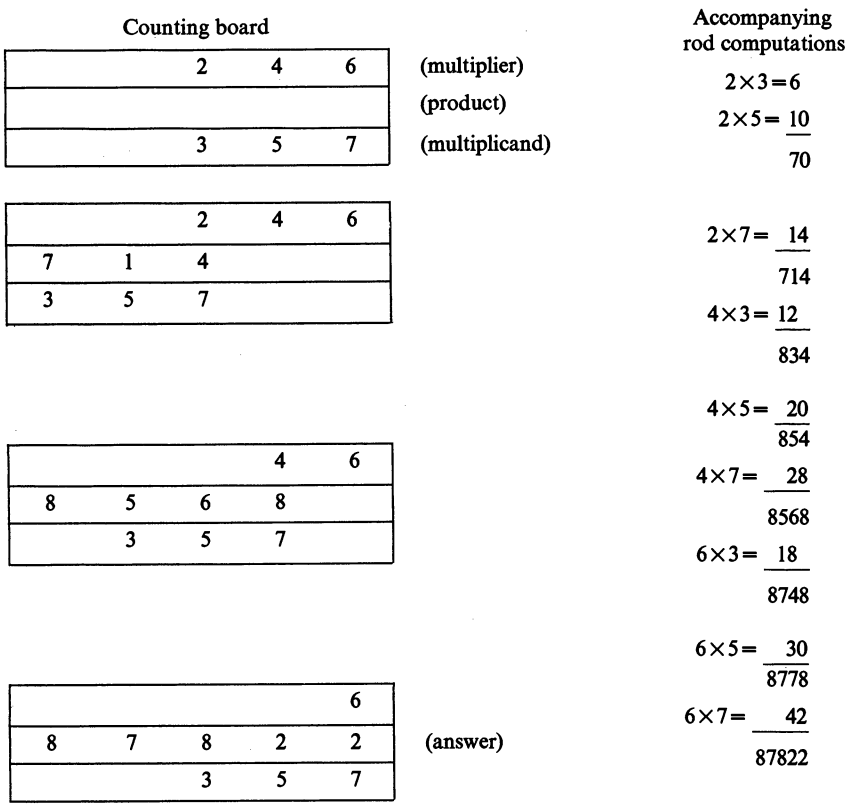


FIGURE 4.

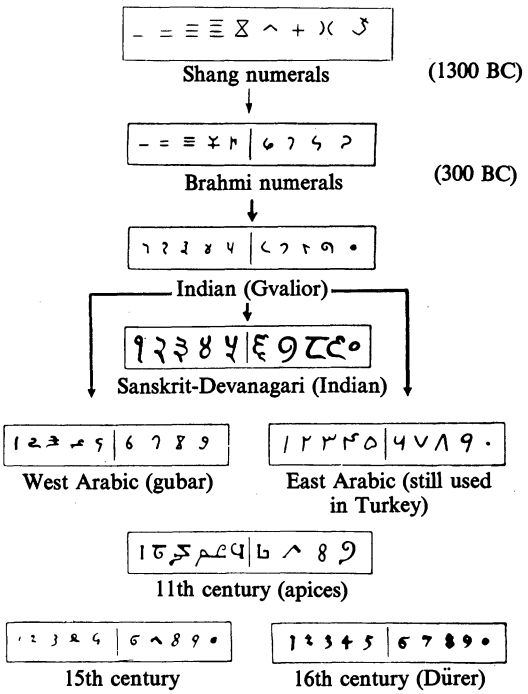


FIGURE 5.

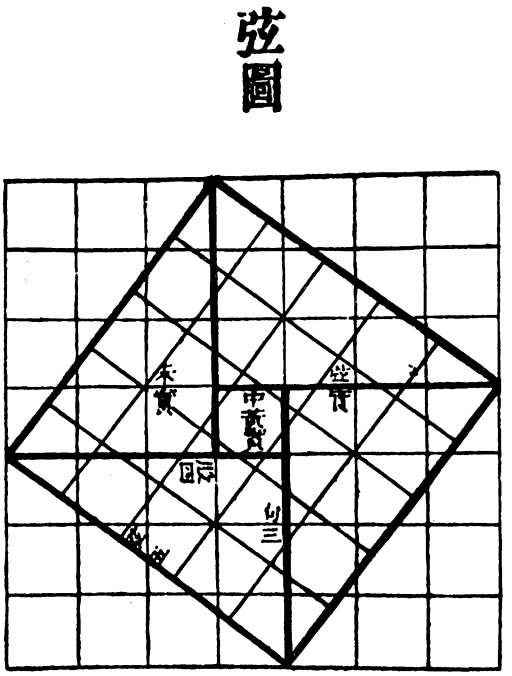


FIGURE 6.

appropriate chronological position, an interesting hypothesis arises, namely that the numeration system commonly used in the modern world had its origins 34 centuries ago in Shang China!

The Systematization of Early Chinese Mathematics

The oldest extant Chinese text containing formal mathematical theories is the *Arithmetic Classic of the Gnomon and the Circular Paths of Heaven*, [*Chou pei suan ching*]. Its contents date before the third century B.C. and reveal that mathematicians of the time could perform basic operations with fractions according to modern principles employing the concept of common denominator. They were knowledgeable in the principles of an empirical geometry and made use of the "Pythagorean theorem." A diagram (see FIGURE 6) in the *Chou pei* presents the oldest known demonstration of the validity of this theorem. This diagram, called the *hsuan-thu* in Chinese, illustrates the arithmetic-geometric methodology that predominates in early Chinese mathematical thinking and shows how arithmetic and geometry could be merged to develop algebraic processes and procedures. If the oblique square of the *hsuan-thu* is dissected and the pieces rearranged so that two of the four congruent right triangles are joined with the remaining two to form two rectangles, then the resulting figure comprised of two rectangles and one small square have the same area as their parent square. Further, since the new configuration can also be viewed as being comprised of two squares whose sides are the legs of the right triangles, this figure demonstrates that the sum of the squares of the legs of a right triangle is equal to the square of the hypotenuse.[8] The processs involved in this intuitive, geometric approach to obtain algebraic results was called *chi-chü* or "the piling up of squares."[9]

The next historical text known to us is also a Han work of about the third century B.C. It is the *Nine Chapters on the Mathematical Art*, [*Chiu chang suan shu*], and its influence on oriental mathematics may be likened to that of Euclid's *Elements* on western mathematical thought. The *Chiu chang's* chapters bear such titles as surveying of land, consultations on engineering works, and impartial taxation, and confirm the impression that the Chinese mathematics of this period centered

on the engineering and bureaucratic needs of the state. Two hundred and forty-six problem situations are considered, revealing in their contents the fact that the Chinese had accumulated a variety of formulas for determining the areas and volumes of basic geometric shapes. Linear equations in one unknown were solved by a rule of false position. Systems of equations in two or three unknowns were solved simultaneously by computing board techniques that are strikingly similar to modern matrix methods. While algebraists of the ancient world such as Diophantus or Brahmagupta used various criteria to distinguish between the variables in a linear equation,[10] the Chinese relied on the organizational proficiency of their counting board to assist them in this chore. Using a counting board to work a system of equations allowed the Chinese to easily distinguish between different variables.

Consider the following problem from the *Chiu chang* and the counting board approach to its solution.

> Of three classes of cereal plants, 3 bundles of the first; 2 of the second and 1 of the third will produce 39 *tou* of corn after threshing; 2 bundles of the first; 3 of the second and 1 of the third will produce 34 *tou*; while 1 of the first, 2 of the second and 3 of the third will produce 26 *tou*. Find the measure of corn contained in one bundle of each class.[11]
>
> (1 *tou* = 10.3 liters)

This problem would be set up on the counting board as:

1	2	3	1st class grain
2	3	2	2nd class grain
3	1	1	3rd class grain
26	34	39	Number of *tou*

Using familiar notation this matrix of numbers is equivalent to the set of equations

$$3x+2y+z=39$$
$$2x+3y+z=34$$
$$x+2y+3z=26$$

which are reduced in their tabular form by appropriate multiplications and subtraction to

$$\begin{aligned} 3x+2y+\ z &= 30 \\ 36y\quad &= 153 \\ 36z &= 99 \end{aligned} \qquad \text{and} \qquad \begin{aligned} 36x\quad &= 333 \\ 36y\quad &= 153 \\ 36z &= 99. \end{aligned}$$

Thus $x=333/36$, $y=153/36$ and $z=99/36$.

A companion problem from the *Chiu chang* involves payment for livestock and results in the system of simultaneous equations:

$$-2x+5y-13z=1000$$
$$3x-9y+3z=0$$
$$-5x+6y+8z=-600.$$

Rules provided for the solution treat the addition and subtraction of negative numbers in a modern fashion; however, procedures for the multiplication and division of negative numbers are not found in a Chinese work until the Sung dynasty (+1299). Negative numbers were represented in the computing scheme by the use of red rods, while black computing rods represented positive numbers. Zero was indicated by a blank space on the counting board. This evidence qualifies the Chinese as being the first society known to use negative numbers in mathematical calculations.

The *Chou pei* contains an accurate process of extracting square roots of numbers. The ancient Chinese did not consider root extraction a separate process of mathematics but rather merely a form

166536 ÷ 648

Counting board layout		Accompanying rod computations	Explanations
2	(quotient)	166500	200 is chosen
		$-120000=(200\times600)$	as the first
166536	(dividend)	46500	partial
		$-8000=(200\times40)$	quotient
648	(divisor)	38500	
		$-1600=(200\times8)$	
		36900	
25		36930	50 is chosen
		$-30000=(50\times600)$	as the second
36936		6930	partial
		$-2000=(50\times40)$	quotient
648		4930	
		$400=(50\times8)$	
		4530	
257		4536	7 is chosen
		$-4200=(7\times600)$	as the third
4536		336	partial
		$-280=(7\times40)$	quotient
648		56	
		$-56=(7\times8)$	
		0	process is finished

FIGURE 7.

of division.[12] Let us examine the algorithm for division and its square root variant. The division algorithm is illustrated in FIGURE 7 for the problem 166536 ÷ 648. The Chinese technique of root extraction depends on the algebraic proposition

$$(a+b+c)^2=a^2+2ab+b^2+2(a+b)c+c^2$$
$$=a^2+(2a+b)b+(2[a+b]+c)c$$

which is geometrically substantiated by the diagram given in FIGURE 8. This proposition is incorporated directly into a form of division where $\sqrt{N}=a+b+c$. The counting board process for extracting the square root of 55225 is briefly outlined in FIGURE 9. Root extraction was not limited to three digit results, for the Chinese were able to continue the process to several decimal places as needed. Decimal fractions were known and used in China as far back as the 5th century B.C. Where a root was to be extracted to several decimal places, the computers achieved greater accuracy by use of the formulae $\sqrt[n]{m}=\sqrt[n]{m10^{kn}}/10^k$.[13] Cube root extraction was conceived on a similar geometric-algebraic basis and performed with equal facility.

Historians of mathematics often devote special consideration to the results obtained by ancient societies in determining a numerical value for π as they believe that the degree of accuracy achieved supplies a comparative measure for gauging the level of mathematical skill present in the society. On the basis of such comparisons, the ancient Chinese were far superior to their contemporaries in computational mathematical ability. Aided by a number system that included the decimalization of fractions and the possession of an accurate root extraction process the Chinese had obtained by the first century a value of π of 3.15147. The scholar Liu Hui in a third century commentary on the *Chiu chang* employed a "cutting of the circle method"—determining the area of a circle with known radius by polygonal approximations—to determine π as 3.141024. A successor, Tsu Chung-chih, refined the method in the fifth century to derive the value of π as 355/113 or 3.1415929.[14] This accuracy was not to be arrived at in Europe until the 16th century.

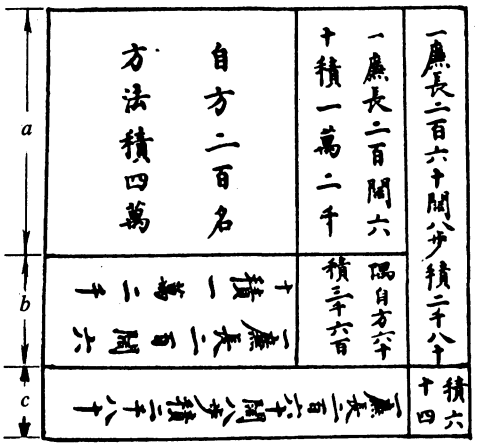


FIGURE 8.

A geometric "proof" (FIGURE 8) of the algebraic proposition (see p. 14) which justifies the calculations (FIGURE 9) leading to $\sqrt{55225}=235$. The 1 in the upper box represents an indexing rod that determines the decimal value of the divisors used. At the beginning of the process, it is moved to the left in jumps of two decimal places until it establishes the largest power of ten that can be divided into the designated number. After each successful division, the rod is indexed two positional places to the right.

Algebraic Significance	Numerical entries on board
N 1	55225 1
a $N-a^2$ $a\times 10000$ 10000	2 15225 20000 10000
$a+b$ $N-a^2$ $(2a+b)b\times 100$ $(2a+b)\times 100$ 100	23 15225 12900 4300 100
$a+b$ $N-[a^2+(2a+b)b]$ $(2a+b)\times 100$ 100	23 2325 4300 100
$a+b+c$ $N-(a+b)^2-[2a(a+b)+c]c$ $2(a+b)+c$	235 0 465

FIGURE 9.

Trends in Chinese Algebraic Thought

While the Chinese computational ability was indeed impressive for the times, their greatest accomplishments and contributions to the history of mathematics lay in algebra. During the Han period, the square and cube root extraction processes were being built upon to obtain methods for solving quadratic and other higher order numerical equations. The strategy for extending the square root process to solve quadratic equations was based on the following line of reasoning. If $x^2=289$, 10 would be chosen as a first entry approximation to the root, then

$$289-(10)^2=189.$$

Let the second entry of the root be represented by y; thus, $x=10+y$ or $(10+y)^2=289$ which, if expanded, gives the quadratic equation $y^2+20y-189=0$. By proceeding to find the second entry of the square root of 289, 7, we obtain the positive root for the quadratic $y^2+20y-189=0$.[15]

By the time of Sung Dynasty in the 13th century, mathematicians were applying their craft to solve such challenging problems as:

> This is a round town of which we do not know the circumference or diameter. There are four gates (in the wall). Three *li* from the northern (gate) is a high tree. When we go outside of the southern gate and turn east, we must walk 9 *li* before we see the tree. Find the circumference and the diameter of the town. (1 *li* = .644 kilometers)

If the diameter of the town is allowed to be represented by x^2, the distance of the tree from the northern gate, a, and the distance walked eastward, b, the following equation results.

$$x^{10}+5ax^8+8a^2x^6-4a(b^2-a^2)x^4-16a^2b^2x^2-16a^3b^2=0.$$

For the particular case cited above, the equation becomes

$$x^{10}+15x^8+72x^6-864x^4-11{,}664x^2-34{,}992=0.$$

Sung algebraists found the diameter of the town to be 9 *li*.[16]

The earliest recorded instance of work with indeterminate equations in China can be found in a problem situation of the *Chiu chang* where a system of four equations in five unknowns results.[17] A particular solution is supplied. A problem in the third century *Mathematical Classic of Sun Tzu*, [*Sun Tzu suan ching*,] concerns linear congruence and supplies a truer example of indeterminate analysis.

> We have things of which we do not know the number; if we count by threes, the remainder is 2; if we count by fives, the remainder is 3; if we count by sevens, the remainder is 2. How many things are there?[18]

In modern form, the problem would be represented as:

$$N\equiv 2 \pmod 3 \equiv 3 \pmod 5 \equiv 2 \pmod 7.$$

Sun's solution is given by the expression

$$70\times 2+21\times 3+15\times 2-105\times 2=23$$

which when analysed gives us the first application of the Chinese Remainder Theorem.

If $m_1,\ldots,m_k$ are relatively prime in pairs, there exist integers x for which simultaneously $x\equiv a_1 \pmod{m_1},\ldots,x\equiv a_k \pmod{m_k}$. All such integers x are congruent modulo $m=m_1m_2\ldots m_k$. The existence of the Chinese Remainder Theorem was communicated to the west by Alexander Wylie, an English translator and mathematician in the employ of the nineteenth century Chinese court. Wylie recorded his findings in a series of articles, "Jottings on the Science of the Chinese; Arithmetic" which appeared in the *North China Herald* (Aug.–Nov.) 1852. The validity of the theorem was questioned until it was recognized as a variant of a formula developed by Gauss.[19]

Perhaps the most famous Chinese problem in indeterminate analysis, in the sense of its transmission to other societies, was the problem of the "hundred fowls" (ca 468).

> A cock is worth 5 *ch'ien*, a hen 3 *ch'ien*, and 3 chicks 1 *ch'ien*. With 100 ch'ien we buy 100 fowls. How many cocks, hens, and chicks are there?
>
> (*ch'ien*, a small copper coin)

The development of algebra reached its peak during the later part of the Sung and the early part of the following Yuan dynasty (13th and 14th centuries). Work with indeterminate equations and higher order numerical equations was perfected. Solutions of systems of equations were found by using methods that approximate an application of determinants, but it wasn't until 1683 that the Japanese Seki Kowa, building upon Chinese theories, developed a true concept of determinants.

Work with higher numerical equations is facilitated by a knowledge of the binomial theorem. The testimony of the *Chiu chang* indicates that its early authors were familiar with the binomial expansion $(a+b)^3$, but Chinese knowledge of this theorem is truly confirmed by a diagram (FIGURE 10) appearing in the 13th century text *Detailed Analysis of the Mathematical Rules in the Nine Chapters*. [*Hsiang chieh chiu chang suan fa*.]. It seems that "Pascal's Triangle" was known in China long before Pascal was even born.

While mathematical activity continued in the post-Sung period, its contributions were minor as compared with those that had come before. By the time of the Ming emperors in the 17th century, western mathematical influence was finding its way into China and the period of indigenous mathematical accomplishment had come to an end.

Conclusions

Thus, if comparisons must be made among the societies of the pre-Christian world, the quality of China's mathematical accomplishments stands in contention with those of Greece and Babylonia, and

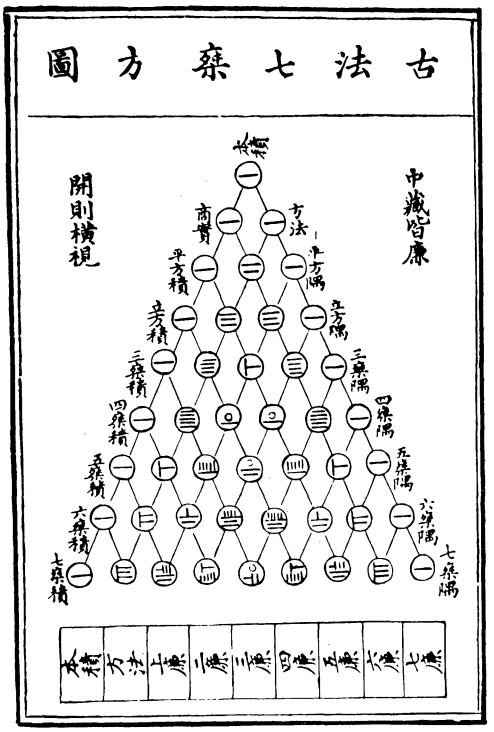


FIGURE 10.

during the period designated in the West as pre-Renaissance, the sequence and scope of mathematical concepts and techniques originating in China far exceeds that of any other contemporary society. The impact of this knowledge on the subsequent development of western mathematical thought is an issue that should not be ignored and can only be resolved by further research. In part, such research will have to explore the strength and vitality of Arabic-Hindu avenues of transmission of Chinese knowledge westward. The fact that western mathematical traditions are ostensibly based on the logico-deductive foundations of early Greek thought should not detract from considering the merits of the inductively-conceived mathematics of the Chinese. After all, deductive systemization is a luxury afforded only after inductive and empirical experimentation has established a foundation from which theoretical considerations can proceed. Mathematics, in its primary state, is a tool for societal survival; once that survival is assured, the discipline can then become more of an intellectual and aesthetic pursuit. Unfortunately, this second stage of mathematical development never occurred in China. This phenomenon—the fact that mathematics in China, although developed to a high art, was never elevated further to the status of an abstract deductive science—is yet another fascinating aspect of Chinese mathematics waiting to be explained.

Notes

1. Morris Kline, *Mathematics: A Cultural Approach* (Reading, Mass.: Addison-Wesley Publishing Co. 1962) p. 12.
2. In his 712 page *A History of Mathematics* (New York: John Wiley & Sons Inc., 1968) Carl Boyer devotes 12 pages to Chinese contributions; the latest revised edition of Howard Eves, *An Introduction to the History of Mathematics*, (New York: Holt, Rinehart and Winston, 1976) contains 6 pages on the history of Chinese mathematics. The contents of these pages are based on information given in an article by D. J. Struik, "On Ancient Chinese Mathematics," *The Mathematics Teacher* (1963), 56: 424–432 and represent little of Eves' own research.
3. John Greenleaf Whittier, "The Chapel of the Hermits."
4. Under this system, the universe is ruled by Heaven through means of a process called the *Tao* ("the Universal way"). Heaven acting through the Tao expresses itself in the interaction of two primal forces, the *Yin* and the

Yang. The *Yang*, or male force, was a source of heat, light and dynamic vitality and was associated with the sun; in contrast, the *Yin*, or female force, flourished in darkness, cold and quiet inactivity and was associated with the moon. In conjunction, these two forces influenced all things and were present individually or together in all physical objects and situations. In the case of numbers, odd numbers were *Yang* and even, *Yin*. For a harmonious state of being to exist, *Yin-Yang* forces had to be balanced.

5. For a fuller discussion of Chinese magic squares, see Schyler Camman, "Old Chinese Magic Squares", *Sinologica* (1962), 7 : 14–53; Frank Swetz, "Mysticism and Magic in the Number Squares of Old China," *The Mathematics Teacher* (January, 1978), 71: 50–56.
6. The evolution of counting rod numerals continued for about 3000 years in China, i.e., 14th century BC–13th century AD. For a discussion of this process, see Joseph Needham, *Science and Civilization in China* (Cambridge: Cambridge University Press, 1955) vol. 3. pp. 5–17.
7. A strong case for this theory has been made by Wang Ling, "The Chinese Origin of the Decimal Place—Value System in the Notation of Numbers". Communication to the 23rd International Congress of Orientalists, Cambridge, 1954.
8. Although a 3, 4, 5 right triangle is used in the demonstration, the Chinese generalized their conclusion for all right triangles. The 3, 4, 5 triangle was merely a didactical aid.
9. See Frank Swetz, "The 'Piling Up of Squares' in Ancient China," *The Mathematics Teacher* (1977), 70: 72–79.
10. Diophantus (275 AD) spoke of unknowns of the first number, second number, etc., whereas Brahmgupta (628 AD) used different colors in written computations to distinguish between variables.
11. *Chiu chang suan shu*, Fang Chheng (chapter 8), problem 1.
12. For a discussion of the Chinese ability at root extraction, see Wang Ling and Joseph Needham, "Horner's Method in Chinese Mathematics: Its Origins in the Root Extraction Procedures of the Han Dynasty", *T'oung Pao* (1955), 43: 345–88; Lam Lay Yong, "The Geometrical Basis of the Ancient Chinese Square-Root Method", *Isis* (Fall, 1970), pp. 92–101.
13. A lengthy discussion of the use of this formula in Europe is given in D. E. Smith, *History of Mathematics* (New York: Dover Publishing Co., 1958 reprint) vol. II, p. 236.
14. The evolution of π in China is traced out in Lee Kiong-Pong, "Development of π in China", *Bulletin of the Malaysian Mathematical Society* (1975), 6:40–47.
15. An actual computational procedure used in solving quadratics can be found in Ho Peng Yoke, "The Lost Problems of the Change Ch'iu-chien Sual Ching, a Fifth Century Chinese Mathematical Manual," *Oriens Extremus* (1965), 12.
16. For a detailed discussion of the solution of this problem see Ulrich Libbrecht, *Chinese Mathematics in the Thirteenth Century* (Cambridge, Mass.: The MIT Press, 1973) pp. 134–40.
17. *Chiu chang suan shu*, chapter 8, problem 13:

    There is a common well belonging to five families; (if we take) 2 lengths of rope of family X, the remaining part equals 1 length of rope of family Y; the remaining part from 3 ropes of Y equals 1 rope of Z; the remaining part from 4 ropes of Z equals 1 rope of V; the lacking part remaining from 5 ropes of V equals 1 rope of U; the remaining part from 6 ropes of U equals 1 rope of X. In all instances if one gets the missing length of rope, the combined lengths will reach (the water). Find the depth of the well and the length of the ropes.

    If we let W equal the depth of the well, the following system of equations result:

$$\begin{aligned} 2X+Y&=W\\ 3Y+Z&=W\\ 4Z+V&=W\\ 5V+U&=W\\ 6U+X&=W \end{aligned}$$

    which are readily reduced to:

$$\begin{aligned} 2X-2Y-Z&=0\\ 2X+Y-4Z-V&=0\\ 2X+Y-5V-U&=0\\ X+Y-6U&=0 \end{aligned}$$

18. *Sun Tzu suan ching*, chapter 3, problem 10.
19. See the discussion of the Chinese Remainder Theorem in Oystein Ore, *Number Theory and its History*, (New York: McGraw-Hill Inc., 1948) pp. 245–49.

NOTES

The Circumradius-Inradius Inequality for a Simplex

MURRAY S. KLAMKIN
University of Alberta
Edmonton, Alberta
Canada T6G 2G1

GEORGE A. TSINTSIFAS
Thessaloniki
Greece

The well-known inequality that the circumradius R of a triangle is at least twice the inradius r follows immediately from the identity [1], $R(R-2r)=OI^2$, where O and I denote the circumcenter and incenter, respectively of the triangle. This ubiquitous inequality occurs in the literature in many different equivalent forms [2,3]. For example, three such forms are

$$abc \geqslant (a+b-c)(b+c-a)(c+a-b), \tag{1}$$

$$1 \geqslant 8\sin\frac{A}{2}\sin\frac{B}{2}\sin\frac{C}{2}, \tag{2}$$

$$(y+z)(z+x)(x+y) \geqslant 8xyz, \tag{3}$$

where a,b,c and A,B,C denote the sides and angles of the triangle and x,y,z are arbitrary non-negative numbers (here $a=y+z$, $2x=b+c-a$, etc.). A proof of (3) follows immediately from the product of the three obvious inequalities

$$y+z \geqslant 2\sqrt{yz}\ , \quad z+x \geqslant 2\sqrt{zx}\ , \quad x+y \geqslant 2\sqrt{xy}\ .$$

A generalization of the inequality from triangles to n-dimensional simplexes is that $R \geqslant nr$ with equality if and only if the simplex is regular. The inequality $R \geqslant 2r$ for triangles is usually attributed to Euler (1765) [2]. However, this result was already known by Chapple (1746) [7]. A nice proof of the inequality $R \geqslant nr$ has been given by Fejes–Tóth [4] by showing that the simplexes of maximum and minimum volume inscribed and circumscribed, respectively, to a given sphere are regular. Here, we give an elementary proof of $R \geqslant nr$ plus some properties of an orthocentric simplex. (The less experienced reader may find it instructive perhaps to check this proof first in the special cases corresponding to $n=2$ and 3, the triangle and the tetrahedron.)

Although our proof is for a simplex, we will take our figure to be a triangle for simplicity (see FIGURE 1). Let A_i and F_i (for $i=1,2,\ldots,n+1$) denote, respectively, the vertices and contents of the opposite $(n-1)$-dimensional faces of an n-dimensional simplex of volume V. Also, let h_i and r_i denote the distances from A_i and from the circumcenter O of the simplex to the face opposite A_i, respectively. Since, $R+r_i \geqslant h_i$, $\Sigma RF_i+\Sigma r_iF_i \geqslant \Sigma h_iF_i$. Moreover, the volume of the n-dimensional simplex is given by h_iF_i/n, so by evaluating the volume of the simplex in three different ways, we get, for every subscript i, that

$$nV = h_iF_i = \Sigma r_iF_i = r\Sigma F_i.$$

Hence,

$$R\Sigma F_i \geqslant \Sigma h_iF_i - \Sigma r_iF_i = (n+1)nV - nV = nr\Sigma F_i$$

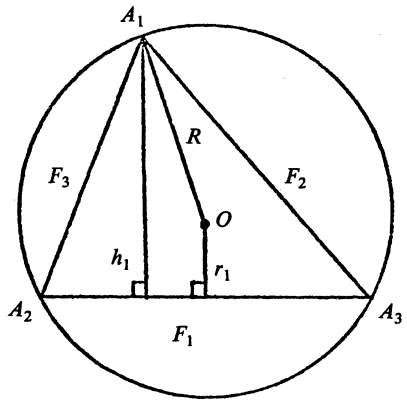


FIGURE 1.

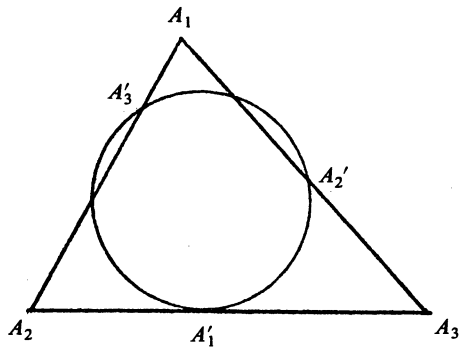


FIGURE 2.

or $R \geqslant nr$. There is equality if and only if $R+r_i=h_i$ for all i, or equivalently, if and only if all the altitudes are all concurrent at the circumcenter. (Such simplexes, whose altitudes are concurrent, are said to be **orthocentric**.)

The latter conditions require the simplex to be regular. For a proof, let $\mathbf{A}_i$ denote vectors from a common origin (which we take to be the orthocenter H for subsequent use) to the vertices of the simplex. Then,

$$\mathbf{A}_k\cdot(\mathbf{A}_i-\mathbf{A}_j)=0, \qquad \text{for all } k\neq i,j \text{ (orthogonal property)}, \tag{4}$$

$$\mathbf{A}_i\cdot\mathbf{A}_i=\mathbf{A}_j\cdot\mathbf{A}_j, \qquad \text{for all } i,j \ (H \text{ and } O \text{ coincide}). \tag{5}$$

Direct computation of the difference of the squares of the lengths of any two concurrent edges of the simplex, using (4) and (5), yields

$$(\mathbf{A}_j-\mathbf{A}_k)^2-(\mathbf{A}_i-\mathbf{A}_k)^2=2\mathbf{A}_k\cdot(\mathbf{A}_i-\mathbf{A}_j)=0. \tag{6}$$

Thus all the edges are congruent and the simplex is regular.

We have just shown that if the orthocenter of an orthocentric simplex coincides with the circumcenter, then the simplex must be regular. Additionally, we now show that if either the incenter I or the centroid G coincides with the orthocenter, the simplex must also be regular. Note, however, that even if O, I and G all coincide for a tetrahedron, the tetrahedron need not be regular.

First, as is well known, the orthocenter divides each altitude into two segments whose product is a constant. This follows immediately from (4) which implies that $\mathbf{A}_i\cdot\mathbf{A}_j=$ constant for all $i\neq j$ (just interpret the dot product). If I coincides with H, each "shorter" segment of each altitude (from H to a face) is of constant length r. Consequently, the other segment of each altitude (from H to a vertex) is also of constant length and equals R. Thus H also coincides with O and, as shown before, the simplex must be regular.

Finally, we know that the centroid G divides each median of the simplex into segments of constant ratio $(n:1)$. Hence, if G also coincides with H, it follows again from the altitude segment property that H also coincides with O and again the simplex must be regular. A somewhat longer proof of this is given in [5]. Also, for a much more comprehensive treatment of orthocentric and general simplexes, see [6] and the references given therein.

The history of our problem involves an interesting alternative proof. In a U.S.A. Olympiad practice session, the first author had assigned the problem of proving $R \geqslant 3r$ for tetrahedra. Only one

student, Miller Puckette (now an undergraduate at M.I.T.), came up with a solution, albeit incomplete. It was completed with some joint effort. A sketch of this proof for triangles follows; the extension to simplexes is almost immediate.

Consider any inscribed triangle $A_1'A_2'A_3'$ to the given triangle $A_1A_2A_3$. Puckette claimed without proof that the circumradius R' of $A_1'A_2'A_3'$ is greater than or equal to r. (Coincidentally, this same result for simplexes was also used subsequently without proof by the second author in a problem proposal submitted to the *American Mathematical Monthly*.) To prove this result, assume $R' \neq r$ so that circumcircle C' must at least intersect one of the three sides of $A_1A_2A_3$ in two distinct points. One possibility is given by FIGURE 2. We now perform a homothetic dilatation with point A_2 as a center on circle C' such that it becomes also tangent to side A_1A_3. Finally, we perform another homothetic dilatation with point A_3 as a center on the previous circle (which is tangent to A_1A_3 and A_2A_3) such that it becomes the inscribed circle. Since all these transformations are proper dilatations, $R' \geqslant r$.

To complete the proof, select points A_i' to be the centroid of the edge (face) opposite A_i. It follows simply by vectors that the inscribed triangle (simplex) is similar to the original one with ratio of sides $1:2$ $(1:n)$. Whence the circumradius of the inscribed figure is $R/2$ (R/n) which is greater than or equal to r. There is equality here (in the general case) if and only if the centroid of each face coincides with the point of tangency of the inscribed sphere with that face. As a challenge to the reader, we leave it as an exercise to show that the last assertion implies that the simplex is regular.

References

[1] R. A. Johnson, Advanced Euclidean Geometry, Dover, New York, 1960, p. 186.
[2] O. Bottema, R. Z. Djordjević, R. R. Janić, D. S. Mitrinović, P. M. Vasić, Geometric Inequalities, Wolters–Noordhoff, Groningen, 1969, p. 48.
[3] M. S. Klamkin, Duality in triangle inequalities, Ford Motor Company Preprint, July 1971 (also, see Notices of AMS, August 1971, p. 782).
[4] L. Fejes–Tóth, Regular Figures, Pergamon Press, London, 1964, pp. 312–313.
[5] C. M. Petty, D. Waterman, An external theorem for n-simplexes, Monat. Math., 59(1955) 320–322.
[6] S. R. Mandan, Altitudes of a simplex in n-space, J. Australian Math. Soc., 2(1961/62) 403–424.
[7] J. S. Mackay, Historical notes on a geometrical theorem and its developments, Proc. Edinburgh Math. Soc., 5(1887) 62–63.

An Attrition Problem of Gambler's Ruin

W. D. KAIGH
University of Texas at El Paso
El Paso, TX 79968

Recently a student solicited my advice in an attempt to improve his skill at a certain table game in which players employ tokens called "armies" and repeatedly "attack" in an attempt to eliminate the "armies" of opposing players. The outcome of an individual conflict between two opposing "armies" is determined randomly through the toss of dice and the token of the loser is removed from the board. An avid enthusiast of the game, my student wished to employ strategy based on sound probabilistic considerations rather than mere intuition.

Our mathematical formulation of this game turned out to be a modification of the classical problem of gambler's ruin which we call the attrition ruin problem. Analysis of this problem provides new insight into the classical ruin problem. Moreover, outcomes of certain athletic events such as the World Series of baseball (each team begins with four "tokens") may be viewed in the context of an attrition ruin problem. Unfortunately, however, the theoretical developments which follow have not

been subjected to the test of empirical verification due to my instigating student's ill-timed discovery of the game of backgammon.

Let us begin our study with a sequence of independent repetitions of a game between two players in which the winner of each trial is determined by the outcome of a certain random experiment. The opponents A and B with initial fortunes a and b respectively compete until the capital of one is exhausted and that player is "ruined". In the context of the classical ruin problem, one unit of capital is transferred from the loser to the winner at the conclusion of each trial. Here we suppose instead that each trial results in the forfeiture of a single unit for the loser and in no change for the winner. As usual, the matters of fundamental concern in this attrition form of the ruin problem are the probabilities of ruin for each player, the expected duration of the contest, and the effect of a change in stakes.

If p denotes the (fixed) probability that player A wins a given trial, then $q=1-p$ is the corresponding probability for player B. Let A^* be the event that player A is the ultimate winner (B is ruined) and denote by D the (random) duration of the contest. Then in the classical ruin problem (see Feller [1], pages 344–349),

$$P(A^*)=\begin{cases} a/(a+b) & \text{if } p=q=1/2 \\ \dfrac{(p/q)^{a+b}-(p/q)^b}{(p/q)^{a+b}-1} & \text{if } p\neq q \end{cases} \tag{1}$$

and

$$E(D)=\begin{cases} ab & \text{if } p=q=1/2 \\ \dfrac{a}{q-p}-\dfrac{a+b}{q-p}\dfrac{1-(q/p)^a}{1-(q/p)^{a+b}} & \text{if } p\neq q \end{cases} \tag{2}$$

To analyse the attrition ruin problem, first note that if player A is the ultimate winner, the total amount of capital L_A which he loses is a random variable satisfying $0\leqslant L_A\leqslant a-1$ so the probability of A's eventually winning is given by

$$P(A^*)=\sum_{x=0}^{a-1} P(A^*;L_A=x).$$

Since the ruin of player B requires the complete loss of his initial capital b, the event A^* and $L_A=x$ involves exactly $b+x$ trials of the random experiment. By a standard counting argument we obtain

$$P(A^*;L_A=x)=\binom{b+x-1}{x}p^bq^x$$

which is the negative binomial probability corresponding to exactly x failures preceding the bth success in independent, identical Bernoulli trials with success probability p. It follows that

$$P(A^*)=\sum_{x=0}^{a-1}\binom{b+x-1}{x}p^bq^x. \tag{3}$$

To determine the probability of the event B^* that player B is the ultimate winner, it suffices to either compute $1-P(A^*)$ or simply replace a,b,p,q by b,a,p,q respectively in (3). To provide a numerical illustration we take $a=b=4, p=0.6, q=0.4$ in (3) and obtain the exact probabilities $P(A^*)=0.710208$, $P(B^*)=0.289792$.

Next we determine the expected duration of the attrition process through a direct computation involving the probability distribution of the random variable D. For any duration x (an integer) with $\min\{a,b\}\leqslant x\leqslant a+b-1$

$$P(D=x)=P(A^*;L_A=x-b)+P(B^*;L_B=x-a)$$

$$=\binom{x-1}{x-b}p^bq^{x-b}+\binom{x-1}{x-a}p^{x-a}q^a.$$

Note that in contrast with this attrition case, the duration of the contest in the classical ruin problem is not a bounded random variable and the derivation of the probability distribution is more difficult.

By manipulating binomial coefficients, we can compute the expected value of the duration:

$$E(D)=\sum_{x=\min\{a,b\}}^{a+b-1} xP(D=x)$$

$$=\sum_{x=b}^{a+b-1} x\binom{x-1}{x-b}p^b q^{x-b}+\sum_{x=a}^{a+b-1} x\binom{x-1}{x-a}p^x$$

$$=(b/p)\sum_{x=b}^{a+b-1}\binom{x}{x-b}p^{b+1}q^{x-b}+(a/q)\sum_{x=a}^{a+b-1}\binom{x}{x-a}p^{x-a}q^{a+1}.$$

Replacing $x-b$ and $x-a$ by x in the above provides

$$E(D)=(b/p)\sum_{x=0}^{a-1}\binom{b+x}{x}p^{b+1}q^x+(a/q)\sum_{x=0}^{b-1}\binom{a+x}{x}p^x q^{a+1}. \tag{4}$$

To illustrate (4) numerically we choose again $a=b=4$, $p=0.6$, $q=0.4$ and compute $E(D)=5.69728$.

The expressions given in (3) and (4) are exact but cumbersome and do not lend themselves to computation when the players have large initial fortunes. Although a relation between the negative binomial distribution and the binomial distribution may be used to evaluate the summations for moderate a and b, we require an approximation for large a and b.

As indicated previously, the summation in (3) represents $P(X<a)$ where X is a random variable with negative binomial distribution and parameters b and p. Such a random variable X has $E(X)=bq/p$, $\mathrm{Var}(X)=bq/p^2$ and may be expressed as the sum of b independent, identically distributed geometric random variables. By an application of the central limit theorem, we have, for large b,

$$P(A^*)=P\left[\frac{X-bq/p}{(bq/p^2)^{1/2}}<\frac{a-bq/p}{(bq/p^2)^{1/2}}\right]$$

$$\approx P\left[Z<(ap-bq)/(bq)^{1/2}\right]$$

where Z is standard normal. In other words, if we denote the cumulative distribution function of the standard normal by Φ, then (for large b)

$$P(A^*)\approx\Phi\left[(ap-bq)/(bq)^{1/2}\right]. \tag{5}$$

Similar analysis on the terms of (4) provides for large a and b

$$E(D)\approx(b/p)\Phi\left[(ap-bq)/(bq)^{1/2}\right]+(a/q)\Phi\left[(bq-ap)/(ap)^{1/2}\right]. \tag{6}$$

In accordance with the usual rule of thumb concerning applications of the central limit theorem, numerical comparisons indicate that the normal approximation is quite adequate when the amounts of initial capital are at least thirty.

From (5) and the symmetry about zero of the standard normal distribution, we see after minor manipulation that (asymptotically) player A is favored in the contest whenever $ap-bq>0$, i.e., the ratio of his capital a to the total capital $a+b$ exceeds his failure probability q.

Next we consider the effect of changing stakes if initial capital remains fixed. If the stakes at each trial are increased to $k>1$ units in the classical ruin problem, $P(A^*)$ increases if $p<q$ and is unchanged if $p=q$. In the attrition ruin problem, examination of (5) and monotonicity of the standard normal cumulative distribution function indicate that (asymptotically) an increase in stakes increases $P(A^*)$ if $a/(a+b)<q$ and has no effect if $a/(a+b)=q$. Thus, the strategy of a player in the attrition

		Initial Capital			Attrition Probability of Ultimate Win		Classical Probability of Ultimate Win	
		Total	A	B	A	B	A	B
p	q	$a+b$	a	b	$P(A^*)$	$P(B^*)$	$P(A^*)$	$P(B^*)$
0.45	0.55	10	4	6	0.17	0.83	0.19	0.81
			5	5	0.38	0.62	0.27	0.73
			6	4	0.64	0.36	0.36	0.64
		30	12	18	0.08	0.92	0.02	0.98
			15	15	0.30	0.70	0.05	0.95
			18	12	0.72	0.28	0.09	0.91
		100	40	60	0.00^+	1.00^-	0.00^+	1.00^-
			50	50	0.17	0.83	0.00^+	1.00^-
			60	40	0.86	0.14	0.00^+	1.00^-
0.50	0.50	10	4	6	0.25	0.75	0.40	0.60
			5	5	0.50	0.50	0.50	0.50
			6	4	0.75	0.25	0.60	0.40
		30	12	18	0.16	0.84	0.40	0.60
			15	15	0.50	0.50	0.50	0.50
			18	12	0.84	0.16	0.60	0.40
		100	40	60	0.03	0.97	0.40	0.60
			50	50	0.50	0.50	0.50	0.50
			60	40	0.97	0.03	0.60	0.40

Ultimate Win Probabilities.

TABLE 1.

game depends on both his failure probability and his share of the total initial capital and a favored player should resist any effort to increase stakes.

As an illustration, we list in Table 1 numerical values obtained from (3) and (5) for $P(A^*)$ for several choices of p,a,b. To facilitate comparison we exhibit the corresponding quantities from (1) for the classical ruin problem. We conclude by noting that whereas in the classical problem a skillful player ($p>1/2$) with small initial capital may well ruin an opponent with larger initial fortune, such an outcome occurs less frequently in an attrition contest.

Reference

[1] W. Feller, An Introduction to Probability Theory and Its Applications, 3rd ed., Wiley, New York (1968).

Lattice Points and Area-Diameter Relation

JOSEPH HAMMER
University of Sydney
Sydney, Australia 2006

For each convex domain S in the plane, let $A(S)$ denote the area and $D(S)$ the diameter of S. Let ϕ denote the unique positive real number satisfying $\sin\phi = \pi/2-\phi$ and set $\lambda=2\sqrt{2}\sin(\varphi/2)$ (≈ 1.144). Recently Scott [2] showed that if $A(S)>r\lambda D(S)$, where r is any positive integer, then S contains at least r lattice points in its interior. In this note we strengthen Scott's lower bound and then give an upper bound for the minimum number of lattice points inside S.

First, we will show that if $A(S)>r\lambda D(S)$, then S contains at least r^2 lattice points. If $r=0$, the result is clear. So suppose $r\geqslant 1$ and consider the similarity transformation

$$S\to S'=\frac{1}{r}S=\left\{\frac{1}{r}Y: Y\in S\right\}.$$

Obviously, $A(S')=A(S)/r^2$ and $D(S')=D(S)/r$. Now let $T=(t_1,t_2)$ be a lattice point with $O\leqslant t_1,t_2\leqslant r-1$ and consider the translate S'' of S' given by $S''=S'-(1/r)T=\{X-(1/r)T: X\in S'\}$. Obviously, $A(S'')=A(S'), D(S'')=D(S')$, and

$$\frac{A(S'')}{D(S'')}=\frac{A(S')}{D(S')}=\frac{1}{r}\frac{A(S)}{D(S)}>\lambda.$$

By Scott's theorem (in the case $r=1$), S'' contains a lattice point G. Hence S' contains the point $G+(1/r)T$, and so S contains the point $P=r(G+(1/r)T)=rG+T$. But $T=(t_1,t_2)$ might have been chosen in r^2 different ways, for we could have selected each of t_1,t_2 in r different ways. Therefore we have r^2 distinct lattice points $P=(p_1,p_2)$ in S. These are distinct, since $p_i\equiv t_i \pmod r (i=1,2)$ and the t_i are a complete set of residues mod r. This means S contains at least r^2 lattice points. In summary: *if $A(S)>r\lambda D(S)$ then the minimum number of lattice points inside S is at least r^2.*

We conclude by establishing an upper bound for the minimum number of lattice points inside S: *If $A(S)>r\lambda D(S)$, then the minimum number of lattice points inside S is at most* $[(2r\lambda)^2/\pi]$, where $[q]$ denotes the integral part of q. This is an immediate consequence of a result in Niven and Zuckerman [1; Thm. 3], namely, if R is a closed, bounded measurable region, then the minimum number of lattice points in R as R moves around the plane, is always less than the measure of R. Now, let C denote a circle of radius $2r\lambda/\pi$. Then $A(C)/D(C)=r\lambda$. Hence, by the theorem of Niven and Zuckerman, we have the result.

References

[1] I. Niven and H. S. Zuckerman, Lattice points in regions, Proc. Amer. Math. Soc., 18 (1967) 364–370.

[2] P. R. Scott, Area-diameter relations for two-dimensional lattices, this MAGAZINE, 47 (1974) 218–221.

Countable Yet Nowhere First Countable

RICHARD WILLMOTT

Queens University

Kingston, Canada K7L 3N6

Examples of countable non-first countable topological spaces tend to be non-first countable at only one point [1, examples 26, 35, 98, 114] or non-elementary [2]. We give here an example of a simple Hausdorff topology on the natural numbers which is nowhere first countable. Our space is 0-dimensional, which assures it of a fairly rich structure. For example, it is T_3 (since 0-dimensional) and Lindelöf and so paracompact.

Let N be the set of natural numbers and let $p(n)$ denote the nth prime number. Call any power of $p(m), p(m)^k, k\geqslant 1$ an **offspring** of m. Thus, for example, $27(=p(2)^3)$ is an offspring of $2(=p(1))$, which is an offspring of 1. Let $\{m\}^1$ be the set of all offspring of m, and for any $n>1$, $\{m\}^n$ be the set of all offspring of $\{m\}^{n-1}$. We will say j is an (nth generation) **descendent** of k, and k is an **ancestor** of j if for some n, $j\in\{k\}^n$. For example, 23^6 has ancestors 9, 2 and 1 $(23=p(9), 9=p(2)^2)$ and is a third generation descendent of 1. A number has an ancestor if and only if it is a prime power and since the

immediate ancestor of any prime power is unique, the set of ancestors of a prime power is strictly linearly ordered by descendence, and the number of the generation of a descendent is unique.

We now introduce a topology on $X=N$ by stipulating that $G\subset X$ is open if for each $x\in G$, there is some M such that for every $m\geqslant M$, $p(x)^m\in G$ (i.e., if $x\in G$, then G contains all offspring of x from some power of $p(x)$ on). Any set U defined by an inductive construction of the following type will be called a **fundamental neighborhood** of x: Choose $M_x\in N$ and let $B_1=\{p(x)^m : m\geqslant M_x\}$, the set of all offspring of x from those of power M_x on. For each $k\in B_1$, choose $M_k\in N$ and let $B_{2,k}=\{p(k)^m : m\geqslant M_k\}$, and then let $B_2=\bigcup_{k\in B_1} B_{2,k}$, a set of second generation descendents of x. Continue, and finally let $U=\{x\}\cup\bigcup_{k\in N} B_k$. It is clear that a fundamental neighborhood of x is open and contains only descendents of x, that the fundamental neighborhoods of x form a neighborhood base at x, and that x has a maximal fundamental neighborhood, namely x along with all of its descendents. The maximal fundamental neighborhood of x will be denoted by $O[x]$.

LEMMA. (a) *If $x,y\in X$ and neither is a descendent of the other, then $O[x]\cap O[y]=\varnothing$.*

(b) *If y is a descendent of x, then there is a fundamental neighborhood of x which is disjoint from $O[y]$.*

Proof. (a) If x and y had a common descendent, z, then both would be ancestors of z, and so one would have to be a descendent of the other.

(b) Suppose y is a descendent of x. Then for some m, $y=p(x)^m$ or y is a descendent of $p(x)^m$. Let

$$U=\{x\}\cup\bigcup_{k>m} O[p(x)^k].$$

Then U is a fundamental neighborhood of x and since $p(x)^m$ and $p(x)^k$ for $k>m$ have no common descendants, $U\cap O[y]=\varnothing$.

It now follows immediately that X is a Hausdorff space. Since X is countable, it also follows immediately that every set is an F_σ set and hence a G_δ set. Yet X is not first countable at any point, as the following shows. Suppose $G_1, G_2, \ldots$ is any countable family of open sets containing $x\in X$. Let b_1 be the least offspring of x in G_1 and $p(b_1)^{m_1}$ the least offspring of b_1 in G_1. Now let b_2 be the least offspring of x in G_2 which is greater than b_1 and $p(b_2)^{m_2}$ the least offspring of b_2 in G_2. Continue inductively and let

$$V=O[x]-\bigcup_{i\in N}\{p(b_i)^{m_i}\}.$$

Then V is an open set containing x which contains none of the G_i. Thus X is not first countable.

Finally we show that any fundamental neighborhood is closed, so that the fundamental neighborhoods of a point form a neighborhood base of open and closed sets. Suppose U is a fundamental neighborhood of x and $y\notin U$. If neither x nor y is a descendent of the other, then by the Lemma $O[x]\cap O[y]=\varnothing$, so $U\cap O[y]=\varnothing$. If x is a descendent of y, then by the Lemma there is a fundamental neighborhood of y disjoint from $O[x]$ and hence from U. If y is an nth generation descendent of x, there is a unique member $b_1, b_2, \ldots, b_{n-1}$ of each generation of descendents of x from which y is descended. Now if $z\in U$, and z is none of these b_i or x, and z is a member of one of the first n generations of descendents of x, then $O[y]\cap O[z]=\varnothing$ by the Lemma. But U is contained in the union of these $O[z]$ with $\{x, b_1, b_2, \ldots, b_{n-1}\}$. Hence $O[y]\cap U=\varnothing$. It follows that the complement of U is open and U is closed. Hence X is zero dimensional.

References

[1] L. Steen and J. Seebach, Counterexamples in Topology, Holt, New York, 1971.

[2] P. W. Harley, III, A countable nowhere first countable Hausdorff space, Canad. Math. Bull., 16 (1973) 441–442.

Inequalities for a Collection

Ralph P. Boas
Northwestern University
Evanston, IL 60201

Some people collect butterflies, some collect stamps, and some collect inequalities. (See the bibliography at the end of this note.) Here is a specimen:

$$\sin^{-1}\left(\sinh\tfrac{1}{2}x\right) \leqslant \sinh\left(\tfrac{1}{2}\sin^{-1}x\right), \qquad 0 \leqslant x \leqslant 1. \tag{1}$$

Like many other kinds of collectibles, inequalities should not appear in one's collection unless they have been authenticated—in this case, by proving them. How can we prove (1)? In the first place, it is not obviously false, since $\sinh\frac{1}{2}x < 1$ when $x \leqslant 1$, so that we can take its inverse sine; and (1) is valid at 0 by inspection and at 1 by computation.

Many inequalities can be established by calculus. For example, since $x - \sin x = 0$ at 0 and has a positive derivative, we deduce that $x - \sin x \geqslant 0$ for $x > 0$, and so $\sin x \leqslant x$. However, (1) does not seem to yield to this approach.

It is sometimes easier to solve a special problem by solving a general problem of which it is a special case. (That is how calculus solves problems about rates, areas, etc.) We might notice that (1) is an instance of a hypothetical inequality

$$f(g(x)) \leqslant g(f(x)). \tag{2}$$

If we can prove (2) under hypotheses that are satisfied by $f(x) = \sin^{-1}x$, $g(x) = \sinh\frac{1}{2}x$, we will have proved (1) and indeed much more. What can we expect to be the necessary properties of f and g?

We can notice several things about the f and g that appear in (1). They both have Maclaurin series with nonnegative coefficients (and $f(0) = g(0) = 0$); hence both are increasing functions with increasing derivatives. Also, $g(1)$ is rather less than $f(1)$ ($g(1) = 0.521$, $f(1) = \pi/2 = 1.57$). It is intuitive that when x is near 1, we have $f(x)$ large and $g(f(x))$ also large (since g increases), whereas $g(x)$ is rather small and so $f(g(x))$ is also rather small. This suggests that (1) should hold at least for x near 1, and that something in the nature of $f(1) > g(1)$ would be useful in (2). Further speculation along these lines suggests both the following theorem and a method for proving it.

Theorem. *Let f be continuous with domain $0 \leqslant x < 1$ or $0 \leqslant x \leqslant 1, f(0) = 0, f(1) > 1$ (including the possibility that $f(1) = +\infty$); let g be continuous with domain the range of f, and $g(1) \leqslant 1$. Let $f(x)/x$ and $g(x)/x$ be strictly increasing on their domains. Finally let $f(x) \neq x$ for $0 < x < 1$. Then $f(g(x)) \leqslant g(f(x))$ for $0 < x < 1$.*

The proof will, of course, specialize to a proof of (1), or of any other case of (2), if we write it out for specific functions f and g. However, the general proof gives us more insight into why (1) and similar inequalities work, without being complicated by irrelevant special properties of the functions concerned. Thus, for example, our observation that f and g in (1) have Maclaurin series with nonnegative coefficients is much more than we actually need.

We observe first that since $f(x)/x$ increases, then if $0 < y \leqslant 1$ we have $f(xy)/(xy) \leqslant f(x)/x$ for $0 < x < 1$ (or $0 < x \leqslant 1$), and consequently

$$f(xy) \leqslant yf(x), \qquad 0 < x < 1, \qquad 0 < y \leqslant 1. \tag{3}$$

We also observe that there must be a number ε_1, $0 < \varepsilon_1 < 1$, such that $f(\varepsilon_1) = 1$ and $f(x) > 1$ for $1 > x \geqslant \varepsilon_1$, because $f(0) = 0 < 1$, whereas $f(1) > 1$, and so f takes the value 1 at least once; and only once because $f(x)$ increases strictly with x (since $f(x)/x$ increases). We then determine ε_k, $k > 1$, recursively: $f(\varepsilon_{k+1}) = \varepsilon_k$ and $f(x) \geqslant \varepsilon_k$ for $1 > x \geqslant \varepsilon_{k+1}$. Clearly $\{\varepsilon_k\}$ is a decreasing sequence of positive

numbers, which consequently has a limit ε. If $\varepsilon>0$, we would have $f(\varepsilon)=\lim_{k\to\infty} f(\varepsilon_k)=\varepsilon$ by continuity; but we assumed $f(\varepsilon)\neq\varepsilon$ for $\varepsilon>0$. Hence $\varepsilon_k\to 0$.

Let $\varepsilon_0=1$. Since $\varepsilon_k\to 0$, each x in $(0,1)$ is in one of the intervals $\varepsilon_{k+1}\leqslant x\leqslant\varepsilon_k, k=0,1,2,\ldots$. In this interval we have $f(x)\geqslant\varepsilon_k$ by the way in which the ε_k were constructed. Then (3) gives us

$$f(g(x))=f(x\cdot g(x)/x)\leqslant f(x)\,g(x)/x, \qquad \varepsilon_{k+1}\leqslant x\leqslant\varepsilon_k; \tag{4}$$

note that (3) is applicable because

$$g(x)/x\leqslant g(1)/1\leqslant 1. \tag{5}$$

In the interval $(\varepsilon_1,1)$ we have $f(x)>1$ and so $g(1)/1\leqslant g(f(x))/f(x)$ because $g(x)/x$ increases; if we substitute this into (5) and then (5) into (4), we get (2) for the interval $(\varepsilon_1,1)$.

If $k\geqslant 1$, then since $g(x)/x$ increases, we can continue (4) to get

$$f(g(x))\leqslant f(x)g(\varepsilon_k)/\varepsilon_k,$$

since $x\leqslant\varepsilon_k$. But $\varepsilon_k\leqslant f(x)$ and g increases, so that

$$f(g(x))\leqslant f(x)g(f(x))/f(x)=g(f(x)).$$

We have therefore established (2) for $\varepsilon_{k+1}\leqslant x<\varepsilon_k$, and consequently for $0<x<1$.

In the inequality (1) from which we started, we have $f(x)=\sin^{-1}x, g(x)=\sinh\frac{1}{2}x$. The Maclaurin series of f and g make it clear that $f(x)/x$ and $g(x)/x$ increase; and $f(x)\neq x$ since $\sin^{-1}x=x$ would imply $x=\sin x$, which is possible only at 0. The number $\frac{1}{2}$ in (1) has no particular significance; it is just a simple number σ for which $\sinh\sigma\leqslant 1$, so any number σ less than $\sinh^{-1}1=0.88137\ldots$ would do. We get

$$\sin^{-1}(\sinh\sigma x)\leqslant\sinh(\sigma\sin^{-1}x), \qquad 0<x<1. \tag{6}$$

Most people seem to find trigonometric functions more congenial than their inverses; you may therefore prefer to take $x=\sin u$, $0<u<\pi/2$; then

$$\sin^{-1}(\sinh(\sigma\sin u))\leqslant\sinh\sigma u, \qquad 0<u<\pi/2. \tag{7}$$

This would look simpler if we took the sine of both sides, which we can do if $\sinh\sigma u<\pi/2$ so that the sine is an increasing function. This requires that $\sigma u\leqslant\sinh^{-1}(\pi/2)$, i.e., $u\leqslant\sigma^{-1}\sinh^{-1}(\pi/2)$ (and of course $0<u<\pi/2$ also). Thus we have (8) $\sinh(\sigma\sin u)\leqslant\sin(\sinh\sigma u)$, $0<u<\min(\pi/2,\sigma^{-1}\sinh^{-1}(\pi/2))$. For $\sigma=\frac{1}{2}$ we have (8) for the same values of u as in (7), i.e. $0<u\leqslant\pi/2$, since $2\sinh^{-1}(\pi/2)=2.47>\pi/2$.

Here are some more applications of the theorem.

A. Let

$$f(x)=-\log(1-x), \qquad g(x)=x^a, \qquad a>1.$$

Then

$$\log\frac{1}{1-x^a}\leqslant\left(\log\frac{1}{1-x}\right)^a, \qquad 0<x<1.$$

Put $x=1-e^{-y}$, $y>0$; we get

$$\log\frac{1}{1-(1-e^{-y})^a}\leqslant y^a, \qquad y>0, \qquad a>1.$$

This inequality was conjectured by E. Beller and first proved by J. H. Wells by more advanced methods than those used here (cf. Problem E2695, *American Mathematical Monthly*).

B. Let

$$f(x)=\sin^{-1}x, \qquad 0<x\leqslant 1; \qquad g(x)=x^a, \qquad a>1.$$

The theorem yields

$$\sin^{-1}(x^a)\leqslant(\sin^{-1}x)^a, \qquad 0\leqslant x\leqslant 1.$$

Since $x\leqslant 1$ we can write $x=\sin y$, $0\leqslant y\leqslant\pi/2$. Then we have

$$\sin^{-1}((\sin y)^a)\leqslant y^a, \qquad 0\leqslant y\leqslant\pi/2.$$

If $y^a < \pi/2$, we can take the sine of both sides and obtain

$$(\sin y)^a \leqslant \sin(y^a), \qquad 0 \leqslant y^a \leqslant \pi/2, \qquad a > 1. \tag{8}$$

In particular,

$$\sin^2 y \leqslant \sin(y^2), \qquad 0 \leqslant y^2 \leqslant \pi/2. \tag{9}$$

Inequality (8) can in fact be established more concisely by more direct methods (cf. Problem E2720, *American Mathematical Monthly*).

C. In our examples so far, the hypothesis $f(x) \neq x$ follows readily because $f(x)/x$ increases strictly and $f'(0) = \lim_{x \to 0+} f(x)/x \geqslant 1$, so that $f(x)/x > 1$ for $x > 0$. It is, however, not essential to have $f'(0) \geqslant 1$. Let us try $f(x) = \alpha \sin^2 \beta x$. Here $f'(0) = 0$ and we have to select α and β so that (a) $f(x)/x$ increases, (b) $f(x) \neq x$, (c) $f(1) > 1$. For (a), we want $(\sin^2 \beta x)/x$ increasing, or, what is the same thing, $\sin^2 y / y$ increasing, where $y = \beta x$. This will be true as long as $(d/dy)(\sin^2 y / y) > 0$, i.e. $2y \sin y \cos y - \sin^2 y > 0$, or (if $0 < y < \pi/2$) $(\tan y)/y < 2$. Now $(\tan y)/y$ starts at 1 at $y = 0$ and increases because its derivative has the sign of $y \sec^2 y - \tan y = \frac{1}{2} \sec^2 y (2y - \sin 2y) > 0$. Hence $(\tan y)/y < 2$ up to the point where $\tan y = 2y$, which occurs (by computation) at approximately $y = 1.1656$. So (a) will hold for $0 < x < 1$ if $\beta < \beta_0 = 1.1656$.

Now consider requirement (b). We need $\alpha \sin^2 \beta x \neq x$ for $0 < x < 1$, where $\beta < \beta_0$; that is, we need $\alpha(\sin^2 \beta x)/x \neq 1$. Since $\alpha(\sin^2 \beta x)/x$ increases (by the choice of β), the largest value of this function on $[0, 1]$ is attained at $x = 1$, where it has the value $\alpha(\sin^2 \beta)/\beta < \alpha(\sin^2 \beta_0)/\beta_0 = 0.7246$. Hence $\alpha \sin^2 \beta x \neq x$ provided that $\alpha < 1/(0.7246) = 1.380 = \alpha_0$, say. Thus both (a) and (b) hold with our present restrictions on α and β.

Finally, $f(1) = \alpha \sin^2 \beta$. Now $\beta < \beta_0$ and $\sin^2 \beta_0 = 0.8446$, and $\alpha_0 \sin^2 \beta_0 > 1.165 > 1$. Hence if we choose β near β_0 and α near α_0 we shall have (c) satisfied.

We can therefore use $f(x) = \alpha \sin^2 \beta x$ in our theorem provided that $\alpha < \alpha_0$, $\beta < \beta_0$, but $\alpha \sin^2 \beta > 1$. If we take, for example, $g(x) = x^2$, we then get $\alpha \sin^2(\beta x^2) \leqslant \alpha^2 \sin^4(\beta x)$, that is,

$$\sin(\beta x^2) \leqslant \alpha^{1/2} \sin^2(\beta x), \qquad 0 \leqslant x < 1,$$

which is something like (9) but with the inequality in the opposite direction. For a specific example, let $\beta = 9/8$, $\alpha = 4/3$; then

$$\sin(9x^2/8) \leqslant (4/3)^{1/2} \sin^2(9x/8), \qquad 0 \leqslant x < 1.$$

D. Let

$$f(x) = (2/\pi)\tan(\pi x/2), \qquad g(x) = x^a, \quad a > 1.$$

(The factor $\pi/2$ is put in to give f the domain $[0, 1)$.) If we replace x by $2y/\pi$ we obtain

$$\tan\{(2/\pi)^{a-1} y^a\} \leqslant (2/\pi)^{a-1} (\tan y)^a,$$

and in particular

$$\tan(2y^2/\pi) \leqslant (2/\pi)\tan^2 y, \qquad 0 \leqslant y < \pi/2.$$

The next example, like the one we started with, has a g that is not a power.

E. Let

$$f(x) = -\log(1 - x), \qquad g(x) = (e^x - 1)/(e - 1).$$

We get

$$\log \frac{e-1}{e-e^x} \leqslant \frac{x}{(1-x)(e-1)}, \qquad 0 \leqslant x < 1.$$

Any number of inequalities of more or less interesting appearance can be written down in the same way.

Annotated Bibliography

The following books constitute a veritable museum of inequalities; you can use them either to see larger collections or to find examples of how inequalities are used in various parts of mathematics.

[1] N. D. Kazarinoff, Analytic inequalities, Holt, Rinehart and Winston, New York, 1961.
Here "analytic" has its old sense of "involving formulas; not geometric." An easy introduction to the subject.

[2] E. F. Beckenbach and R. Bellman, An introduction to inequalities, New Mathematical Library, vol. 3, 1961.
Another easy introduction.

[3] G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities, Cambridge University Press, 1934.
This is the classical reference; no longer quite up to date, but worth reading, if only for its style. The 1952 edition is not very different. The Russian edition (1948) contains much new material in its appendices; some of these were translated in Amer. Math. Soc. Transl. (2) 14 (1960), 1–29.

In 1929 Hardy wrote, "I think that it was Harald Bohr who remarked to me that 'all analysts spend half their time hunting through the literature for inequalities which they want to use and cannot prove'". Ten years later someone asked Hardy whether the book had improved the situation. His reply was that he never seemed able to find in it exactly what he wanted.

[4] E. F. Beckenbach and R. Bellman, Inequalities, Springer-Verlag, New York, 1965.
Some overlap with [3] but contains many more modern topics and methods, and more applications.

[5] D. S. Mitrinović, Analytic inequalities, Springer-Verlag, New York–Heidelberg–Berlin, 1970.
An encyclopedic reference book containing some material not easily accessible elsewhere; perhaps unnecessarily comprehensive. Readers of Serbo-Croatian may prefer the edition in that language (Analitičke Nejednakosti, Belgrade, 1970).

The Distribution of Primes in a Special Ring of Integers

RUFUS ISAACS
Johns Hopkins University
Baltimore, MD 21218

There is a simple binary operation on the nonnegative integers familiar to all who know a common strategy for the game of Nim. Our sum may be obtained by adding in binary notation with no carrying. More precisely, to add a and b, write them above each other in binary notation; addition then yields 0 in any column where a and b have the same digit and 1 in any column where the two summands have different digits. We call this the dot sum, $a \dot{+} b$. Thus $(2 \dot{+} 3)_{10} = (10)_2 \dot{+} (11)_2 = (1)_2 = (1)_{10}$. Berge in [1] explains how this operation fits general games of Nim when it is applied to what he has named, in honor of its discoverer, the Grundy function, and this is the source of our ideas. Whether the algebraic development which follows has further relevance to Nim is an open question.

The nonnegative integers with dot sum form an Abelian group in which each element is its own inverse. An innovation analogous to dot sum for multiplication makes the system distributive. Thus the two new operations make the set of nonnegative integers into a ring G. This ring turns out to be Euclidean so that the nonnegative integers are uniquely factorable (in G) into primes.

We have here a new number-theoretic arena and yield to the temptation to examine the corresponding versions of the classical tantalizers. A list of the new (dot) primes resembles superficially the familiar but irregular classical array. However, a probe of the allocation of primes in G finds them less recalcitrant than in the traditional case. We give a more or less complete answer—and a

fairly sharp one—to the distribution question in Theorem 1. Unexpectedly, a slight detour in the proof leads to one of the Gaussian generalizations of Fermat's theorem.

Our first task is to develop our new dot multiplication: apply the elementary school scheme to the numbers written in binary, but do the final addition dotwise. For example, $5 \cdot 11 = 39$ since

$$\begin{array}{r} 1011 \\ 101 \\ \hline 1011 \\ 10110 \\ \hline 100111. \end{array}$$

Formally, if a and b are written as sums of distinct powers of 2 (tantamount to their binary representations), i.e., if $a = \Sigma_i 2^{\alpha_i}$ and $b = \Sigma_j 2^{\beta_j}$, we define $a \cdot b = \Sigma_{i,j} 2^{\alpha_i + \beta_j}$ where the summation is carried out using our new sum operation. It is clear that this operation is associative and it is not hard to verify distributivity. In short, $G(\dot{+}, \cdot)$ is a ring.

Division in G likewise admits of an elementary format: it is ordinary long division where the internal subtractions are replaced by dot addition. (Recall that in G each element is its own inverse.) For example $50 = 7 \cdot 11 \dot{+} 3$, since

$$\begin{array}{r} 1011 \\ 111\overline{)110010} \\ 111 \\ \hline 101 \\ 111 \\ \hline 100 \\ 111 \\ \hline 11. \end{array}$$

Here, at each intermediate division, the result (digit of the quotient) is 0 or 1 according as the number of digits in the divisor exceeds or equals that of the current remainder.

The reader may easily probe other amusing aspects of our ring. A number is dot divisible by 3 if and only if there are an even number of 1's among its binary digits. A dot square is obtained by inserting a 0 between each consecutive pair of binary digits. The ring G luxuriates in Pythagorean triples.

Readers versed in Galois fields (see [2], for example) will perceive that G is not utterly novel. Let R be the ring of polynomials in x with coefficients the integers reckoned modulo 2. R and G can be seen to be isomorphic, with $\Sigma \epsilon_i 2^i$ corresponding to $\Sigma \epsilon_i x^i$ where each $\epsilon_i = 0, 1$. Thus all our results can be expressed in the language of R, which is well trod algebra, and so doubtless some are intrinsically not new. But the viewpoint matters: G gives us a striking analogue to common number theory unlikely to be noticed in the realm of R.

We start by showing that G is a Euclidean ring. Then each of its elements will have a unique dot factorization into primes [2]. Define $I_k = \{a : 2^k \leqslant a < 2^{k+1}\}$ for $k = 0, 1, 2, \ldots$. If $a \in I_k$, let us denote by $g(a)$ the number of binary digits in a, less 1. From the longhand scheme of multiplication, it is clear that if neither a nor b is 0, $g(a \cdot b) = g(a) + g(b)$. As $g(a) \geqslant 0$ (in fact $g(1) = 0$ and $g(a) > 0$ otherwise), $g(a \cdot b) \geqslant g(b)$. From the long division format we also see that q and r exist uniquely such that $b = a \cdot q \dot{+} r$ with either $r = 0$ or $g(r) < g(a)$. Thus G is Euclidean.

Let us hunt primes. Those up to 128, i.e., with $g \leqslant 6$, are

$$2, 3, 7, 11, 13, 19, 25, 31, 37, 41, 47, 55$$
$$59, 61, 67, 73, 87, 91, 97, 103, 109, 115, 117.$$

We are struck by what appears to be the same inscrutable irregularity that has fascinated mankind with the usual primes. We note familiar absentees and unexpected interlopers. But, as we will see, the

dot primes are much more easily counted than the long struggle with the classic prime number theorem would suggest.

We observe that dot multiplication by a power of 2 is the same as in ordinary multiplication. In fact both multiplications simply adjoin zeros to the right end of the binary representation. Thus in our factorization studies we can confine ourselves to odd integers on the grounds that factors of the form 2^n offer no novelty. (See TABLE 1 for some factorizations and also the Multiplication TABLE 4.) To count the dot primes, we use the inclusion-exclusion principle involving the Möbius function μ. It is defined by $\mu(1)=1$, $\mu(p_1p_2\cdots p_m)=(-1)^m$ where the p_i are distinct primes (in the classical sense), $\mu(c)=0$ if c contains a duplicated prime factor.

THEOREM 1. *Let P_k be the number of dot primes in I_k, then*

$$P_k=\frac{1}{k}\sum_{d|k}2^d\mu\left(\frac{k}{d}\right) \tag{1}$$

where μ is the Möbius function and the sum extends over all positive divisors of k.

Thus, for example, the number of dot primes in I_{10}, the interval [1024,2048), is $P_{10}=(1/10)[2^{10}-2^5-2^2+2^1]=99$. The theorem also guarantees that the right side of (1) is an integer. Is such true for numbers other than 2? A famous result of Fermat states that a prime p always divides a^p-a. Gauss generalized this result by showing that any natural number n divides

$$\sum_{d|n}a^d\mu\left(\frac{n}{d}\right). \tag{2}$$

(Notice that, for n a prime p, (2) is Fermat's a^p-a.) A detour (given at the end) from our later proof of (1) yields as a by-product a proof of the Gaussian generalization. Before doing this, however, we shall use Theorem 1 to develop a rough comparison of the classical and dot prime distributions.

Rewrite (1) as

$$P_k=\frac{1}{k}\left[2^k-\sum_p 2^{k/p}+\sum_{p,p'}2^{k/pp'}-\cdots\right],$$

where the first summation extends over all (classical) prime divisors of k, the second over pairs of distinct prime divisors, etc. Generally, the first term in the bracket will be by far the largest—the second being at most its square root—and we may write $P_k\cong 2^k/k$. But 2^k is the number of numbers in I_k and so we see that the density of dot primes in I_k is roughly $1/k$. Each n in G is contained in some I_k where k is found (roughly) by $2^k=n$ or $k=\log n/\log 2$. Thus the dot prime density near n is roughly $\log 2/\log n$.

The renowned prime number theorem states that $\pi(x)$, the number of primes not exceeding x, approaches $x/\log x$ as $x\to\infty$. That is, the density of primes up to n approaches $1/\log n$ with growing n. Thus, for large n, the ratio of the densities of the dot and classical primes should be $\log 2=.693\ldots$. TABLE 2 offers a reasonable confirmation where π_k denotes the number of classical primes in I_k.

3=11	$17=3^{4\cdot}$
5=3·3	**19=10011**
7=111	$21=7^{2\cdot}$
9=3·7	23=3·13
11=1011	**25=11001**
13=1101	$27=3^{2\cdot}\cdot 7$
15=3·3·3 or $3^{3\cdot}$	29=3·11
	31=11111.

Some prime factorizations: primes are printed bold face and, in place of factorization, we give their binary representations.

TABLE 1.

k	P_k	π_k	P_k/π_k	% Error from log2	Extremes of I_k
9	56	75	.742	7.7	512–1024
10	99	137	.723	4.3	1024–2048
11	186	255	.729	5.2	2048–4096
12	335	464	.722	4.2	4096–8192
15	2182	3030	.720	3.9	32,768–65,536
16	4080	5709	.715	3.1	65,536–131,072
17	7710	10749	.717	3.5	131,072–262,144
18	14532	20390	.713	2.8	262,144–524,288

A comparison of the number of dot primes P_k and the number of classical primes π_k in the interval $I_k=[2^k,2^{k+1})$. The higher π_k are from a table made by D. V. Vandelinde.

TABLE 2.

TABLE 3.

A Table of Dot Addition

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	+
0	3	2	5	4	7	6	9	8	11	10	13	12	15	14	17	16	19	18	21	1
	0	1	6	7	4	5	10	11	8	9	14	15	12	13	18	19	16	17	22	2
		0	7	6	5	4	11	10	9	8	15	14	13	12	19	18	17	16	23	3
			0	1	2	3	12	13	14	15	8	9	10	11	20	21	22	23	16	4
				0	3	2	13	12	15	14	9	8	11	10	21	20	23	22	17	5
					0	1	14	15	12	13	10	11	8	9	22	23	20	21	18	6
						0	15	14	13	12	11	10	9	8	23	22	21	20	19	7
							0	1	2	3	4	5	6	7	24	25	26	27	28	8
3		5						0	3	2	5	4	7	6	25	24	27	26	29	9
5		15	17						0	1	6	7	4	5	26	27	24	25	30	10
7		9	27	21						0	7	6	5	4	27	26	25	24	31	11
9		27	45	63	65						0	1	2	3	28	29	30	31	24	12
11		29	39	49	63	69						0	3	2	29	28	31	30	25	13
13		23	57	35	101	127	81						0	1	30	31	28	29	26	14
15		17	51	45	119	105	75	85						0	31	30	29	28	27	15
17		51	85	119	153	187	221	255	257						0	1	2	3	4	16
19		53	95	121	139	173	199	225	291	261						0	3	2	5	17
21		63	65	107	189	151	233	155	325	367	273						0	1	6	18
																		0	7	19
		3	5	7	9	11	13	15	17	19	21								0	20

A Table of Dot Multiplication

TABLE 4

These two tables exhibit some of the interesting patterns found in the operation of dot addition and multiplication. In the multiplication table only odd numbers are included because dot and common multiplication by 2 are the same. Thus to use the multiplication table to find $10 \cdot 28$ we need apply the table only to the product in parenthesis: $10 \cdot 12 = 2 \times 4 \times (5 \cdot 7) = 8 \times 27 = 216$.

There is a subclass of dot primes with no apparent classical analogue. The binary representation of an odd integer a begins and ends with 1. Let a' denote the integer a with the digit order reversed. We call a **palindromic** if $a = a'$. A little study of the longhand dot multiplication scheme makes it clear that $(a \cdot b)' = a' \cdot b'$. It follows readily that if a is prime, so is a'. Thus, all dot primes are either palindromic or occur in mirror-image pairs. From TABLE 1 we see that 11 and 13, 19 and 25 are such pairs, while the lowest palindromic dot primes are 3, 7, 31, 73, 313.

Just as for primes, we may give a precise formula for the number of palindromic dot primes in I_k:

THEOREM 2. *Let S_k be the number of palindromic dot primes in I_k.*

$$S_k = \begin{cases} 1 & \text{if } k = 1 \\ 0 & \text{if } k \text{ is odd and } k > 1, \text{ and} \\ \dfrac{1}{k} \displaystyle\sum_{d \mid c} \mu\left(\frac{c}{d}\right) 2^{2^{h-1}d} & \text{if } k = 2^h c \text{ with } h > 0 \text{ and } c \text{ odd.} \end{cases}$$

While it is possible to give a proof of Theorem 2 similar to that given below for Theorem 1, the two lemmas below furnish a more direct and less tedious approach. The proofs we give are skeletal; the reader may find adding the flesh enjoyable and instructive.

Let us call the primes relative to the palindromic numbers relative palindromic dot primes or RPP for short. Thus "n is an RPP" means: n has no palindromic dot divisors other than itself or 1. Let R_k be the number of RPP in I_k.

LEMMA 1. *For odd k, $R_k = 0$; for even k, R_k is the number of odd dot primes in $I_{k/2}$.*

Proof. A palindromic integer is divisible by 3 unless it has an odd number of binary digits with the central one being 1. The number of such primes in I_{2m} is the same as the number of odd integers in

I_m. Each of the latter has a canonical dot factorization $\Pi_i p_i^{s_i}$ with $\Sigma_i s_i g(p_i)=m$. We can show this inductively on m isomorphic factorizations for the former with RPP replacing the p_i.

LEMMA 2. *For an even k, put $k=2^h c$ with $h>0$ and c odd. Then*

$$S_k=\sum_{j=0}^{h-1}\frac{P_{2^j c}}{2^{h-j}}.$$

Proof. Of the RPP in I_k, some are dot primes and the rest have nonpalindromic divisors. The latter can only consist of one pair, p and p', with $g(p)=k/2$. Thus $R_k=S_k+\frac{1}{2}(P_{k/2}-S_{k/2})$. We now substitute repeatedly, using Lemma 1.

Theorem 2 now follows by using Theorem 1 to eliminate the P_i. These same techniques can be used to show that the density of palindromic dot primes near the integer n is approximately $(\frac{2}{3}\log 2)/\sqrt{n}\ \log n$.

There is abundant territory still to be explored. First, the classical analogs. Does a dot Fermat theorem hold? In other words, does a dot prime p always divide $a^{\dot{p}}\dot{+}a$? What about a dot Wilson Theorem? Or quadratic forms, or perfect numbers, or twin dot primes?

Further, we might replace the base 2 by another integer. If our new base is prime, we can expect success, but is this question not rather pedestrian? However, the familiar base 10 is curious: commonplace arithmetic with no carrying. We get a ring but not a Euclidean one. The reader may probe the fanciful way that integers here factor; the algebra teacher may find in it a simple example to show that factorization can be far from unique in general rings.

The remainder of this note consists of a proof of Theorem 1 with a subsidiary proof of the Gauss-Fermat theorem. If we decompose an integer $a\in I_n$ (that is, with $g(a)=n$) into dot prime factors, the fact that $g(a\cdot b)=g(a)+g(b)$ limits us to the following type of outcome: if there are x_j prime factors for which g is equal to j, then we must have

$$\sum_{j=1}^{n} jx_j=n. \tag{3}$$

Let us denote by X a particular set of x_j satisfying (3). We inquire as to how many factorizations of various members of I_n pertain to this X.

For each $j(j=1,2,\ldots,n)$ we can choose from among P_j possible dot primes. Let us denote the number of possible products with x_j such primes by $L(P_j,x_j)$. (In other words, $L(P_j,x_j)$ is the answer to the following combinatorial problem. If we have unlimited stocks of each of P_j kinds of fruit, how many ways are there of filling a fruit basket with x_j pieces?) By standard reasoning we ascertain

$$L(P_j,x_j)=\frac{P_j(P_j+1)\cdots(P_j+x_j-1)}{x_j!}=(-1)^{x_j}\binom{-P_j}{x_j}. \tag{4}$$

Returning to our main reasoning, the number of factorizations with a given X is thus the product $L(P_1,x_1)L(P_2,x_2)\cdots L(P_n,x_n)$ and the total of *all* factorizations is this quantity summed over all possible sets X. But, because of unique decomposition into dot primes, this total is also the number of members of I_n which is 2^n. Thus

$$\sum_{X}\prod_{x_j\in X} L(P_j,x_j)=2^n. \tag{5}$$

Let us now consider the formal power series

$$S_j(u)=1+L(P_j,1)u^j+L(P_j,2)u^{2j}+\cdots=\sum_{i=0}^{\infty}L(P_j,i)u^{ij}$$

where $j=1,2,\ldots k$ and $L(p,0)$ is taken as 1. Since we are going to manipulate such series formally, only as a convenient device to obtain relations among the coefficients, convergence need not concern

us. What is the coefficient of u^n when all these series are formally multiplied together? If this coefficient arises through the use of x_jth terms in the jth series, it must be that the former terms satisfy $\Sigma_j jx_j = n$, which is (3). Thus, each set X contributes a product of $L(P_j, x_j)$ to the coefficient of u^n and its value is the sum of such for all possible X. In other words, the coefficient is the left side of (5). We conclude that

$$\prod_j S_j(u) = \sum_n 2^n u^n. \tag{6}$$

Now, from (4), $S_j(u)$ is $(1-u^j)^{-P_j}$, while the series on the right of (6) is $(1-2u)^{-1}$. Thus

$$(1-u)^{P_1}(1-u^2)^{P_2}(1-u^3)^{P_3}\cdots = 1-2u. \tag{7}$$

Take (formal) logarithms of both sides:

$$\sum_j P_j \sum_{i=1}^{\infty} \frac{u^{ij}}{i} = \sum_{v=1}^{\infty} \frac{2^v u^v}{v}.$$

Note that u^v occurs on the left each time that i and j are a pair of factors of v. Thus, equating coefficients of like powers leads to $\Sigma_{d|v} dP_d = 2^v$. The Mobius inversion theorem now avers that

$$kP_k = \sum_{d|k} \mu\left(\frac{k}{d}\right) 2^d, \tag{8}$$

our sought result.

If we return to (7) but with the 2 on the right replaced by an arbitrary positive integer a, we observe that if the P_j can be determined *integrally* so that the equation formally balances, this would then prove the Gauss-Fermat Theorem (2). If the binomials are multiplied out, then together, and coefficients of like powers equated, each P_j will appear for the first time in an equation with the following properties: P_j will appear once, lineally with coefficient ± 1. The other terms will be products of binomial coefficients of the type $\binom{P_k}{k}$ (and a when $j=1$). Thus the P_j are successively and uniquely determined as integers. We then reason as before, arriving at (8) with a replacing 2 (and, of course, new P_k).

References

[1] Claude Berge, The Theory of Graphs and Its Applications, Methuen, 1962.

[2] B. L. Van der Waerden, Modern Algebra, Ungar, 1949.

Iterated Absolute Differences

PETER ZVENGROWSKI

University of Calgary

Calgary, Canada T2N 1N4

If one starts with any sequence s of four non-negative integers (a_0, a_1, a_2, a_3), forms the "derived" sequence $Ds = (|a_0 - a_1|, |a_1 - a_2|, |a_2 - a_3|, |a_3 - a_0|)$, and iterates this process, then ultimately the sequence $\theta = (0,0,0,0)$ is obtained. Let us illustrate this curious process with a few examples:

(1,9,8,0)	(1,9,7,9)	(90, 1,174,123)
(8,1,8,1)	(8,2,2,8)	(89,173, 51, 33)
(7,7,7,7)	(6,0,6,0)	(84,122, 18, 56)
(0,0,0,0)	(6,6,6,6)	(38,104, 38, 28)
	(0,0,0,0)	(66, 66, 10, 10)
		(0, 56, 0, 56)
		(56, 56, 56, 56)
		(0, 0, 0, 0)

A proof of this result is given by Honsberger [5], who attributes the problem to an Italian mathematician named E. Ducci. A generalization of this result, stated as the theorem below, is given by Ciamberlini and Marengoni [3]. In this note we shall give a short proof of this generalization which takes advantage of some basic properties of polynomial rings, and mention one or two related questions.

Let $S_r=\{(a_0,a_1,\ldots,a_{r-1}): a_i\in Z, a_i\geqslant 0\}$, and let $D: S_r\to S_r$, where

$$D(a_0,a_1,\ldots,a_{r-1})=(|a_0-a_1|,|a_1-a_2|,\ldots,|a_{r-1}-a_0|).$$

If $s\in S_r$, we say s **converges to zero** if $D^n s=\theta$ for some n. Finally, for any $q\geqslant 0$ let $q\cdot s=(qa_0,\ldots,qa_{r-1})$ and let $\max s=\max\{a_0,a_1,\ldots,a_{r-1}\}$.

THEOREM. *All sequences $s\in S_r$ converge to zero if and only if r is a power of 2.*

Proof. First consider the theorem "modulo 2". That is, suppose $a_0,\ldots,a_{r-1}\in Z_2$. In this context $Ds=(a_0+a_1,\ldots,a_{r-1}+a_0)$ where $s=(a_0,\ldots,a_{r-1})$. Now consider the polynomial $p_s(t)=a_0t^{r-1}+a_1t^{r-2}+\cdots+a_{r-1}$ with coefficients in Z_2. The ring $R=Z_2[t]/(t^r+1)$ can be thought of as the polynomial ring $Z_2[t]$ with t^r identified with 1. But p_s has a special property in R, namely, $p_{Ds}(t)=(1+t)p_s$; this may be seen by direct calculation (remembering to replace t^r by 1). So by iteration, $p_{D^n s}(t)=(1+t)^n p_s$.

Two observations allow us to study $D^r s \pmod 2$. First $s\equiv\theta \pmod 2$ if and only if p_s is the zero polynomial in R. Secondly, $(1+t)^{2^m}=1+t^{2^m}$ in R since all the intermediate binomial coefficients vanish (mod 2) (see [4]). So if $r=2^m$,

$$p_{D^r s}(t)=(1+t)^{2^m}p_s(t)=(1+t^{2^m})p_s(t)=(1+t^r)p_s(t)=0$$

in R. This shows that $D^{2^m}s\equiv\theta \pmod 2$.

If $D^r s\equiv\theta \pmod 2$, $D^r s=2\cdot s'$, for some $s'\in S_r$, $D^{2r}s=2\cdot D^r s'=4s''$ for some $s''\in S_r$, and so on. Thus all the terms in $D^n s$ become divisible by arbitrarily high powers of 2 (with increasing n). But since the terms of $D^n s$ are bounded by $\max s$, for sufficiently large n the terms of $D^n s$ are in fact all zero.

Suppose now that r is not a power of 2. Choose $s=(0,0,0,\ldots,0,1)$. Then $p_s(t)=1$ and $p_{D^n s}(t)=(1+t)^n$. Now $(1+t)^n$ can be zero in R if and only if $(1+t)^n$ is divisible by t^r+1 in $Z_2[t]$. But $Z_2[t]$ is a unique factorization domain [2]. Since $1+t$ is prime in $Z_2[t]$, this means $t^r+1=(1+t)^r$, which is impossible since $r\neq 2^m$ (see [4]). Hence, for $r\neq 2^m$, not all sequences converge to zero.

This theorem clearly generalizes to non-negative rational numbers. It seems interesting to ask whether a uniform bound on the number of derivations necessary for zero-convergence exists (for a fixed $r=2^m$). For sequences of length 4, a simple construction is given in [3] which shows the existence of sequences in S_4 requiring arbitrarily many derivations for zero-convergence. This is also shown in [1]. However, from the considerations in the last part of the proof, one readily sees that $r([\log_2(\max s)]+1)$ is a bound valid for each s individually.

References

[1] E. R. Berlekamp, Design of slowly shrinking squares, Math. Comp., 29 (1975) 25–27.

[2] G. Birkhoff and S. MacLane, A Survey of Modern Algebra (rev. ed.), Macmillan, New York, 1953, p. 76.

[3] C. Ciamberlini and A. Marengoni, Su una interessante curiosità numerica, Period. Mat. Ser. 4, 17 (1937) 25–30.

[4] N. J. Fine, Binomial coefficients modulo a prime, Amer. Math. Monthly, 54 (1947) 589–592.

[5] R. Honsberger, Ingenuity in Mathematics, Random House, New York, 1970.

Estimating a Population Proportion Using Randomized Responses

JAY L. DEVORE
California Polytechnic State University
San Luis Obispo, CA 93407

A standard problem in the statistician's repertoire is that of estimating a population proportion p. The usual approach to this problem is to let X represent the number of individuals in a sample of size n who possess the characteristic which defines p, assume that the individual sample responses comprise a binomial experiment (n independent homogeneous trials, each with the same dichotomous responses) so that X is a binomial random variable, and use the estimator $\hat{p} = X/n$. When the underlying assumptions are satisfied, the estimator $\hat{p}$ is both the minimum variance unbiased estimator and the maximum likelihood estimator of p.

The reasonableness of this estimator, as well as its optimality properties, depend crucially on the presumption that individual responses are truthful. If, for example, we are interested in estimating the true proportion of all presently enrolled college students who have seen the movie "Star Wars," then there is no reason to suppose that a randomly chosen student will respond untruthfully when asked, "Have you seen "Star Wars"?" There are, however, many situations in which one or perhaps both responses may embarrass or stigmatize the respondent. Examples include questions regarding drug use, sexual experience, honor code violations, and abortions. Even if a randomly selected student has violated the honor code at his or her college, it is not clear that the response to "Have you ever violated the college honor code?" will be "yes." In such a situation, the estimator $\hat{p}$ will probably have a bias of unknown, and thus uncorrectable, magnitude (and even direction, if both a "yes" or "no" response can stigmatize.)

In recent years a method called the randomized response technique has been introduced to circumvent the problem of untruthful responses. To illustrate this technique, suppose that an investigator has identified a random sample of size n from the population of interest. The investigator has in hand a deck of 100 cards, of which 50 are type I cards and the other 50 are type II cards. The type I cards ask for a (truthful) response to the question of interest (e.g., "Have you violated the honor code?"), while the other 50 ask for a response to an "unrelated" or "irrelevant" question such as "Is the fifth digit of your home telephone number a 0, 1, or 2?" A respondent is asked to examine the deck to confirm the stated mix of type I and type II cards, then to shuffle the deck until it is well-mixed, select a card and respond truthfully to the question on the card, and lastly reinsert the card into the deck without having revealed to the investigator which type of card was selected.

What prompts a truthful response in this scheme? The key point is that the questioner does not know whether the respondent has answered the question of interest or the irrelevant question, so that the respondent will not be stigmatized by a truthful response. The model presumes that the proportion of type II cards in the deck is high enough to offer enough protection against stigma so that truthful responses are forthcoming. In practice the tendency for an untruthful response may not be completely eliminated by the technique, but should certainly be drastically reduced. Goodstadt and Gruson ([5]) report on the successful use of the technique in a survey on drug use.

Presuming that responses using the randomized technique are truthful, an estimator for p can be derived by defining Y as the number of yes responses in the sample using the randomized scheme, and λ as the probability of a yes response using the randomized scheme. Then a simple probability argument yields

$$\begin{aligned}\lambda &= P\{\text{yes}|\text{type I card}\}P\{\text{type I card}\} + P\{\text{yes}|\text{type II card}\}P\{\text{type II card}\} \\ &= .5p + .15. \end{aligned} \tag{1}$$

The final expression in (1) depends on knowing the probability of a yes response to the irrelevant question; we assume here that the fifth digit of a telephone number is uniformly distributed among the ten possibilities, yielding $P\{\text{yes}|\text{type II card}\}=.3$. Solving (1) for p as a function of λ yields $p=2\lambda-.3$. Since λ can be estimated by $\hat{\lambda}=Y/n$, substituting this estimator in place of λ gives $\hat{p}=2\hat{\lambda}-.3$ as an estimator for p.

The usefulness of the technique, ease of exposition, and use of the elementary probability calculus in the derivation has led to the appearance of this material in expository and introductory sources (e.g., [1], [3], [8]), as well as to refinements and extensions of the technique in the research literature.

In general, the basic equation for the irrelevant question model from which an estimator is derived is

$$\lambda=\Pi p+(1-\Pi)\theta \tag{2}$$

where λ and p are as defined earlier, Π equals the probability of responding to the question of interest (a type I card), and θ equals the probability of a yes response to the irrelevant question. Solving this equation for p gives

$$p=\frac{\lambda-(1-\Pi)\theta}{\Pi}. \tag{3}$$

The estimator for p is now obtained by substituting $\hat{\lambda}=Y/n$ into (3) in place of λ.

One problem with the use of the irrelevant question model is that even with the randomization, it is still true for many questions of interest (e.g. cheating, abortions) that only a yes response can stigmatize. There is another model, which was actually introduced prior to the irrelevant question model, which is suitable when both responses might possibly stigmatize. In this earlier model, the irrelevant question on the type II card is replaced by the negation of the question of interest appearing on the type I card. The relationship between λ and p is then obtained by replacing θ in (2) by $1-p$. Provided that $\Pi\neq\frac{1}{2}$(i.e. the deck is not evenly divided between type I and type II cards), p can be estimated by using Y/n to estimate λ. If $\Pi=\frac{1}{2}$, then the probability of a yes response equals $\frac{1}{2}$ for all values of p; thus the probability distribution of Y does not depend on p, so p cannot be estimated.

Both the unrelated question estimator and the negation estimator are linear functions of the binomial variable Y, which makes investigation of their properties straightforward. In particular, since $\hat{\lambda}$ is unbiased for λ, both estimators are unbiased for p. Using the subscripts "un" to denote the unrelated question estimator and "neg" to denote the negation estimator, the variances are

$$\begin{aligned}\operatorname{Var}(\hat{p}_{\text{un}})&=[\Pi p+(1-\Pi)\theta][1-\Pi p-(1-\Pi)\theta]/n\Pi^2\\ \operatorname{Var}(\hat{p}_{\text{neg}})&=\left[\Pi p+(1-\Pi)(1-p)\right]\left[1-\Pi p-(1-\Pi)(1-p)\right]/n\Pi^2.\end{aligned} \tag{4}$$

For most reasonable parameter values, the unrelated question estimator has smaller variance than the negation estimator ([1]), so that the former is now much more widely used than the latter. Using either the unrelated question model or the negation model, it follows from standard statistical theory that Y is a complete sufficient statistic for λ (and thus p), so that both estimators have minimum variance among all unbiased estimators ([6], chapter 7). There is, though, a serious problem with both estimators: it is possible for an estimate to be either negative or greater than one. A specific example of this is obtained by taking the original values $\Pi=.5$, and $\theta=.3$ in (3). Then if $\hat{\lambda}=.1$, $\hat{p}=-.1$, while if $\hat{\lambda}=.8$, $\hat{p}=1.3$.

From a practical viewpoint, the use of an estimator yielding estimates which are not possible parameter values should certainly be avoided. Furthermore, this property of $\hat{p}$ directly contradicts the claim made in several recent papers in widely read journals ([1], [4], [7]), including one co-authored by the originator of the method (Warner), that the estimators are maximum likelihood. For the definition of a maximum likelihood estimator guarantees that any such estimate will be a possible value of the parameter. Since the parameters λ and p are linearly related, the invariance property of maximum likelihood estimators ([2], p. 291) implies that if the maximum likelihood estimator for λ is substituted into (3), the resulting estimator for p will be a maximum likelihood estimator. This in turn implies that

$\hat{\lambda} = Y/n$ cannot be the maximum likelihood estimator for λ, i.e., the sample proportion of yes responses is not the maximum likelihood estimator for the probability of a yes response. To see why this is so, consider again equation (1). As p varies between 0 and 1, λ varies only between .15 and .65, so that $\hat{\lambda} = .1$ is not a maximum likelihood estimate because .1 is not a possible value for λ.

Since Y is a binomial variable, the maximum likelihood estimator is obtained by finding

$$\sup_{.15 \leqslant \lambda \leqslant .65} \binom{n}{y} \lambda^y (1-\lambda)^{n-y} \tag{5}$$

with the maximum likelihood estimator being the value of λ for which the supremum is achieved. Denoting the maximum likelihood estimator by $\tilde{\lambda}$, it easily follows that

$$\tilde{\lambda} = \begin{cases} .15 & \text{if } Y/n < .15 \\ Y/n & \text{if } .15 \leqslant Y/n \leqslant .65 \\ .65 & \text{if } Y/n > .65. \end{cases} \tag{6}$$

In general $\tilde{\lambda}$ will be Y/n truncated at the upper and lower bounds for λ obtained by varying p in (2). Now substitution of $\tilde{\lambda}$ into (3) gives the maximum likelihood estimator for p, which is simply the unbiased estimator truncated at 0 and 1.

It might be noted that this produces a very nice and realistic example of a minimum variance unbiased estimator which can take on ridiculous values over some parts of the sample space. A standard but artificial textbook example of this phenomenon is to estimate $\exp(-2\lambda)$ based on a single observed value of a Poisson variable X having parameter λ.

If the respondents in the sample of size n can be gathered in one location, then a "without replacement" version of the randomized response scheme can be used to produce an estimator which is both minimum variance unbiased and the maximum likelihood estimator. Suppose that n is even and that the investigator has a deck containing n cards, $n/2$ having the question of interest and the other $n/2$ instructing the respondent to give a yes response. This deck is shown to the respondents, well shuffled, and one card is given to each respondent. Then there will be $n/2$ responses to the question of interest, but the investigator will again not know which responses are to the question of interest and which are automatic yes responses. Letting $X =$ the total number of yes responses from the sample, $X - n/2$ is a binomial variable based on $n/2$ trials with success parameter p, so the maximum likelihood estimator and best unbiased estimator is

$$\hat{p} = \frac{X - n/2}{n/2}, \tag{7}$$

which obviously assumes only values between 0 and 1 inclusive. The variance of this without-replacement estimator is

$$\operatorname{Var}(\hat{p}) = 2p(1-p)/n. \tag{8}$$

Comparison of (8) with the variances of (4) when $\Pi = .5$ (so that for all three models, the question of interest appears on half the cards in the deck) shows that the without-replacement estimator is preferable to the other two estimators. Of course, in many situations it will not be possible to gather all respondents together in one location.

References

[1] Cathy Campbell and Brian Joiner, How to get the answer without being sure you've asked the question, Amer. Statist., 27 (1973) 229–231.

[2] Morris DeGroot, Probability and Statistics, Addison-Wesley, 1975.

[3] Richard Drogin and Michael Orkin, Vital Statistics, McGraw-Hill, 1975.

[4] Sven Eriksson, A new model for randomized response, Int. Stat. Rev., 41 (1973) 101–113.

[5] Michael Goodstadt and Valerie Gruson, The randomized response technique: a test on drug use, J. Amer. Statist. Assoc., 70 (1975) 814–823.

[6] Robert Hogg and Allen Craig, Introduction to Mathematical Statistics, Macmillan, 1970.

[7] Frederick Leysieffer and Stanley Warner, Respondent jeopardy optimal designs in randomized response models, J. Amer. Statist. Assoc., 71 (1976) 649–656.

[8] Gottfried Noether, Introduction to Statistics (2nd ed.), Houghton Mifflin, 1976.

Monochrome Lines in the Plane

JONATHAN M. BORWEIN
Dalhousie University
Halifax, Nova Scotia
Canada B3H 3J5

Recently Tingley [3] has shown that, given any two disjoint, connected, compact sets, A and B, in the plane which do not both lie entirely on one line, one can find a line which passes through at least two points of one of these sets and misses the other. Such a line is said to be **monochrome** and the set it passes through is said to have a monochrome line. Tingley's theorem can be viewed as a variant of Motzkin's result (see [3] for discussion) that any two disjoint finite sets jointly **spanning** R^2 (that is, which do not both lie on the same line) possess a monochrome line. In this note we first show that Tingley's conditions can be considerably weakened. Then we give conditions for uncountably many monochrome lines to exist.

The following notational conventions will be useful: ab denotes the line through points a and b; $[a,b]$ denotes the closed segment between a and b; $P(\bar{b},B)$ denotes the "pencil" consisting of all lines through points $b \in B$ and a fixed $\bar{b} \notin B$; co B denotes the **convex hull** of B, i.e., the smallest convex set containing B; $\overline{B}$ denotes the closure of B; $A - B$ denotes the points of A which are not in B; and if L is a line, L^+ and L^- denote the two closed half planes it generates.

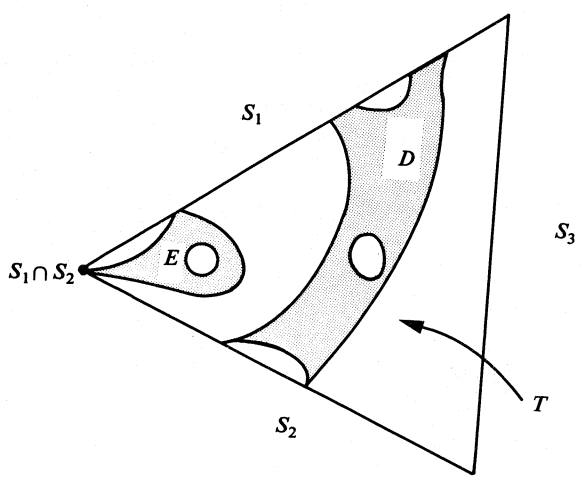


FIGURE 1.

We also take note of the following two results concerning connected sets which will be used frequently. Recall that a set A is **connected** if it cannot be written as the union of two nonempty sets B and C such that $\overline{B} \cap C = \overline{C} \cap B = \varnothing$. In particular, a closed set is connected if it is not the union of two nonempty disjoint closed subsets. Newman's book [2] is an excellent source for results about connectedness.

PROPOSITION 1. *Suppose D and E are closed connected disjoint sets lying in a triangle T with sides s_1, s_2 and s_3. Suppose also that* (i) $s_1 \cap s_2 \in E$, (ii) $s_1 \cap D \neq \emptyset$, (iii) $s_2 \cap D \neq \emptyset$. *Then* $s_3 \cap E = \emptyset$.

This is easily proved (assuming only that either D or E is closed) using the methods described in Newman's book [2, pg. 102]. It also follows from a theorem in [1, pg. 636]. The situation is indicated in FIGURE 1. It is worth pointing out that if neither D nor E is closed, the result is false. This is the main subject of [1].

PROPOSITION 2. *If D is closed and connected and L is a line such that $D \cap L$ is at most a singleton, then both $D \cap L^+$ and $D \cap L^-$ are connected.*

This is a simple exercise in handling connected sets. The situation is shown in FIGURE 2. We are now prepared to show how Tingley's result may be extended to a pair of sets A and B where only one is bounded.

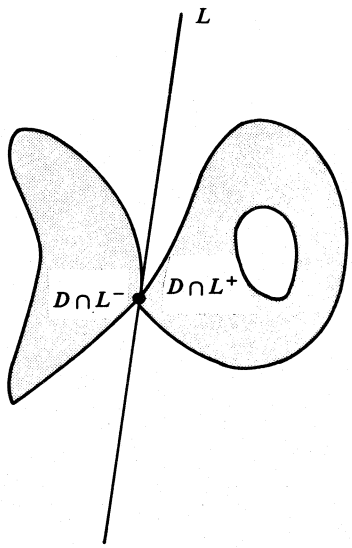


FIGURE 2.

THEOREM 1. *Suppose A and B are two sets in R^2 which jointly span the plane and suppose B is bounded. Suppose further that $\overline{A}$ and $\overline{B}$ are disjoint and connected. Then there exists a monochrome line for A and B.*

Proof. We will proceed through a number of stages. Our first aim is to show that $\overline{A}$ and $\overline{B}$ possess a monochrome line. We may assume that some $\bar{a} \in \overline{A}$ exists with $\bar{a} \notin \text{co}\overline{B}$. If A is unbounded, this is clear, while if both $\overline{A}, \overline{B}$ are compact, either (i) $\overline{A} - \text{co}\overline{B} \neq \emptyset$ or (ii) $\overline{B} - \text{co}\overline{A} \neq \emptyset$. For if both (i) and (ii) were false, this would imply that $\text{co}\overline{A} = \text{co}\overline{B}$. And then two compact sets with identical convex hulls must of necessity intersect, since $\text{co}\overline{A} = \text{co}\overline{B}$ implies that the furthest point from $(0,0)$ in $\text{co}\overline{A}$ must belong to $\overline{A}$ and equally to $\overline{B}$. Thus if $\overline{A} \cap \overline{B} = \emptyset$, one of (i) of (ii) must hold. By symmetry we assume below that (i) holds.

Consider the pencil $P(\bar{a}, A \cup B - \{\bar{a}\})$ of lines through $\bar{a}$. Either a monochrome line exists in A or $P(\bar{a}, \overline{A} \cup \overline{B} - \{\bar{a}\}) = P(\bar{a}, \overline{B})$. We shall shorten $P(\bar{a}, \overline{B})$ to P'. Since $\overline{B}$ is compact there are extreme lines

L_1 and L_2 of P containing points b_1 and b_2 of $\bar{B}$. Since $A \cup B$ spans R^2 and $\bar{A}$ does not contain a monochrome line, L_1 is not equal to L_2. Also since $a \notin \text{co}\bar{B}$, L_1 cannot coincide with L_2 in the opposite direction. Thus $\bar{A} \cup \bar{B}$ lies in two closed cones with vertex $\bar{a}$ and vertex angles properly between 0 and π. Moreover, all the points of $\bar{B}$ must, by Proposition 2, lie in one of these cones; call it P^+. $\bar{A}$ may have points in the other cone which we shall call P^-. This configuration is indicated in FIGURE 3.

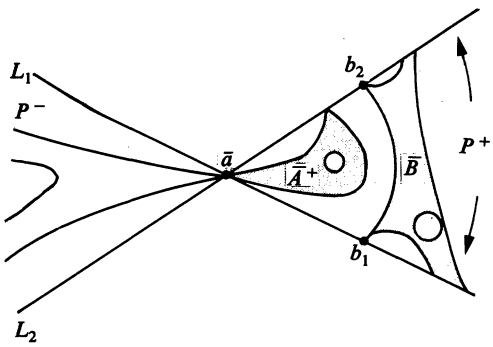


FIGURE 3.

We make the following claims:

(a) $\bar{A}^+ = \bar{A} \cap P^+$ is connected.

(b) $\bar{A}^+$ is bounded, hence compact.

The first claim follows from Proposition 2 applied to $\bar{A}$ and the external bisector of L_1 and L_2. To establish (b), consider translations of the line b_1b_2 away from $\bar{a}$. Since $\bar{B}$ is compact, eventually some such translate N is disjoint from $\bar{B}$ and since we have assumed that $\bar{A}$ has no monochrome line, there is at most one point of $\bar{A}^+$ on N. Let T_1 be the triangle bounded by N, $\bar{a}b_1$, and $\bar{a}b_2$. By Proposition 2, $\bar{A}^+ \cap T_1$ is connected, and by Proposition 1 (with $T = T_1$, $E = \bar{A}^+ \cap I_1$ and $D = \bar{B}$) we find that $N \cap (\bar{A}^+ \cap T_1) = N \cap \bar{A}^+ = \varnothing$. Since $\bar{A}^+$ is connected, $\bar{A}^+$ must lie in T_1 and so is bounded. The situation is indicated in FIGURE 4.

Now either b_1b_2 is monochrome or there is a point of $\bar{A}$ (which must lie in $\bar{A}^+$) on $[b_1, b_2]$. In this latter case there is (by (b)) a line parallel to b_1b_2 meeting $\bar{A}^+$ and such that no line parallel to b_1b_2 and farther from $\bar{a}$ meets $\bar{A}^+$. Let a_1 be a point of $\bar{A}^+$ on this line and let c_1 be a point on $\bar{a}a_1$ beyond a_1 with $[a_2, c_1] \cap \bar{B} = \varnothing$. Note that a_1 may be on b_1b_2 while c_1 may be on L_1 or L_2.

Consider the line M parallel to b_1b_2 and running through c_1. By construction $\bar{a}$ and c_1 are on opposite sides of b_1b_2. Since $M \cap \bar{A} = \varnothing$, either M is monochrome of $M \cap \bar{B}$ is at most a singleton. In this latter case we consider the triangle T_2 whose sides are bounded by $\bar{a}b_1, \bar{a}b_2$ and M. This is shown in FIGURE 5. By Proposition 2, $\bar{B} \cap T_2$ is connected. $\bar{A}^+ \cup [a_1, c_1]$ is also closed and connected. We may now apply Proposition 1 with $T = T_2, D = \bar{B} \cap T_2, E = \bar{A}^+ \cup [a_1, c_1]$ and discover (noting that D and E are disjoint) that $E \cap M = \varnothing$, which is impossible. The only other possibility was that M was monochrome. We have now established that either

(1) $\bar{A}$ has a monochrome line,

or

(2) some line $b_1'b_2'$ parallel to b_1b_2 is monochrome in $\bar{B}$.

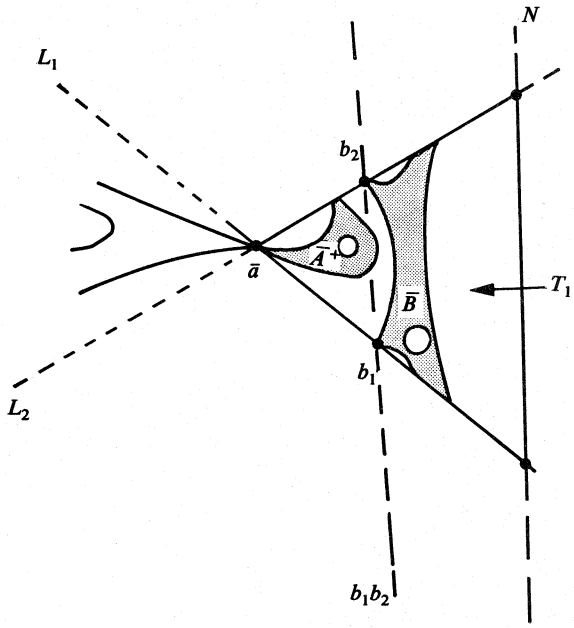


FIGURE 4.

It remains to show that monochrome lines actually exist in A or B. In case (2) we may pick sequences $\{b_{1n}\}$ and $\{b_{2n}\}$ in B with $b_{1n}\to b_1', b_{2n}\to b_2'$. Now for n large enough, $b_{1n}b_{2n}$ will not meet P^- (since $b_1'b_2'$ neither meets P^- nor is parallel to L_1 or L_2). It follows that any sequence $\{a_n\}\subset\bar{A}$ with a_n on $b_{1n}b_{2n}$ must lie in $\bar{A}^+$. Since that set is compact $\{a_n\}$ would contain a subsequence converging to a point of $\bar{A}$ on $b_1'b_2'$. This is impossible and it follows that for n large enough $b_{1n}b_{2n}$ is monochrome.

The first case follows similarly, but now we need not consider P^- because $\bar{B}$ is assumed compact. Thus in either case we have produced a monochrome line in one of A or B.

Looking back on our proof we see that we cannot just apply Tingley's theorem to $\bar{A}^+$ and B because the "monochrome" line so produced may intersect $A\cap P^-$. This is avoided in the proof by considering lines parallel to b_1b_2. Also note that in the last paragraphs of the proof all lines $b_{1n}b_{2n}$ may coincide with $b_1'b_2'$. Theorem 2 (below) is relevant here.

In [3] it is shown that even for A and B compact, Theorem 1 may fail if even one of A or B is disconnected. The construction of pleasant examples of this kind is a most diverting occupation. In Tingley's example, for instance, one of the sets is connected and the other has only two components. It is also shown there that the theorem fails if both are unbounded. However the example of Figure 4 of [3] does not satisfy the conditions of Theorem 1.

If in Theorem 1 we require A and B to be closed and to *separately* span R^2, we can strengthen the conclusion considerably.

THEOREM 2. *If A, B are disjoint closed sets each of which spans R^2, there are uncountably many monochrome lines.*

Proof. Suppose first that in the construction in Theorem 1 case (2) obtains. By hypothesis there is some point $b_3\in B-b_1'b_2'$, whence, by the connectedness of B, every translate L_t of $b_1'b_2'$ between b_3 and $b_1'b_2'$ contains a point b_t of B. (Here L_0 is $b_1'b_2'$, L_1 is the translate through b_3 and L_t tends continuously to L_0 as $t\to 0$.) Since either $\|b_t-b_1'\|\geq\frac{1}{2}\|b_2'-b_1'\|=d$ or $\|b_t-b_2'\|\geq\frac{1}{2}\|b_2'-b_1'\|$ we may

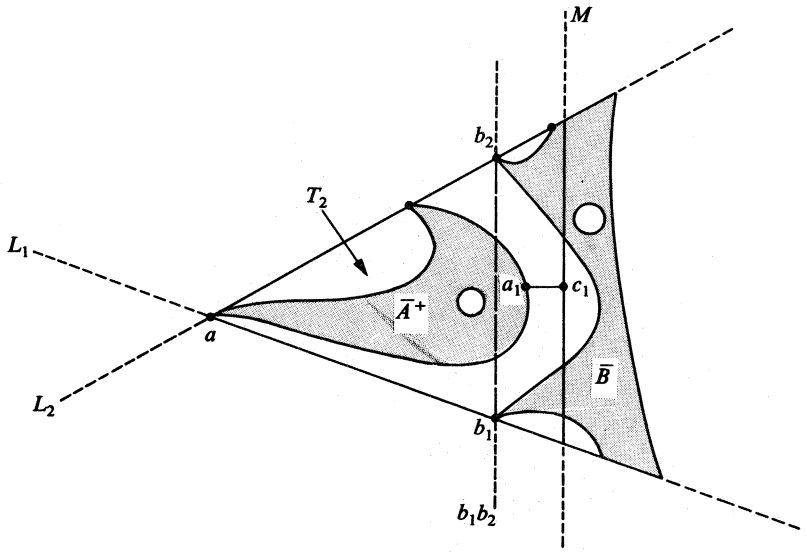


FIGURE 5.

assume that for uncountably many t, $||b_t - b_1'|| \geqslant d$. Then we consider $b_t\, b_1'$ for these t. If $\theta_t = \angle\, b_t b_1' b_2'$ then $\theta_t \to 0$ (since $d(L_0, b_t)/||b_t - b_1'|| \to 0$). It follows that, for uncountably many t tending to 0, $b_t b_1'$ does not meet P^-. If for even a countable sequence $\{t_n\}$ of these t converging to 0, $b_{t_n} b_1'$ contained points of A then these points would lie in $\bar{A} \cap P^+$. As before, the compactness of $\bar{A} \cap P^+$ would produce a contradiction. Case (1) follows similarly.

The spanning requirements of Theorem 2 are necessary, for if

$$A = \{(x,y) | x = 0, 0 \leqslant y \leqslant 1 \text{ or } y = 0, 0 \leqslant x \leqslant 1\},$$
$$B = \{(x,y) | 0 \leqslant x \leqslant 2, y = x - 2\},$$

(only B spans R^2) there is only one monochrome line even when both A and B are compact. In a similar way it is easy to see that even if A is unbounded and spans R^2, all monochrome lines may lie in B. For example, if

$$A = \{(x,y) | x = 0, y \geqslant 0 \text{ or } y = 0, -1 \leqslant x \leqslant 1\},$$
$$B = \{(x,y) | x^2 + y^2 = 2, y \leqslant 0\},$$

then all monochrome lines are in B. It would be interesting to know when an unbounded set must possess a monochrome line.

Finally, the outstanding question of interest concerning monochrome lines in the plane is: *When do two disjoint countable compact sets have a monochrome line?* Motzkin's result answers this if both are finite. One can show that two disjoint planar convergent sequences which do not lie entirely on one line possess a monochrome line. Nothing more appears to be known other than the existence of easy examples which show that some compactness conditions are clearly necessary. Since countable sets tend to be badly disconnected, the present methods appear useless. Equally, the finite case appears to have little direct application.

References

[1] D. Morrison, M. Kiang and J. Wright, On disjoint connected subsets of the square containing pairs of antipodal points, Can. Math. Bull., 9:5 (1966) 631–638.

[2] M.H.A. Newman, Elements of the Topology of Plane Sets of Points, 2nd ed., University Press, Cambridge, 1951.

[3] D. Tingley, Monochromatic lines in the plane, this MAGAZINE 46 (1975) 271–274.

PROBLEMS

DAN EUSTICE, Editor
LEROY F. MEYERS, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before October 1, 1979.

1058. Is it true that a square matrix that is not a scalar multiple of the identity is always similar to a matrix with all non-zero elements? [*H. Kestelman, University College, London.*]

1059. How should n given non-negative real numbers be indexed to minimize (maximize) $a_1a_2+a_2a_3+\cdots+a_{n-1}a_n+a_na_1$? [*David E. Daykin, Reading University.*]

1060. Prove or disprove: There exists a function f defined on $[-1,1]$ with f'' continuous such that $\Sigma_{n=1}^{\infty} f(1/n)$ converges but $\Sigma_{n=1}^{\infty}|f(1/n)|$ diverges. [*Peter Ørno, The Ohio State University.*]

1061. In how many ways can n^2 distinct real numbers be arranged into an $n\times n$ array (a_{ij}) such that $\max_j \min_i a_{ij} = \min_i \max_j a_{ij}$? [*Edward T. H. Wang, Wilfrid Laurier University.*]

1062. a. Let (x_1,y_1), (x_2,y_2), and (x_3,y_3) be three points in the Cartesian plane. Assume the points and their negatives are all distinct. Show that there is an ellipse, centered at the origin, passing through the three points if and only if

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \cdot \begin{vmatrix} x_1 & y_1 & -1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \cdot \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & -1 \\ x_3 & y_3 & 1 \end{vmatrix} \cdot \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & -1 \end{vmatrix} > 0.$$

Interpret this condition geometrically.

b.* Find a necessary and sufficient condition for the existence of an ellipsoid, centered at the origin, passing through four given points in 3-space. [*G. A. Edgar, The Ohio State University.*]

ASSISTANT EDITORS: DON BONAR, *Denison University*; WILLIAM A. MCWORTER, JR., *The Ohio State University. We invite readers to submit problems believed to be new. Proposals should be accompanied by solutions, when available, and by any information that will assist the editors. Solutions to published problems should be submitted on separate, signed sheets. An asterisk (*) will be placed by a problem to indicate that the proposer did not supply a solution. A problem submitted as a Quickie should be one that has an unexpected succinct solution. Readers desiring acknowledgment of their communications should include a self-addressed stamped card. Send all communications to this department to Dan Eustice, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.*

1063. Let M be an $n \times n$ matrix of integers whose inverse is also a matrix of integers. Prove that the number of odd entries in M is at least n and at most n^2-n+1, and that these are the best possible bounds. [*D. A. Moran, Michigan State University.*]

1064. For each positive integer n, define

$$L(n)=\int_0^{\infty}\left(\frac{\sin x}{x}\right)^n dx.$$

It is well known that $L(1)=L(2)=\frac{\pi}{2}$.

a. Find $L(3)$, $L(4)$, and $L(5)$.

b.* Is there a formula for $L(n)$ for general n? [*Edward T. H. Wang, Wilfrid Laurier University.*]

1065. A is an $n+1$ by $n+1$ matrix; its (1, 1)th element is 0 and all others are 1. Find a formula for the elements of A^k when $k \geqslant 2$. [*H. Kestelman, University College, London.*]

Quickies

Solutions to Quickies appear at the conclusion of the Problems section.

Q656. For each positive integer n, show that either

$$\sum_{k=1}^{n} k \equiv 1 \pmod 5 \quad \text{or} \quad \sum_{k=1}^{n} k^2 \equiv 0 \pmod 5.$$

[*Warren Page, New York City Community College.*]

Q657. Find all solutions to the Diophantine equation

$$1!+2!+\cdots+n!=m^2.$$

[*Edward T. H. Wang, Wilfrid Laurier University.*]

Solutions

Three Solutions **March 1977**

1012. Find all solutions (x,y) of $x^y=y^{x-y}$, where x and y are positive integers. [*Gerald E. Gannon and Harris S. Shultz, California State College at Fullerton.*]

Solution: If one of the numbers is 1, then the other must also be 1. Therefore, assume that (x,y) is a solution with $x \geqslant 2$ and $y \geqslant 2$. Then $x^y = y^{x-y} > 1$, so that $x > y$. Dividing both sides by y^y yields $(x/y)^y = y^{x-2y}$. Since $x > y$, $x/y > 1$ and $(x/y)^y = y^{x-2y} > 1$. Thus $x-2y$ is a positive integer and thus $x/y > 2$ and $(x/y)^y$ is a positive integer. This implies that x/y is a positive integer. If $x/y \geqslant 5$, then

$$x/y = y^{(x/y)-2} \geqslant 2^{(x/y)-2} > x/y.$$

Thus $2 < x/y < 5$ or x/y must equal 3 or 4. Taking $x/y = 3$ yields $y = 3$ and $x = 9$. Taking $x/y = 4$ yields $y = 2$ and $x = 8$. Hence the only solutions in positive integers of $x^y = y^{x-y}$ are (x,y) equal to $(1,1)$, $(9,3)$, and $(8,2)$.

ROBERT CLARK, student
Central High School
Philadelphia, Pennsylvania

Also solved by Mangho Ahuja, George Berzsenyi, W. J. Blundon (Canada), Stephen D. Bronn, Stephen Eberhart, Michael W. Ecker, Milton Eisner, Thomas E. Elsner, Win Emmons, Robert S. Fisk & Gary D. Peterson, Marjorie Fitting, Marguerite Gerstell, Michael Goldberg, William E. Gould, M. G. Greening (Australia), Richard A. Groeneveld & Edward Pollak, David Hammer, Daniel L. Hansen, George C. Harrison, Dinh Thế Hùng, J. P. Lambert, Jordan I. Levy, Graham Lord (Canada), James C. McKim, John S. Maginnis, Jerry Metzger, David Montana, William Myers, Roger B. Nelsen, Bob Prielipp, Edith E. Risen, James T. Sandefur, Arthur Solomon, J. M. Stark, David R. Stone, Y. H. Yiu (Hong Kong), and the proposers.

Fibonacci Holes **May 1977**

1013. Let the sequence $\{S_n\}$ be defined by $S_1 = a$, $S_2 = a + b$, and $S_{n,1} = S_n + S_{n-1}$ for $n > 2$, where a and b are distinct positive integers. Define a hole of $\{S_n\}$ as an integer which is not expressible as a sum of distinct terms of $\{S_n\}$. Find a general formula for $J(k)$, the number of holes of $\{S_n\}$ between S_k and S_{k-1}. [*James Propp, 1976 U.S.A. Mathematical Olympiad Training group.*]

Solution: Let $p = S_k + j$ $(1 \leqslant j < S_{k-1})$ be a number between S_k and S_{k+1}. The largest possible sum of distinct terms of $\{S_n\}$ which does not exceed p and which cannot be expressed as a sum using S_k is $S_{k-1} + S_{k-3} + S_{k-5} + \cdots + S_\varepsilon$, where $\varepsilon = 1$ or $\varepsilon = 0$. (If any other S_i were used, we could simplify the sum to obtain a sum using S_k.) Now, $S_{\varepsilon+1} = S_{\varepsilon+3} - S_{\varepsilon+2}$ implies $S_{\varepsilon+3} > S_\varepsilon + S_{\varepsilon+2}$. Also, $S_{\varepsilon+3} = S_{\varepsilon+5} - S_{\varepsilon+4} > S_\varepsilon + S_{\varepsilon+2}$ implies $S_{\varepsilon+5} > S_\varepsilon + S_{\varepsilon+2} + S_{\varepsilon+4}$. Continuing in this fashion we obtain

$$S_k > S_{k-1} + S_{k-3} + S_{k-5} + \cdots + S_\varepsilon.$$

Thus any sum of distinct S_i that equals p contains S_k. Hence $J(k)$ is the number of holes of $\{S_n\}$ between 1 and S_{k-1}. Thus

$$J(k) = \sum_{i=0}^{k-2} J(i) = J(k-2) + \sum_{i=0}^{k-3} J(i) = J(k-2) + J(k-1), \tag{1}$$

where $J(0) = a - 1$ and $J(1) = b - 1$. ($J(0)$ is the number of holes less than S_1.) Therefore we obtain a Fibonacci-type sequence $\{J(n)\}$ with the given initial conditions.

DANIEL S. FREED, student
Harvard University

Editor's Comment. Let f_k denote the kth Fibonacci number ($f_0 = 0$, $f_1 = 1$, and $f_n = f_{n-1} + f_{n-2}$). As many solvers pointed out, the following two formulas are easily proved by induction using (1):

$J(k)=S_k-f_{k+1}$, $k\geqslant 1$, and $J(k)=(a-1)f_{k-1}+(b-1)f_{k-2}$, $k\geqslant 2$. As a corollary, these formulas show the well-known result that every positive integer can be expressed as a sum of distinct Fibonacci numbers.

Also solved by S. Floyd Barger, Donald Batman, J. C. Binz (Switzerland), Clayton W. Dodge, Michael W. Ecker, Thomas E. Elsner, Donald C. Fuller, George C. Harrison, John S. Maginnis, Robert Patenaude, Scott Smith, J. M. Stark, Gillian W. Valk, Robert Wood, and Randall Dougherty & the proposer.

Power of Ten **May 1977**

1016. For n a positive integer, describe all n-digit numbers x with the property that there exists a permutation y of the digits of x such that $x+y=10^n$. [*Michael W. Ecker, City University of New York.*]

Solution: Let x be an n-digit solution. Since $10^k x$ is an $(n+k)$-digit solution, we may assume without loss of generality that

$$x=a_n10^{n-1}+a_{n-1}10^{n-2}+\cdots+a_2 10+a_1$$

where $a_1\neq 0$. Let $y=10^n-x=\bar{a}_n10^{n-1}+\bar{a}_{n-1}10^{n-2}+\cdots+\bar{a}_2 10+\bar{a}_1$. Then $a_1+\bar{a}_1=10$ and $a_j+\bar{a}_j=9$ for $j=2,3,\ldots,n$. Hence

$$\sum_{i=1}^{n}(a_i+\bar{a}_i)=2\sum_{i=1}^{n}a_i=9(n-1)+10$$

and so

$$\sum_{i=1}^{n}a_i=9(n-1)/2+5.$$

Therefore n is odd and, since for $j=2,\ldots,n$, a_j is paired with a distinct $\bar{a}_j$ so that $a_j+\bar{a}_j=9$, by "casting out nines" in x we get $(n-1)/2$ pairs of nines and $a_1=\bar{a}_1=5$.

Thus x is described as follows: $a_1=5$ and the digits $a_2,\ldots,a_n$ consist of $(n-1)/2$ pairs of integers, each pair having sum 9, arranged in arbitrary order. The general solution is obtained by multiplying x by any power of 10.

RICHARD A. GIBBS
Fort Lewis College

Also solved by Milton Eisner, Thomas Elsner, Daniel S. Freed, Tim Fynskov, George C. Harrison, Dinh The Hūng, Eli L. Isaacson, John S. Maginnis, Victor Pambuccian (Romania), Gillian W. Valk, Edward T. H. Wang (Canada), Robert Wood, and the proposer.

$A^p=I$ **May 1977**

1017. (a) Given an $n\times n$ matrix A over the rationals, show that $A^p=I$ for a prime $p>n+1$ implies $A=I$.

(b) For each k, $1<k\leqslant n+1$, show that there exists an $n\times n$, non-identity matrix over the rationals such that $A^k=I$. [*Stanley Friedlander, Bronx Community College*]

Solution: (a) For an $n\times n$ matrix A, $A^k=I$ if and only if $c(x)$, the minimum polynomial of A, divides $x^k-1=(x-1)(x^{k-1}+x^{k-2}+\cdots+1)$. If $A^k=I$ and k is a prime, then the two factors of x^k-1 are irreducible over the rationals and, if $k>n+1$, then the degree of the second factor is greater than n and thus greater than the degree of $c(x)$. Consequently, $c(x)$ must be $x-1$ and $A=I$.

(b) If $k\leqslant n+1$, then a suitable matrix A is the direct sum of the $(k-1)\times(k-1)$ companion matrix of $x^{k-1}+x^{k-2}+\cdots+1$ and the $(n+1-k)\times(n+1-k)$ identity matrix.

Remark: It follows that there exists an $n\times n$ matrix $A\neq I$ over the rationals such that $A^k=I$ if and only if k has at least one prime factor $p\leqslant n+1$. For, if $k=pb$ with $p\leqslant n+1$, then the matrix $A\neq I$ constructed in (b) so that $A^p=I$ is such that $A^k=A^{pb}=I$. On the other hand, if the prime

factorization of k is $p_1p_2\cdots p_m$ with each $p_i>n+1$, then for any $n\times n$ matrix $A\neq I$, the repeated use of (a) shows that $A^k=A^{p_1p_2\cdots p_m}\neq I$.

NORMAN M. RICE
Queen's University, Kingston, Canada

Also solved by Donald C. Fuller, Paul K. Garlick, Eli L. Isaacson, Steve Kahn, Richard Levaro, St. Olaf Problems Group, J. M. Stark, L. Van Hamme (Belgium), Edward G. H. Wang (Canada), Qazi Zameeruddin (India), and the proposer.

Rotating a Polygon **May 1977**

1018. Let $P_1,P_2,\ldots,P_n$ be the vertices in order of a convex n-gon with θ_r, $0<\theta_r<\pi$, as the angle at P_r. Rotations $R_1,R_2,\ldots,R_n$ are defined as follows: R_1 rotates $2\theta_1$ about P_1, R_2 rotates $2\theta_2$ about $R_1(P_2)$, R_3 rotates $2\theta_1$ about $R_2R_1(P_1)$, etc. Prove that $R_nR_{n-1},\ldots,R_2R_1$ is the identity. [*H. Kestelman, University College, London.*]

Solution: Let G_0 be the original position of the polygon, and $R_1(G_0)=G_1$, $R_2R_1(G_0)=G_2$, and so on. Let H be the reflection of G_0 in P_1P_n, and $P_2',\ldots,P_{n-1}'$ be the reflections of $P_2,P_3,\ldots,P_n$ respectively in P_1P_n. The rotation R of G_0 can be executed by first reflecting G_0 in P_1P_n, thus getting H, and then reflecting H in P_1P_2'. Thus $G_1=R_1(G_0)$ is the reflection of H in P_1P_2'. The rotation $R_2R_1(G_0)$ can be obtained by first reflecting G_1 in $P_2'P_1$, thus getting H, and then reflecting H in $P_2'P_3'$. Thus G_2 is the reflection of H in $P_2'P_3'$. Continuing in this manner, $R_{n-1}\ldots R_3R_2R_1(G_0)=G_{n-1}$ is the reflection of H in P_nP_{n-1}'. Finally, $R_nR_{n-1}\ldots R_3R_2R_1(G_0)=G_n$ is the reflection of H in P_nP_1. But H is the reflection of G_0 in P_1P_n, hence its reflection in P_1P_n is G_0 itself. Hence $R_nR_{n-1}\ldots R_2R_1$ is the identity.

MANGHO AHUJA
Southeast Missouri State University

Also solved by Clayton W. Dodge, Howard Eves, Daniel S. Freed, Michael Goldberg, Steven L. Hergenrother, Graham Lord (Canada), James T. Smith, J. M. Stark, and the proposer.

Characteristic of a Ring **May 1977**

1019. Let R be a ring for which there is an integer n, $n>1$, such that $x^n=x$ for each element x of R. Prove that the characteristic of R is a (square-free) product of distinct primes p such that $(p-1)|(n-1)$. [*Daniel Mark Rosenblum, Carnegie-Mellon University.*]

Solution: For each x in R and k an integer let $kx=(x+\cdots+x)$ (k times). Then $(kx)^n=(kx)$ implies $k^nx=kx$, or $(k^n-k)x=0$. Therefore, if the characteristic of R is c, then $c|k^n-k$ for all integers k. In particular, if $p^2|c$ then $p^2|p^n-p$, which is impossible, so c is square-free.

If $p|c$ where p is prime, then let a be an integer such that $\{0,a,a^2,\ldots,a^{p-1}\}$ is a complete set of residues mod p. Since $p|c$, then $p|a^n-a$ or $a^n\equiv a \pmod p$, which implies $a^{n-1}\equiv 1 \pmod p$. But the least positive number m such that $a^m\equiv 1 \pmod p$ is $p-1$; therefore, the greatest common divisor of $p-1$ and $n-1$ is $p-1$ which implies $(p-1)|(n-1)$.

TOBIAS ORLOFF
Massachusetts Institute of Technology

Also solved by Thomas E. Elsner, Paul K. Garlick, Lee O. Hagglund, David Hammer, Richard Levaro, John S. Maginnis, Jerry Metzger, James J. Reynolds, Gazi Zameeruddin (India), and the proposer. A. Wilansky points out that such rings are characterized in Problem 5972, *Amer. Math. Monthly*, 83 (1976) 66–67.

Determinant Equation **May 1977**

1020. For $i=1$, 2, and 3, let the circle C_i have center (h_i,k_i) and radius r_i. Find a determinant equation for the circle orthogonal to these three given circles which generalizes the well-known result for the circle through three points. [*Leon Gerber, St. John's University.*]

Solution: Let the required circle C have center (h,k) and radius r. Since C is orthogonal to C_i, $i=1,2,3$, we have

$$(h-h_i)^2+(k-k_i)^2=r^2+r_i^2,$$

or

$$h_i^2+k_i^2-r_i^2-2hh_i-2kk_i+h^2+k^2-r^2=0, \quad i=1,2,3. \tag{1}$$

But an equation of C is

$$x^2+y^2-2hx-2ky+h^2+k^2-r^2=0. \tag{2}$$

From (1) and (2) it now follows, by elimination of h, k and $h^2+k^2-r^2$, that C is also given by

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ h_1^2+k_1^2-r_1^2 & h_1 & k_1 & 1 \\ h_2^2+k_2^2-r_2^2 & h_2 & k_2 & 1 \\ h_3^2+k_3^2-r_3^2 & h_3 & k_3 & 1 \end{vmatrix}=0. \tag{3}$$

When $r_1=r_2=r_3=0$, (3) reduces to the familiar determinant equation for the circle through the three points $(h_1,k_1),(h_2,k_2),(h_3,k_3)$.

HOWARD EVES
University of Maine at Machias

Also solved by Mangho Ahuja, Michael Goldberg, Sister Stephanie Sloyan, J. M. Stark, and the proposer. Michael Goldberg supplied a reference to Analytical Conics *by D. M. Y. Sommerville, G. Bell and Sons, London,* 1924, *pp.* 88–89. *See also* A Treatise on Conic Sections *by G. Salmon, Chelsea,* 1954, *pp.* 102 *and* 130.

Converges to One **September 1977**

1021. Prove or disprove that a countably infinite set of positive real numbers with a finite non-zero cluster point can be arranged in a sequence, $\{a_n\}$, so that $\{(a_n)^{1/n}\}$ is convergent. [*Peter Ørno, The Ohio State University.*]

Solution: Let $\{x_n\}$ denote the real numbers of concern, and let a be a non-zero finite cluster point. Choose A such that $1/A<a<A$, and let a_1 denote the x_n of smallest subscript which lies in the interval $(1/A,A)$. Assuming $a_1,a_2,\ldots,a_{k-1}$ to have been chosen, pick a_k to be the x_n of smallest subscript different from the $n-1$ already chosen lying in the interval $(A^{-\sqrt{n}},A^{\sqrt{n}})$. Since a is a cluster point such an x_n can be found, and since $A^{\sqrt{n}}\to\infty$ and $A^{-\sqrt{n}}\to 0$ it is clear that every x_n eventually is included in an interval of the form $(A^{-\sqrt{n}},A^{\sqrt{n}})$ and will therefore eventually become part of the sequence $\{a_n\}$. For each n we have $1/A^{\sqrt{n}}<a_n<A^{\sqrt{n}}$ and therefore $(1/A)^{1/\sqrt{n}}<a_n^{1/n}<A^{1/\sqrt{n}}$. But $\lim_{n\to\infty} A^{1/\sqrt{n}}=\lim_{n\to\infty}(1/A)^{1/\sqrt{n}}=1$ and therefore $\lim_{n\to\infty} a_n^{1/n}=1$.

TIM HESTERBERG, student
Saint Olaf College

Also solved by Michael W. Ecker, Donald C. Fuller, Marguerite Gerstell, G. A. Heuer, K. E. Hirst, Ralph Jones, William Myers, J. M. Stark, Samuel Weinberger, Joseph E. Yukich, and the proposer.

1022. We have n cards numbered 1 through n. Find the expected number of drawings needed to put the cards in order by each of the following strategies:

(a) The shuffled cards are drawn without replacement until card 1 is drawn. The remaining $n-1$ cards are shuffled and drawn without replacement until card 2 is drawn. This process is continued until all the cards are drawn and put in linear order.

(b) A card, say card k, is drawn from the shuffled deck. The remaining cards are shuffled and drawn without replacement until either card $k-1$ or card $k+1$ is drawn. We identify card $1-1$ as card n and card $n+1$ as card 1. This process is continued until all the cards are drawn and put in circular order. [*Joe Dan Austin, Emory University.*]

Solution: Let $X_1, X_2, X_3, \ldots, X_n$ be the number of drawings needed to put the 1st, 2nd, 3rd, ..., nth card in order. Then the expected number of drawings is $\sum_{i=1}^{n} X_i$.

(a) For each i, $1 \leqslant i \leqslant n$, cards are drawn without replacement from a deck of $n-i+1$ cards until card i is drawn. The probability of this occurring on the jth draw, $1 \leqslant j \leqslant n-i+1$, is

$$P[X_i=j] = \frac{n-i}{n-i+1} \cdot \frac{n-i-1}{n-i} \cdot \frac{n-i-2}{n-i-1} \cdot \cdots \cdot \frac{n-i-j+2}{n-i-j+3} \cdot \frac{1}{n-i-j+2}$$
$$= \frac{1}{n-i+1}.$$

Hence

$$E[X_i] = \sum_{j=1}^{n-i+1} j \cdot \frac{1}{n-i+1} = \frac{n-i+2}{2}$$

and the expected number of drawings is

$$\sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{n-i+2}{2} = \frac{1}{4} n(n+3).$$

(b) $X_1 = X_n = 1$, and, for each i, $2 \leqslant i \leqslant n-1$, cards are drawn without replacement from a deck of $n-i+1$ cards until either one of two required cards is drawn. The probability of this occurring on the jth draw, $1 \leqslant j \leqslant n-i+1$, is

$$P[X_i=j] = \frac{n-i-1}{n-i+1} \cdot \frac{n-i-2}{n-i} \cdot \frac{n-i-3}{n-i-1} \cdot \cdots \cdot \frac{n-i-j+1}{n-i-j+3} \cdot \frac{2}{n-i-j+2} = \frac{(n-i-j+1)2}{(n-i+1)(n-i)}.$$

Hence, for $2 \leqslant i \leqslant n-1$,

$$E[X_i] = \sum_{j=1}^{n-i+1} j \cdot \frac{(n-i-j+1)2}{(n-i+1)(n-i)} = \frac{n-i+2}{3}$$

and the expected number of drawings is

$$\sum_{i=1}^{n} E[X_i] = 1 + \sum_{i=2}^{n-1} \frac{n-i+2}{3} + 1 = 1 + \frac{1}{6} n(n+1).$$

P. J. PEDLER
Mount Lawley College, Australia

Also solved by Bern Problem Solving Group (Switzerland), J. C. Binz (Switzerland), Nicholas Birkett (Canada), Walter Bluger (Canada), Landy Godbold, Richard A. Groeneveld, Keith Hodge, William Myers, Western Maryland College Problem Seminar, and the proposer. Partial solutions by John Atkins, Stephen D. Bronn, Adam Chu, and G. A. Heuer.

Many Superheros **September 1977**

1023*. Call a triangle *super-Heronian* if it has integral sides and integral area, and the sides are consecutive integers. Are there infinitely many distinct super-Heronian triangles? [*Steven R. Conrad, Benjamin N. Cardozo, H. S., Bayside, N.Y.*]

Editor's Comment. Because of the large number of solvers of this problem, we must forgo the printing of their names. Most solvers, by letting the sides of the triangle be $a-1$, a, and $a+1$, and observing that a must be even for integral area, reduced the problem to Pell's equation $n^2-3m^2=-3$ with $a=2m$. The infinitely many solutions to Pell's equation yield triangles with sides $2m-1$, $2m$, and $2m+1$ and area nm. A few solvers used different methods. The following solution was one of the most original. While it does not find all super-Heronian triangles, it shows that there are infinitely many.

Solution: With sides $a-1$, a, and $a+1$, the area A is $(a/4)\sqrt{3(a^2-4)}$. A is an integer if $\frac{1}{4}\sqrt{3(a^2-4)}$ is an integer. Now $a_{n+1}=a_n^2-2$ with $a_1=14$ defines a recurrent sequence such that $\frac{1}{4}\sqrt{3(a_n^2-4)}$ is always an integer. To see this, we find for $n=1$ that $\frac{1}{4}\sqrt{3\cdot 192}=6$. Now suppose that $\frac{1}{4}\sqrt{3(a_n^2-4)}$ is an integer. Then

$$\tfrac{1}{4}\sqrt{3(a_{n+1}^2-4)}=\tfrac{1}{4}\sqrt{3\left((a_n^2-2)^2-4\right)}=\tfrac{1}{4}\sqrt{3(a_n^4-4a_n^2)}=\frac{a_n}{4}\sqrt{3(a_n^2-4)}$$

is also an integer. Therefore, there exists for every positive integer n a super-Heronian triangle with sides a_n-1, a_n, and a_n+1 and area $(a_n/4)\sqrt{3(a_n^2-4)}$. They are easily seen to be distinct and non-similar.

PETER ADDOR, student
University of Bern, Switzerland

Percentage vs. Games Behind **September 1977**

1024. In many athletic leagues the progress of teams is reported both in terms of winning percentage and in terms of "games behind" the league leader, defined as the difference in games won minus the difference in games lost, divided by 2. Sports fans often observe, especially early in the season, that the league leader in percentage (the official standard) is behind some other team in games.

Suppose Team A is the percentage leader, but Team B is ahead of Team A in games. Assume no ties.

a. Which team has played more games?

b. What is the minimum difference in number of games played?

c. Characterize possible won/lost records for the two teams if the difference in number of games played is minimal.

d. Is it possible for this anomaly to occur late in the season?

[*David A. Smith, Duke University.*]

Solution: Let w_B and w_A be the number of games won by Team B and Team A respectively. Similarly n_B and n_A are the number of games played, and p_B and p_A are the winning percentages.

From the hypotheses, we see that (1) $w_B-w_A>(n_B-n_A)/2$ and (2) $w_B-w_A<p_A(n_B-n_A)$ and (3) $p_A>1/2$.

a. From the above inequalities it follows directly that $\frac{1}{2}(n_B-n_A)<p_A(n_B-n_A)$, which from (3) gives $n_B-n_A>0$.

b. Since w_B-w_A is a positive integer and $p_A\leqslant 1$, inequalities (1) and (2) force $n_B-n_A\geqslant 3$. That n_B-n_A is acceptable is clear from the case where $w_A=n_A=1$ and $w_B=3$, $n_B=4$.

c. If $n_B - n_A = 3$, then $3/2 < w_B - w_A < 3p_A \leq 3$, so $w_B - w_A = 2$. Since $w_A/n_A > (w_A+2)/(n_A+3)$ iff $p_A > 2/3$ we conclude that if Team A has won W games and lost L, then $W > 2L$ and Team B has won $W+2$ games and lost $L+1$.

d. That the anomaly can occur late in the season is evident from part c. However, if $p_A < 2/3$, the difference in games played must be larger than 3. In 1977 the league leaders in major league baseball near the end of the season had values of p_A near .63. This requires a games played difference of at least 5, which may not be likely. (An example of this situation is when $w_A = 94$, $w_B = 97$, $n_A = 150$, and $n_B = 155$. Then $p_A = .627$ and $p_B = .626$.)

LANDY GODBOLD
Westminster Boys School

Also solved by The Augusta College Problems Group, Douglas A. Bies, Noël Cortey, Milton Eisner, Thomas E. Elsner, Robert S. Fisk, Donald C. Fuller, Steve Kahn, Frederick R. Lane, Richard Vogt, Western Maryland Problem Seminar, and the proposer.

Convergent Subseries **November 1977**

1025. Let $\{a_n\}$ be a sequence of positive real numbers with $\Sigma a_n = \infty$ and $\Sigma a_n^2 < \infty$. For a given $C > 0$, the sequence $\{m_i\}$ of positive integers is such that $\Sigma a_n > C$, the sum being over those n such that $m_i < n \leq m_{i+1}$.

a. Prove that there is a sequence $\{p_i\}$ with $m_i < p_i \leq m_{i+1}$, such that $\Sigma a_{p_i} < \infty$.

b. Show by an example that Σa_{p_i} need not converge for all such $\{p_i\}$. [*W. C. Waterhouse, The Pennsylvania State University.*]

Solution: a. Let $M_i = \{n : m_i < n \leq m_{i+1}\}$. For each i, there exists a p_i such that $a_{p_i} = \min\{a_n : n \in M_i\}$. Then

$$Ca_{p_i} < \sum_{n \in M_i} a_n a_{p_i} \leq \sum_{n \in M_i} a_n^2.$$

Thus

$$\sum_{i=1}^{\infty} a_{p_i} < \frac{1}{C} \sum_{i=1}^{\infty} \sum_{m \in M_i} a_n^2 = \frac{1}{C} \sum_{m=1}^{\infty} a_n^2.$$

b. Let

$$a_n = \begin{cases} 1/n, & \text{if } n \text{ is not a positive integral power of 2,} \\ 1/m, & \text{if } n = 2^m. \end{cases}$$

Then $a_n \geq 1/n$, so $\Sigma a_n \leq \Sigma 1/n = \infty$. But for any $k > 0$,

$$\sum_{n=1}^{k} a_n^2 \leq \sum_{n=1}^{k} \frac{1}{n^2} + \sum_{m=1}^{k} \frac{1}{m^2} \leq 2 \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Thus $\Sigma_{n=1}^{\infty} a_n^2 < \infty$. Now let $C = 1/3$ and $m_i = 2^{i-1}$. Using M_i as in part a. above, we have $a_n \geq 1/2^i$ for $n \in M_i$, and

$$\sum_{n \in M_i} a_n \geq \frac{m_i}{2^i} = \frac{1}{2} > \frac{1}{3} = C.$$

But we may take $p_i = 2^i$. Then $p_i \in M_i$ and $\Sigma a_{p_i} = \Sigma 1/i = \infty$. Hence not every such p_i will work.

CURT MCMULLEN, student
Williams College

Also solved by Adam Riese, St. Olaf Problems Group, Benjamin L. Schwartz, J. M. Stark, Lou Thurston, Joseph E. Yukich, and the proposer.

1026. A decomposition of a positive integer n is an ordered tuple $(n_1, n_2, \ldots, n_k)$ of positive integers such that $\sum_{i=1}^{k} n_i = n$. Find the total number of decompositions of n which are palindromes. For example, for $n=4$, there are four such, namely: (4), (2,2), (1,1,1,1), and (1,2,1). [*Michael Capobianco, St. John's University.*]

Solution: For each palindromic decomposition of n, we construct two such decompositions of $n+2$, according to these rules:

(a) Append a 1 to each end of the original decomposition.

(b) Add 1 to the first and last numbers of the original decomposition (or add 2 if the original decomposition consisted of exactly one number).

That this generates all possible palindromic decompositions of $n+2$, without duplicating any, can be seen by imagining the reverse process, which is a function well-defined on the domain of palindromic decompositions of any $n+2$:

(a′) If a palindromic decomposition has 1's at both ends, delete them.

(b′) If a palindromic decomposition does not have 1's at both ends, subtract 1 from the number at each end (or 2, if there is only one number).

Since the number of palindromic decompositions of 1 and 2 are 1 and 2 respectively, we have $2^{[n/2]}$ palindromic decompositions of each positive integer n, where $[n/2]$ denotes the greatest integer less than or equal to $n/2$.

MARGUERITE GERSTELL
Florida Institute of Technology

Also solved by Bern Problem Solving Group (Switzerland), Richard Bland College Problems Group, Elwyn H. Davis, H. O. Eberhart, Michael W. Ecker, Milton Eisner, Howard Eves, Phil K. Garlick, Richard A. Gibbs, Landy Godbold, Heiko Harborth (West Germany), Dinh Thê Hūng, Jimmy Johnston, Kantonsschule Zürcher Unterland Problems Solving Group (Switzerland), Kirchenfeld Gymnasium Problems Group (Switzerland), Lew Kowarski, Alan M. Kriegsman, Graham Lord (Canada), Fred Leung, Mary Helen Manning, Jerry Metzger, William Myers, Bill Peters, Reinhard Razen (Austria), Elizabeth Rentmeesters, James J. Reynolds, Robert Scherrer, F. G. Schmitt, Jr., Harry T. Sedinger, Maurice Shrader-Frechette, Blair Spearman, Philip Straffin, Terry Therneau, Lou Thurston, Gillian W. Valk, Matt Wyneken, Ken Yocum, and the proposer. There was one unsigned solution. Razen supplied a reference to V. E. Hoggatt, Jr. and M. Bicknell, Palindromic compositions, Fibonacci Quart., 13(1975) 350–356.

Answers

Solutions to the Quickies which appear near the beginning of the Problems section.

Q656. $5\sum_{k=1}^{n} k^4 = \dfrac{n(n+1)(2n+1)(3n^2+3n-1)}{6} = \left(\sum_{k=1}^{n} k^2\right)\left(6\sum_{k=1}^{n} k - 1\right).$

Q657. Since the last digit of $n!$ is 0 for all $n \geqslant 5$, the last digit of the left hand side is 3 for all $n \geqslant 5$. Clearly m must be odd. Since the last digit of the square of an odd integer must be 1, 5, or 9, there are no solutions for $n \geqslant 5$. Examining the cases when $n \leqslant 4$ shows that there are exactly two solutions given by $n=m=1$ and $n=m=3$.

REVIEWS

PAUL J. CAMPBELL, Editor
Beloit College

PIERRE MALRAISON, Editor
Control Data Corp.

Assistant Editor: Eric S. Rosenthal, Princeton University. Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of the mathematics literature. Some reviews of books are adapted from the Telegraphic Reviews in the American Mathematical Monthly.

Lin, C.C., *Education of applied mathematicians,* SIAM Review 20 (October 1978) 838-845.

The author makes a number of recommendations for education at the undergraduate level, not the least of which is for a two-year sophomore-junior sequence of courses surveying applied mathematics, in addition to the study of calculus, complex variables, differential equations, and linear algebra.

Bartusidle, Marcia F., *Calculatoritis,* Science News 80 (18 November 1978) 347.

A humorous note on the mental laziness calculators seem to spawn and "the fear of performing the simplest arithmetic problems without running to a calculator."

Gurland, Robert H., *Teaching mathematics,* in Steven M. Cahn (Ed), Scholars Who Teach: The Art of College Teaching, Nelson-Hall, 1978; pp. 75-100; xi + 246 pp.

The author, professor of philosophy and formerly professor of mathematics, offers useful (if general) advice built around the enterprise of dealienating students from mathematics.

Kolata, Gina Bari, *Computer science: surprisingly fast algorithms,* Science 202 (24 November 1978) 857-858.

The new fast algorithms are for manipulating polynomials and power series. One of the results shows how nonintuitive they are: any power of a polynomial can be computed as quickly as the square of it!

Moler, Cleve and Van Loan, Charles, *Nineteen dubious ways to compute the exponential of a matrix,* SIAM Review 20 (October 1978) 801-836.

The article offers an excellent example of the analysis of algorithms.

Press, William H., *Mathematical theory of the waterbed,* American J. Physics 46 (October 1978) 966-970.

After several pages of modelling, the important basic problems remain unsolved: "For what range of mattress fullness and separation do flat bodies tend to slide toward each other?...How full should a waterbed be?"

Steen, Lynn A., *A new perspective on infinity*, New Scientist (9 November 1978) 448-451.

Extended popular account of nonstandard analysis, its development and its significance.

Kolata, Gina Bari, *Anti-semitism alleged in Soviet mathematics*, Science 202 (1978) 1166-1167.

Article on alleged instances and causes of anti-semitism specifically directed at Soviet mathematicians (including G. Margoulis who was not allowed to go to Helsinki to accept his Fields medal). Includes a list of suggestions from Soviet mathematicians on how to protest effectively.

Mbili, L.S.R., Mathematical Challenge! 100 Problems for the Olympiad Enthusiast, Dept. of Mathematics, University of Cape Town, South Africa, 1978; ii + 50 pp, 80 c (P).

Collection of original Olympiad-type problems by a 23-year-old Zulu actuarial student.

Box, Joan Fisher, R.A. Fisher: The Life of a Scientist, Wiley, 1978; xii + 512 pp.

R.A. Fisher, a founding father of modern statistical reference, worked mainly in genetics and evolution. This biography by his daughter deals primarily with the development of his scientific ideas but also discusses his personal life--not as separate from that development, but as an integral part of it.

Woodcock, Alexander and Davis, Monte, Catastrophe Theory, Dutton, 1978; viii + 152 pp, $9.95.

This is catastrophe theory for T.C. Mits (The Common Man in the Street). After discussing the history of the subject and some of its controversy, the authors present a large collection of applications, from relatively uncontroversial ones in the hard sciences to more speculative ones in the "soft" sciences. The book itself is not deep in mathematics, but the short bibliography points to more sophisticated treatments.

Schaaf, William L., Mathematics and Science: An Adventure in Postage Stamps, NCTM, 1978; 152 pp, $7 (P).

How mathematics evolved, its creators, and its relation to science, illustrated with postage stamps.

Beruskin, Alan D., *Tuning the ill-tempered clavier*, American J. Physics 46 (August 1978) 792-795.

Discussion in mathematics courses of musical scales generally concludes with equal tempering. The author points out that pianos are *not* tuned to equal temper, but to a "stretched" tuning (sharp in treble, flat in bass) *unequivocally* preferred by listeners. Data from tuned pianos allow fitting a curve to the deviation from equal temper, and the article points out that the "stretched" tuning can then be accomplished using modern instrumentation (pocket calculator, digital frequence meter).

Abbott, J.C., The Chauvenet Papers: A Collection of Prize-Winning Expository Papers in Mathematics, 2 vols., MAA, 1978; xvii + 595 pp.

Along with the splendid papers themselves, the volumes contain updates (for the older papers) and biographical sketches of the authors. Among the authors are Hardy, Halmos, Chern, Kac, Henkin, and MacLane; the topics cover the broad spectrum of modern mathematics.

Kastner, Bernice, Applications of Secondary School Mathematics, NCTM, 1978; vi + 106 pp, (P).

A valuable collection of topics to answer "what's it good for?": perception, growth models, wave motion, optics, economic models, music and mathematics, relativity, structure and synthesis of protein models. Intended for the teacher, this booklet makes a handy supplement for student use as well.

Brown, Stephen I., Some "Prime" Comparisons, NCTM, 1978; x + 106 pp, $6 (P).

This booklet provides a thought-provoking introduction to elementary number theory by comparing concepts in N to their analogue in E = 1,2,4,6,8,10,12,... . It could provide excellent enrichment material for intellectually-alive high school students, serve as a module for a course in liberal arts mathematics, or serve as a "meta-text" in courses in teacher education.

Ladany, Shaul P. and Machol, Robert E., Optimal Strategies in Sports, North-Holland/American Elsevier, 1977; xx + 231 pp, $20.

The 34 short articles here, some of them adapted from elsewhere, apply quantitative methods and systems analysis to sports competition. Most make specific strategic recommendations to managers and competitors. An American orientation prevails, with 13 articles on baseball and 5 on football; but basketball, hockey, cricket, tennis, golf, swimming, horse racing, rowing, and track and field all come under scrutiny. A short preface explains technical terms and notation for the benefit of sports managers and fans, and an annotated bibliography concludes the collection.

Lockwood, James R. and Runion, Garth R., Deductive Systems: Finite and Non-Euclidean Geometries, NCTM, 1978; vi + 90 pp, $4 (P).

Short treatment of finite, Lobachevskian, and Riemannian geometries, suitable for use as a supplement at high school or college level.

Bartlett, Albert A., *Forgotten fundamentals of the energy crisis*, American J. Physics 46 (September 1978) 876-888.

"Many people find it hard to believe that when the rate of consumption is growing a mere 7%/year, the consumption in one decade exceeds the total of all of the previous consumption." So saying, the author relies upon "the pristine simplicity of elementary mathematics" (of exponential growth) to offer an unequivocal understanding of the origins, magnitude, and implications of the energy crisis.

Ball, John A., Algorithms for RPN Calculators, Wiley, 1978; xii + 330 pp.

The "RPN" stands for "reverse Polish notation;" the most prominent manufacturer of such calculators is Hewlett-Packard. At least four other companies produce them, however, and the author cannot be accused of propagandizing on behalf of any one manufacturer. He succinctly compares the properties of several recent models, demonstrates the economy of the RPN notation, emphasizes key techniques in writing algorithms, and suggests improvements in design of the machines. About 40% of the book is devoted to a collection of algorithms for various purposes.

Peelle, Howard A., *Learning mathematics with recursive computer programs*, Journal of Computer-Based Instruction 3 (Feburary 1977) 97-102.

Suggests the use of recursive computer programs to introduce the concept of recursion, as a stimulant, to students learning mathematics. Three sample APL programs are listed and analyzed (formula for triangular numbers, Newton's rule for square roots, Eratosthenes' sieve).

NEWS & LETTERS

GEOMETRY CONFERENCE

The Seventh Annual Mathematics and Statistics Conference at Miami University, Oxford, Ohio, will be held September 28-29, 1979. The theme for this year's conference will be "Geometry." Featured speakers will include Professors Branko Grünbaum of the University of Washington, and Ernest E. Schult of Kansas State University. There will be sessions of contributed papers, which should be suitable for a general audience of mathematicians and students who are not necessarily experts in geometry. Abstracts should be sent by June 1, 1979 to Professor David Kullman, Department of Mathematics and Statistics, Miami University, Oxford, Ohio 45056. Information concerning preregistration, housing, etc., may also be obtained from the above address.

The Ohio Delta Chapter of Pi Mu Epsilon will also hold its annual student conference September 28-29, 1979. Undergraduate mathematics students are invited to contribute papers, and should send abstracts to Professor Milton Cox, Department of Mathematics and Statistics, Miami University, Oxford, Ohio 45056.

TRIPLE JEOPARDY

If a Pythagorean triple consists of positive integers, then the parameter n used in H. Klostergaard's "Tabulating All Pythagorean Triples" (this *Magazine*, September 1978, pp. 226-227) is just the inradius of the corresponding triangle.

It is well known (B.M. Stewart, *Theory of Numbers*, Macmillan, 1975, pp. 90-91) that all positive primitive Pythagorean triples (x,y,z) having inradius n can be found by factoring n in all possible ways as a product uv, where u and v are relatively prime positive integers and u is odd. For each factorization, put $x = 2n + \min(u^2, 2v^2)$, $y = 2n + \max(u^2, 2v^2)$, and $z = 2n + u^2 + 2v^2$. Then $x < y < z$ and the divisor d in Klostergaard's note is the quantity $\min(u^2, 2v^2)$. When one wants to generate a table of primitive triplets only, this method has the computational advantage that it requires a search for odd divisors u of n which are relatively prime to n/u, rather than a search for divisors d of $2n^2/d$.

The estimates in Klostergaard's note can be improved slightly. If a table has been made for $1 < n < N$, then all $x < 2N + 3$, all $y < (N+1)(2+2\sqrt{2})$, and all $z < (N+1)(2+2\sqrt{2})$ which are members of Pythagorean triples will have appeared.

Simple formulas for the number of Pythagorean triplets corresponding to a given inradius can be found in the book by Stewart cited above.

Daniel Drucker
Wayne State University
Detroit
Michigan 48202

FROM 272 TO 6000

In Professor Erdös' enlightening "A Property of 70" (this *Magazine*, September 1978, pp. 238-240) he describes an infinite set of sequences and states that for all sufficiently large n, at least one of the terms of the sequences beginning with n, other than the first, is the product of exactly two primes. Professor Pomerance proved this result for n greater than 6000, and he showed that when $n = 272$, the sequence contains no term which is the product of exactly two primes.

I have now strengthened the theorem by proving that for $n > 272$, at least one of the terms is the product of exactly two distinct primes. To do this, I assumed that if a sequence con-

tains terms which are products of exactly two distinct primes, then the first of these terms will occur early in the sequence, and the labor of finding it is not prohibitive especially if the researcher develops useful techniques as he goes along. For the region of interest, this proved to be convenient and true.

Using only pencil, paper, and a factor book, and entirely without the use of a computer or calculator, I was able to list for each n above 272 and below 6001 one factor which satisfies the theorem. This work was completed in one night, except for later checking to correct minor errors.

Samuel Yates
104 Brentwood Drive
Mt. Laurel
New Jersey 08054

In "On a Class of Relatively Prime Sequences," that will appear in Volume 10 of the *Journal of Number Theory*, Erdös and I prove that the set n for which no term in the sequence beginning with n is the product of two distinct primes is {1, 2, 3, 4, 6, 7, 8, 11, 12, 15, 17, 18, 22, 23, 24, 29, 30, 35, 39, 43, 44, 69, 70, 103, 104, 119, 268, 271, 272}. The proof uses some fairly deep estimates involving $\pi(x)$, the number of primes less than x, and is in fact valid for all $n \geq 26569$.

David E. Penney
University of Georgia
Athens
Georgia 30605

AN UNSETTLING AMENDMENT

I wonder how many people are aware that the Bill of Rights consists of the ten ratified amendments out of a dozen submitted to the states for ratification in 1789. Here is the text of the first article, which was not ratified:

Article I. After the first enumeration required by the first article of the Constitution, there shall be one Representative for every thirty thousand, until the number shall amount to one hundred, after which the proportion shall be so regulated by Congress, that there shall be not less than one hundred Representatives, nor less than one Representative for every forty thousand persons, until the number of Representatives shall amount to two hundred; after which the proportion shall be so regulated by Congress, that there shall not be less than two hundred Representatives, nor more than one Representative for every fifty thousand persons. Reference: The Constitution of the United States of America, House Document No. 124, U.S. Government Printing Office, 1967.)

My source does not explain why it was rejected, but I can imagine a number of criticisms. It is confusing, unnecessary, and ignores the ticklish problem of how to round off fractional representatives for a particular state. It appears to presume that the population will always increase--at least reverting to a smaller House following a plague or war would be forbidden. It is not clear whether its restrictions must be satisfied continuously, or only once a decade following the census.

My object, however, is to point out an unintended inconsistency. If we let n be the size of the House, p be the population, then $r = p/n$ is the number of people represented by each congressman. In algebraic notation, Article I provides: (1) For $n < 100$, $r = 30{,}000$; (2) for $100 \leq n < 200$, $r \leq min\ (40{,}000, p/100)$; (3) and for $n \geq 200$, $50{,}000 \leq r \leq p/200$.

But (2) can apply only if $p \leq 7{,}960{,}000$ whereas (3) requires $p = nr \geq 200 \cdot 50{,}000 = 10{,}000{,}000$! Imagine the dire measures that would be required to satisfy this article whenever our population edged into the forbidden range!

It is curious that the article was worded to restrict r (and implicitly p) as a function of n rather than limiting r as a function of p. Since the wrong variable is independent, we would be trapped into adjusting p to conform with this restrictive amendment. I suppose we could always deport a couple million settlers! (cf., the title of this note.)

Allen J. Schwenk
U.S. Naval Academy
Annapolis
Maryland 21402

1978 WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

A-1. Let A be any set of 20 distinct integers chosen from the arithmetic progression 1, 4, 7, ... , 100. Prove that there must be two distinct integers in A whose sum is 104.

A-2. Let a, b, p_1, p_2, ... , p_n be real numbers with $a \neq b$. Define $f(x) = (p_1-x)(p_2-x)(p_3-x)\ldots(p_n-x)$. Show that the determinant of

$$\begin{bmatrix} p_1 & a & a & a & \ldots & a & a \\ b & p_2 & a & a & \ldots & a & a \\ b & b & p_3 & a & \ldots & a & a \\ b & b & b & p_4 & \ldots & a & a \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ b & b & b & b & \ldots & p_{n-1} & a \\ b & b & b & b & \ldots & b & p_n \end{bmatrix}$$

equals $[bf(a)-af(b)]/(b-a)$.

A-3. Let $p(x) = 2 + 4x + 3x^2 + 5x^3 + 3x^4 + 4x^5 + 2x^6$. For k with $0 < k < 5$, define

$$I_k = \int_0^\infty \frac{x^k}{p(x)}\,dx.$$

For which k is I_k smallest?

A-4. A "bypass" operation on a set S is a mapping from $S \times S$ to S with the property $B(B(w,x), B(y,z)) = B(w,z)$ for all w, x, y, z in S.

(a) Prove that $B(a,b) = c$ implies $B(c,c) = c$ when B is a bypass.

(b) Prove that $B(a,b) = c$ implies $B(a,x) = B(c,x)$ for all x in S when B is a bypass.

(c) Construct a table for a bypass operation B on a finite set S with the following three properties:

(i) $B(x,x) = x$ for all x in S.

(ii) There exist d and e in S with $B(d,e) = d \neq e$.

(iii) There exist f and g in S with $B(f,g) \neq f$.

A-5. Let $0 < x_i < \pi$ for $i = 1, 2, \ldots, n$, and set $x = (x_1 + x_2 + \ldots + x_n)/n$. Prove that

$$\prod_{i=1}^{n} \frac{\sin x_i}{x_i} \leq \left(\frac{\sin x}{x}\right)^n$$

A-6. Let n distinct points in the plane be given. Prove that fewer than $2n^{3/2}$ pairs of them are unit distance apart.

B-1. Find the area of a convex octagon that is inscribed in a circle and has four consecutive sides of length 3 units and the remaining four sides of length 2 units. Give the answer in the form $r + s\sqrt{t}$ with r, s, and t positive integers.

B-2. Express

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{m^2n + mn^2 + 2mn}$$

as a rational number.

B-3. The sequence $\{Q_n(x)\}$ of polynomials is defined by

$$Q_1(x) = 1 + x, \quad Q_2(x) = 1 + 2x,$$

and, for $m \geq 1$, by

$$Q_{2m+1}(x) = Q_{2m}(x) + (m+1)xQ_{2m-1}(x),$$

$$Q_{2m+2}(x) = Q_{2m+1}(x) + (m+1)xQ_{2m}(x).$$

Let x_n be the largest real solution of $Q_n(x) = 0$. Prove that $\{x_n\}$ is an increasing sequence and that $\lim_{n\to\infty} x_n = 0$.

B-4. Prove that for every real number N, the equation

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4$$

has a solution for which x_1, x_2, x_3, x_4 are all integers larger than N.

B-5. Find the largest A for which there exists a polynomial $P(x) = Ax^4 + Bx^3 + Cx^2 + Dx + E$, with real coefficients, which satisfies $0 \leq P(x) \leq 1$ for $-1 \leq x \leq 1$.

B-6. Let p and n be positive integers. Suppose that the numbers $c_{h,k}$ (h=1, 2, ... , n; k=1, 2, ... , ph) satisfy $0 \leq c_{h,k} \leq 1$. Prove that $\Sigma(c_{h,k}/h)^2 \leq p\Sigma c_{h,k}$, where each summation is over all admissible ordered pairs (h,k).

HOUGHTON MIFFLIN UPDATE:

Four essential texts by Doris S. Stockton

University of Massachusetts, Amherst

Stockton texts feature a mastery approach, with lists of objectives, detailed explanations, a wealth of examples and exercises, marginal references to related exercises, self-scoring quizzes, and final tests. All have Instructor's Manuals.

ESSENTIAL ALGEBRA AND TRIGONOMETRY

588 pages • 1978
For students not necessarily preparing for calculus.

ESSENTIAL PRECALCULUS

598 pages • 1978

ESSENTIAL COLLEGE ALGEBRA

About 526 pages • Early 1979
Assumes intermediate algebra.

ESSENTIAL TRIGONOMETRY

About 350 pages • Early 1979
Intermediate algebra a prerequisite.

Keller / Zant

BASIC MATHEMATICS
Third Edition

M. Wiles Keller, Purdue University
James H. Zant, Oklahoma State University
About 576 pages • paper • perforated
Instructor's Annotated Edition • Early 1979

Keller and Zant's workbook-text provides an intuitive introduction to the concepts and techniques of arithmetic, algebra, and trigonometry. Emphasizes the building of problem-solving skills. Diagnostic tests help instructor and student identify areas for in-depth study, review, or omission. Abundant exercises for building computational skills.

Rector / Zwick

FINITE MATHEMATICS AND ITS APPLICATIONS

Robert E. Rector and Earl J. Zwick
both of Indiana State University
About 432 pages • Instructor's Manual
Early 1979

Primarily for students in business administration and the social sciences, Rector/Zwick introduces concepts through applications. Computer options in BASIC permit use of computer for tedious computations.

Aufmann / Barker

ARITHMETIC
An Applied Approach

Richard N. Aufmann and Vernon C. Barker
both of Palomar College
512 pages • paper • Instructor's Manual • 1978

Brett / Sentlowitz

ELEMENTARY ALGEBRA BY EXAMPLE

William Brett and Michael Sentlowitz
both of Rockland Community College
497 pages • paper • Instructor's Manual • 1977

Haldi

BASIC MATHEMATICS
Skills and Structure

John F. Haldi, Spokane Community College
397 pages • paper • Instructor's Manual • 1978

Hecht / Hecht

MODUMATH: Arithmetic

Miriam Hecht, Hunter College,
City University of New York
Caroline Hecht
598 pages • paper • Instructor's Manual • 1978

Lyng / Meconi / Zwick

APPLIED TECHNICAL MATHEMATICS

Merwin J. Lyng, Mayville State College
L. J. Meconi, University of Akron
Earl J. Zwick, Indiana State University
496 pages • Instructor's Manual • 1978

Anderson / Sclove

AN INTRODUCTION TO THE STATISTICAL ANALYSIS OF DATA

T. W. Anderson, Stanford University
Stanley L. Sclove
University of Illinois, Chicago Circle
704 pages • Solutions Manual • 1978

Comprehensive introduction that blends data analysis and statistical inference. Many examples, problems, and applications to social, biological, physical, and administrative sciences.

Daniel

APPLIED NONPARAMETRIC STATISTICS

Wayne W. Daniel, Georgia State University
503 pages • Instructor's Manual • 1978

Nonmathematical treatment emphasizing applications and methods for the student/researcher. Worked-out examples for each technique; exercises based on real data.

Christensen

STATISTICS STEP BY STEP

Howard B. Christensen
Brigham Young University
670 pages • paper • Instructor's Manual with Solutions • 1977

Daniel

INTRODUCTORY STATISTICS WITH APPLICATIONS

Wayne W. Daniel, Georgia State University
475 pages • Study Guide • Solutions Manual and Instructor's Guide • 1977

For adoption consideration, request examination copies from your regional Houghton Mifflin office.

Houghton Mifflin

Dallas, TX 75235 Geneva, IL 60134 Hopewell, NJ 08525 Palo Alto, CA 94304 Boston, MA 02107

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, N.W.
Washington, DC 20036

MATHEMATICS MAGAZINE VOL. 52, NO. 1, JANUARY 1979