INDEX

AUTHORS

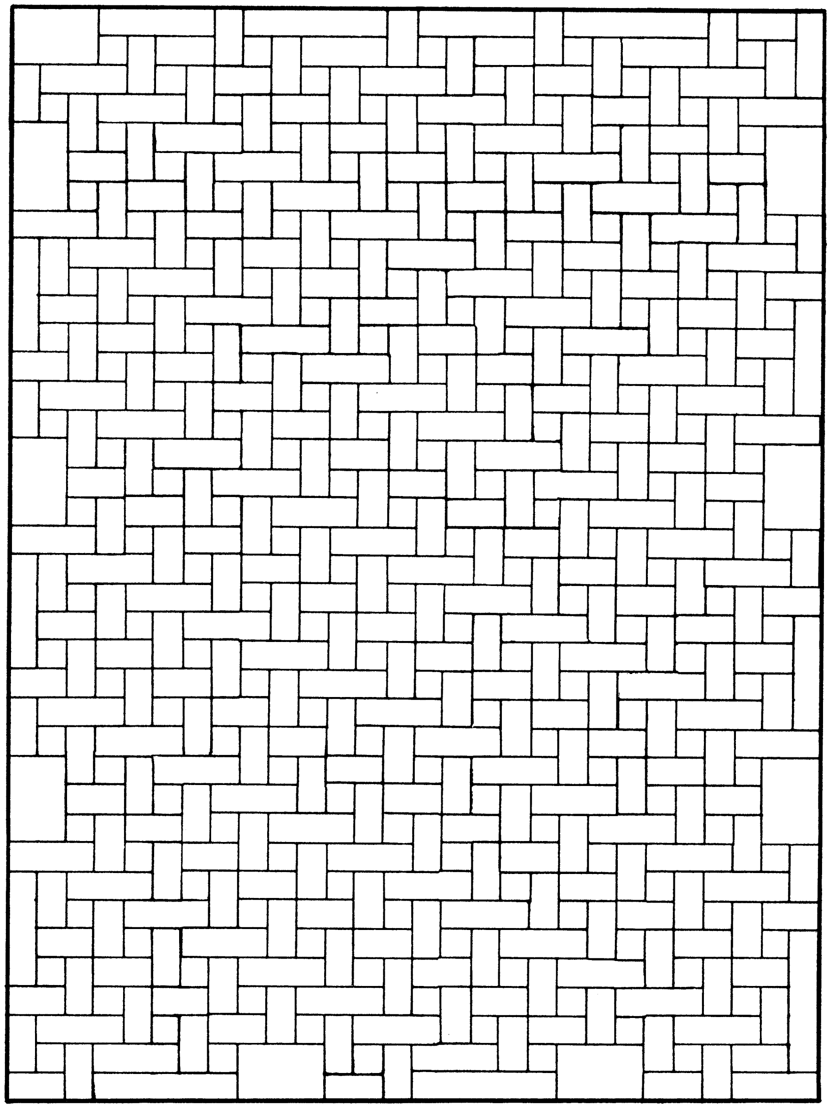
TITLES

PROBLEMS

P, S, *and* Q *refer to Proposals, Solutions and Quickies, respectively. Page numbers are given in parentheses. Thus,* P1158(300) *refers to proposal number* 1158 *which appears on page* 300.

NEWS

MATHEMATICS MAGAZINE



Vol. 55 No. 5
November 1982

ITERATIVE METHODS • TOOTH TABLES

TILING RECTANGLES • PERIODIC SUMS

MATHEMATICS MAGAZINE

Volume 55, Number 5, November 1982
(ISSN—0025-570X)

COVER: A simple tiling with average element area close to 11/6. See p. 290.

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

INDEX TO VOLUME 55

EDITORIAL POLICY

Mathematics Magazine is a journal of collegiate mathematics which aims to provide inviting, informal mathematical exposition of interest to undergraduate students. Manuscripts accepted for publication in the *Magazine* should be written in a clear and lively expository style and stocked with appropriate examples and graphics. Our advice to authors is: say something new in an appealing way or say something old in a refreshing way. The *Magazine* is not a research journal and so the style, quality, and level of articles submitted for publication should realistically permit their use to supplement undergraduate courses. The editor invites manuscripts that provide insight into the history and application of mathematics, that point out interrelationships between several branches of mathematics and that illustrate the fun of doing mathematics.

New manuscripts should be sent to: Doris Schattschneider, Editor, Mathematics Magazine, Moravian College, Bethlehem, Pennsylvania 18018. 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.

AUTHORS

Clifford Wagner ("A Generic Approach to Iterative Methods") grew up in Cincinnati, Ohio, and studied mathematics at the Universities of Cincinnati and Michigan before earning a Ph.D. in topology at the State University of New York at Albany. An expository lecture by Professor Gerald Rising (SUNY at Buffalo) sparked his interest in contraction mappings. Additional inspiration was derived from Donald Wagner and Ray Charles. Dr. Wagner is Associate Professor of Mathematics at The Capitol Campus, an upper-division and graduate campus of The Pennsylvania State University.

ILLUSTRATIONS

Paul Berge caught Brisben in the act, p. 285.

All other illustrations were provided by the authors.

Authors planning to submit manuscripts should read the full statement of editorial policy which appears in the News and Letters section of this *Magazine*, Vol. 54, pp. 44–45. Additional copies of the policy are available from the Editor.

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. The annual subscription price for Mathematics Magazine to an individual member of the Association is $8, included as part of the annual dues of $40. Students receive a 50% discount. Bulk subscriptions (5 or more copies to a single address) are available to colleges and universities for distribution to undergraduate students at a 35% discount. The library subscription price is $25.

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.

ARTICLES

A Generic Approach to Iterative Methods

Variations on a theme by Stefan Banach

CLIFFORD H. WAGNER
The Pennsylvania State University
Middletown, PA 17057

Iterative processes usually appear in the undergraduate curriculum as unrelated topics, each having a specific purpose and independent existence. Yet such topics as Newton's method, Picard's method for showing the existence and uniqueness of solutions to initial value problems, the Economic Cobweb Theorem for describing supply and demand equilibria, the convergence theorem for the long-run behavior of regular Markov chains, as well as many other iterative processes, are simply different manifestations of the general theme set forth in the **Contraction Mapping Principle**. The first abstract setting for this principle is credited to Stefan Banach [2],[17],[27], who showed that, under a very general hypothesis, all sequences generated by the repeated evaluation of a distance-decreasing function must converge to a unique fixed point. This convergence is the essence of an iterative technique which can be used in a variety of applications to find an approximate solution, to assert that a unique solution must exist, or to show that a given sequence converges to a known solution. Such applications of the Contraction Mapping Principle are the substance of this article.

A good example of a simple iterative process was shown to me by my five-year-old son Donald: Turn on a scientific calculator and repeatedly press the COS button. On a typical calculator, this algorithm computes iterated values of the calculator function COS(x), degree mode (a discrete rational function that approximates the continuous real function $f(x)=\cos(2\pi x/360)$), and displays successive terms of the sequence

$$\begin{aligned} x_0 &= 0, \\ x_1 &= 1, \\ x_2 &= .9998476952, \\ x_3 &= .9998477415, \\ x_4 &= .9998477415, \\ &\vdots \\ x_n &= .9998477415 \qquad \text{for } n \geqslant 3. \end{aligned}$$

This iterative sequence, starting at $x_0=0$ and generated by the function COS(x), converges to a fixed point $x^*=.9998477415$ for which COS(x^*) $=x^*$. In fact, you can choose any starting value you like and, *regardless of your choice for* x_0, the repeated evaluation of the COS function always produces the same fixed point x^* after at most four iterations. The Elementary Contraction Mapping Principle explains this phenomenon.

The Elementary Contraction Mapping Principle and simple fixed points

In general, a **fixed point** of the function $f(x)$ is a value x^* (in both the domain and range of the function) for which $f(x^*) = x^*$. A **contraction** (also called a **contraction mapping**) is a real-valued function $f(x)$ for which there is a constant K (called a **Lipschitz constant**) such that

$$0 \leqslant K < 1 \text{ and } |f(x) - f(y)| \leqslant K|x - y| \tag{1}$$

for all x and y in the domain.

For contractions defined on a closed interval I, the existence of a fixed point implies that the range is contained in I since (1) implies

$$|f(x) - x^*| = |f(x) - f(x^*)| \leqslant |x - x^*|$$

for each x in the interval. The converse of this proposition is not only valid but also more noteworthy.

THEOREM 1 (Elementary Contraction Mapping Principle). *If $f(x)$ is a contraction which maps a closed interval $[a, b]$ into itself, then*

(i) *there is a unique fixed point x^* in the interval $[a, b]$,*
(ii) *every iterative sequence generated by $f(x)$ converges to x^* (that is, x^* is the limit of every sequence $x_0, x_1, x_2, \ldots$ for which x_0 is in $[a, b]$ and $x_n = f(x_{n-1})$ for $n \geqslant 1$), and*
(iii) *if $\{x_n\}$ is any iterative sequence generated by $f(x)$, then*

$$|x_n - x^*| \leqslant \frac{K}{1-K}|x_n - x_{n-1}| \leqslant \frac{K^n}{1-K}|x_1 - x_0|, \tag{2}$$

where K is a Lipschitz constant for the contraction $f(x)$ and n is a positive integer.

The inequalities given in (2) can be used to estimate the error in approximating the fixed point x^* by the terms x_n of an iterative sequence generated by $f(x)$. The Elementary Contraction Mapping Principle (ECMP) is a special case of the General Contraction Mapping Principle (GCMP) given below. Many applications of the ECMP are facilitated by the following corollary of the Mean Value Theorem.

THEOREM 2 (Bounded Derivative Condition). *If the function $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b), and if there is a constant K such that $|f'(x)| \leqslant K < 1$ for $a < x < b$, then $f(x)$ is a contraction with Lipschitz constant K.*

We can apply Theorems 1 and 2 to see why every iterative sequence generated by the function $f(x) = \cos(2\pi x/360)$ discussed earlier converges very rapidly to the unique fixed point $x^* \doteq .9998477415$. We take the interval $I = [-1, 1]$, and the constant $K = 4 \times 10^{-4}$. This K is an acceptable Lipschitz constant for $f(x)$ because for $-1 \leqslant x \leqslant 1$, the maximum magnitude of $f'(x)$ is

$$|f'(1)| = \frac{\pi}{180} \sin\left(\frac{\pi}{180}\right) \doteq .000305.$$

The graph of $f(x)$ is symmetric about the y-axis and monotonically decreasing for $0 \leqslant x \leqslant 1$; therefore, regardless of the starting value x_0 in I, the first three terms of the sequence are bounded by the indicated constants:

$$\begin{aligned} -1 &\leqslant x_0 \leqslant 1, \\ f(1) &\leqslant x_1 \leqslant f(0), \\ f^2(0) &\leqslant x_2 \leqslant f^2(1), \text{ and} \\ f^3(1) &\leqslant x_3 \leqslant f^3(0). \end{aligned}$$

In this example, inequality (2) leads to

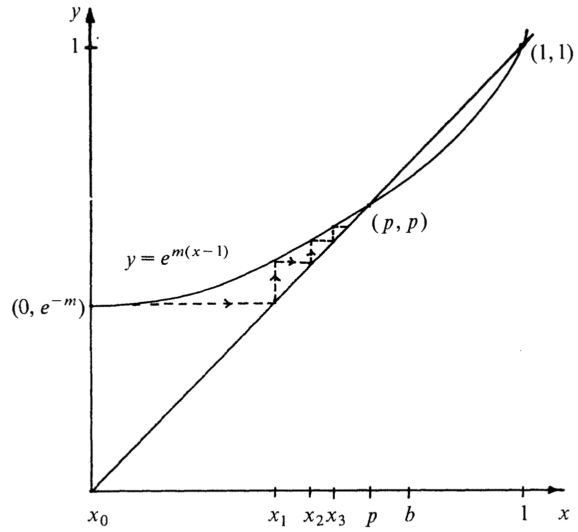


FIGURE 1. **The iterative sequence generated by $f(x)=e^{m(x-1)}$ when $m>1$.**

$$|x_3-x^*|\leqslant\frac{.0004}{.9996}|x_3-x_2|\leqslant\frac{1}{2499}\left[f^3(0)-f^2(0)\right]$$
$$\doteq\frac{.9998477415-.9998476952}{2499}\doteq1.85\times10^{-11}.$$

Thus x^* is approximated with an accuracy of at least ten significant places by the third term of every iterative sequence which starts in $I=[-1,1]$ and is generated by $\cos\left(\frac{2\pi x}{360}\right)$. Iterative sequences which start with an x_0 of magnitude greater than 1 must have x_1 in I, and then x_4 approximates x^* with the same accuracy. Similarly, the ten-place calculator function COS(x), degree mode, always reaches its fixed point by the fourth iteration.

Another simple fixed-point problem arises in R. A. Fisher's model [7] for the progeny of a single mutant gene in a cross-pollinated cereal plant. In Fisher's model, the probability for the ultimate extinction of the mutant genes is the smallest positive solution of

$$e^{m(x-1)}=x, \tag{3}$$

where m is the *average* number of offspring per mutant gene [**18**]. When $m\leqslant1$, ultimate extinction is certain and $x^*=1$ is the smallest solution of (3). However when $m>1$, the extinction probability is less than 1 and may be approximated by the iterative method of this section.

When successive generations are increasing in size ($m>1$), the function $f(x)=e^{m(x-1)}$ has two fixed points in the interval $[0,1]$. The extinction probability p is the smaller fixed point and the iterative sequence generated by $f(x)$ and starting at $x_0=0$ must converge to this p because there is a number b between p and 1 for which $f'(b)<1$, and then $f(x)$ is a contraction mapping of $[0,b]$ into itself with Lipschitz constant $K=f'(b)$ (see FIGURE 1). For example, Fisher considered the case $m=1.01$, a rate of increase which is only 1% greater than the rate for exact replacement (on average) of each generation, and estimated the extinction probability to be .9803 by computing the iterative sequence generated by $f(x)$.

Fixed-point problems also occur very naturally in calculus. For example, the critical points of the function $(\sin x)/x$, where $x>0$, are found by solving the equation $(x\cos x-\sin x)/x^2=0$, and this is equivalent to solving $\tan x=x$.

For positive values of x, the fixed points of $f(x)=\tan x$ occur near the discontinuities at $x=3\pi/2$, $5\pi/2$, etc. (See FIGURE 2). But the tangent function is not contractive in any interval,

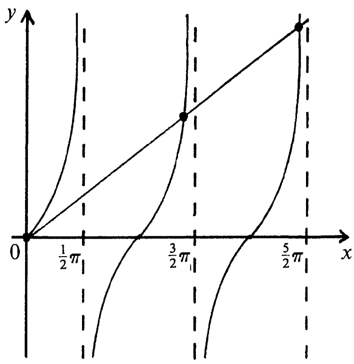


FIGURE 2. **Solutions of $x = \tan x$, with $x > 0$.**

and therefore one cannot apply the Contraction Mapping Principle with $f(x) = \tan x$. This problem is solved more easily by finding the roots of $\tan x - x = 0$. In the next section we describe techniques for finding roots of such equations; these too use the Contraction Mapping Principle in a special way.

Finding roots

A general method for approximating a root of $g(x) = 0$ is to find an auxiliary function $A(x)$ which never vanishes and a closed interval I in such a way that

$$f(x) = x - A(x)g(x)$$

is a contraction mapping from I into itself. Then all iterative sequences generated by $f(x)$ will converge to a value x^* that is both a fixed point of $f(x)$ and a zero of $g(x)$. Newton's method [**5**], [**10**] uses the auxiliary function $A(x) = 1/g'(x)$.

THEOREM 3 (Newton's Method). *If the equation $g(x) = 0$ has a root x^* somewhere in an open interval J where $g'(x)$ and $g''(x)$ are continuous, and where $g'(x)$ never vanishes, then J contains a closed subinterval $I = [a, b]$ such that*

(i) $a < x^* < b$,

(ii) *the function* $f(x) = x - \left[\dfrac{1}{g'(x)}\right] g(x)$ *is a contraction mapping from I into itself, and*

(iii) *the root x^* is the limit of every iterative sequence $x_0, x_1, x_2, \ldots$ which starts in I and is determined recursively by*

$$x_{n+1} = x_n - \left[\frac{1}{g'(x_n)}\right] g(x_n),$$

for $n \geqslant 0$.

Proof. The key here is that $f(x)$ satisfies the Bounded Derivative Condition (THEOREM 2) for some interval containing x^*.

The function $f(x)$ is defined for all x in J and its derivative,

$$f'(x) = \frac{g(x)g''(x)}{[g'(x)]^2},$$

is continuous throughout J with $f'(x^*) = 0$. Consequently, for *any* choice of K, where $0 < K < 1$, there is a closed subinterval $I = [a, b]$ for which $a < x^* < b$ and

$$|f'(x) - f'(x^*)| = |f'(x)| \leqslant K < 1 \text{ for all } x \text{ in } I.$$

Thus $f(x)$ is a contraction when its domain is restricted to I, and since I contains the fixed point, f must map I into itself. The ECMP (Theorem 1) is used to complete the proof.

Although every convergent iterative sequence based on the Contraction Mapping Principle must have at least linear convergence in the sense that a known bound for $|x_{n+1} - x^*|$ is proportional to $|x_n - x^*|$, it is well known that Newton's method actually generates quadratic convergence where the bound for each successive error is proportional to the *square* of the previous error. (Quadratic convergence occurs because $f'(x^*) = 0$ and a Taylor expansion of $f(x)$ about $x = x^*$ gives

$$x_{n+1} = f(x_n) = x^* + 0(x_n - x^*) + \frac{f''(c)}{2}(x_n - x^*)^2$$

for some c in the interval I.)

The next theorem illustrates a simpler iterative method using a constant auxiliary function, $A(x) = 1/M$. This method is useful when linear convergence is sufficient or when $g'(x)$ is not easily evaluated.

THEOREM 4. *If the equation* $g(x) = 0$ *has a root* x^* *in a closed interval* $I = [a, b]$ *where* $g(x)$ *is continuous on* I *and where* $0 < m \leqslant g'(x) \leqslant M$ *for* x *in* (a, b), *then* $f(x) = x - \frac{1}{M}g(x)$ *is a contraction mapping of* I *into itself and* x^* *is the limit of every iterative sequence generated by* $f(x)$ *and starting in* I.

Proof. The proof is similar to that of Theorem 3. In this case, a suitable Lipschitz constant is $K = 1 - m/M$ since

$$0 \leqslant f'(x) = 1 - \frac{g'(x)}{M} \leqslant K.$$

COROLLARY. *In the context of* Theorem 4, *if* $g'(x)$ *is always negative with* $0 < m \leqslant |g'(x)| \leqslant M$ *for* x *in* (a, b), *then* x^* *is the limit of every iterative sequence starting in* I *and generated by* $f(x) = x + \frac{1}{M}g(x)$.

Proof. Apply Theorem 4 to the function $[-g(x)]$.

We can now complete the solution of the fixed-point problem for the function $g(x) = \tan x - x$ and the interval $I = [4.4, 4.5]$ which contains the smallest positive root of $g(x) = 0$. The hypothesis of Theorem 4 is satisfied with $m = 9.504$ and $M = 21.6$ because $g'(x) = \tan^2(x)$ and thus

$$\min\{g'(x) : x \in I\} = \tan^2(4.4) \doteq 9.587 \geqslant m,$$

and

$$\max\{g'(x) : x \in I\} = \tan^2(4.5) \doteq 21.505 \leqslant M.$$

A Lipschitz constant for $f(x) = x - \frac{g(x)}{M}$ is

$$K = 1 - \frac{m}{M} = .56$$

Starting with $x_0 = 4.5$, the iterative sequence has the values

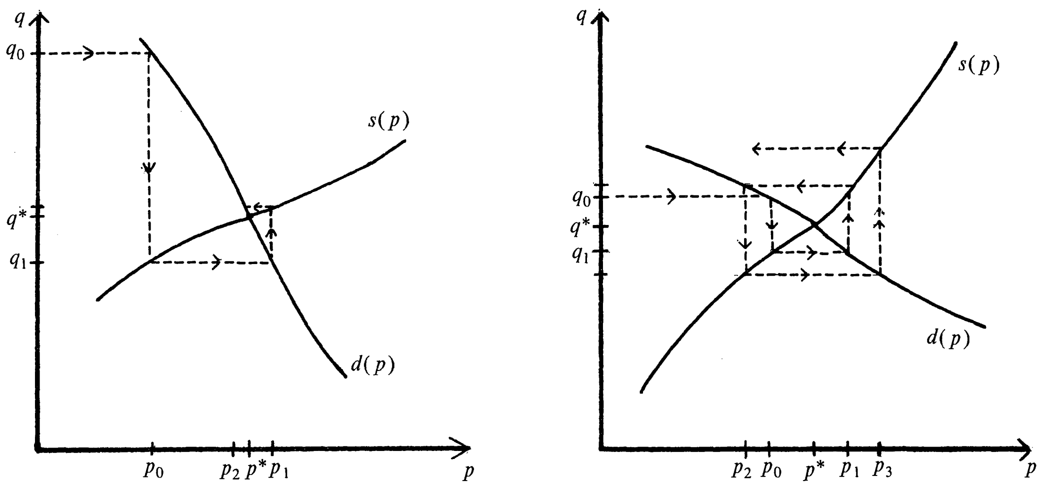
$$x_1 = f(4.5) \doteq 4.49364, \quad x_2 = f(x_1) \doteq 4.49342, \text{ etc.}$$

One might suspect that x_2 has an accuracy of at least three significant decimal places; this is in fact so, since

$$|x_2 - x^*| \leqslant \frac{K}{1-K}|x_2 - x_1| \doteq 2.8 \times 10^{-4} < 5 \times 10^{-4}.$$

Thus 4.49342 approximates the smallest positive zero of $\tan x - x$ which is also the smallest positive critical value of $(\sin x)/x$.

By comparison, Newton's method starting with $x_0 = 4.5$ gives $x_1 = 4.49361390$ and $x_2 = 4.49340966$, where x_2 is actually correct to six significant decimal places.



Converging to equilibrium **Diverging from equilibrium**

FIGURE 3. **Two economic cobwebs.**

The Cobweb Theorem

In mathematical economics, the market for a particular product often is modelled using a supply function $s(p)$ to represent the total quantity of the product that sellers are willing to supply at a given price level p, and a demand function $d(p)$ to represent the total quantity of the product that buyers are willing to purchase at a given price level p. It is customary to assume that these functions are strictly monotonic and intersect at some equilibrium point (p^*, q^*) where $q^* = d(p^*) = s(p^*)$.

The **cobweb model** [1], [21], [22] concerns a qualitative analysis of markets in which supply adjustments have a time lag and demand adjustments occur with no delay. For example, if an agricultural product is harvested once a year, then supply adjustments which producers make now in reaction to current price levels will affect the quantity to be offered for sale in next year's market. After this lag of one year, competitive market behavior will lead immediately to a new price level consistent with the buyer's demand curve.

Starting with price and quantity levels p_0 and $q_0 = d(p_0)$ the cobweb model determines subsequent levels by the relations

$$q_n = s(p_{n-1}),$$

and

$$p_n = d^{-1}(q_n)$$

for $n = 1, 2, 3, \ldots$. A relevant economics problem is to find conditions for the supply and demand functions to guarantee long term market stability in the sense that the sequence (p_n, q_n) converges to the equilibrium (p^*, q^*). The cobweb graphs of FIGURE 3 illustrate two possible situations.

Convergence occurs if, at each price level p, suppliers' reactions to small price changes, as measured by $|s'(p)|$, are smaller than buyers' reactions as measured by $|d'(p)|$.

THEOREM 5 (The Cobweb Theorem). *If the functions $d(p)$ and $s(p)$ have continuous nonvanishing derivatives throughout a closed interval I, if I contains a value p^* such that $d(p^*) = s(p^*)$ and if*

$$0 < |s'(p)| < |d'(p)|$$

for all p in I, then the cobweb sequences $p_n = d^{-1}(q_n)$ and $q_n = s(p_{n-1})$ converge to p^* and $q^* = d(p^*)$, respectively, for every initial price, p_0, in I.

Proof. Define $f(q)=s[d^{-1}(q)]$ for q in the closed interval $J=d(I)$. The derivative,

$$f'(q)=s'\left[d^{-1}(q)\right]\cdot\frac{1}{d'\left[d^{-1}(q)\right]},$$

is a continuous function in J, and because $|s'(p)|$ is always less than $|d'(p)|$ there is a constant K such that $0<|f'(q)|\leqslant K<1$ for all q in J. This means that $f(q)$ must be a contraction. As observed previously, a contraction mapping with a fixed point maps the domain into itself.

By the ECMP (Theorem 1), the iterative sequence generated by f, starting with $q_0=d(p_0)$ and satisfying

$$q_n=f(q_{n-1})=s\left[d^{-1}(q_{n-1})\right]=s(p_{n-1})$$

for $n\geqslant 1$, converges to the unique fixed point q^*. And since d^{-1} must be continuous, the sequence, $p_0=d^{-1}(q_0)$, $p_1=d^{-1}(q_1)$, etc., must converge to $p^*=d^{-1}(q^*)$.

The General Contraction Mapping Principle

So far, the Elementary Contraction Mapping Principle has been applied to contractions defined on closed intervals. In the more general setting of contractions on complete metric spaces, the General Contraction Mapping Principle (GCMP) brings unity to an even greater variety of iterative methods.

Before stating the GCMP, we recall several definitions. Suppose X is a metric space (with metric function d) and $x_0, x_1, x_2, \ldots$ is a sequence in X; then $\{x_n\colon n=0,1,2,\ldots\}$ is a **Cauchy sequence** if for each $\varepsilon>0$ there is an integer N such that

$$d(x_n, x_m)<\varepsilon \text{ whenever } n>N \text{ and } m>N.$$

The metric space X is **complete** (with respect to the metric d) if every Cauchy sequence in X converges to a limit (which must be unique). (For example, a closed interval in $\mathbb{R}$ with the usual metric, $d(x,y)=|x-y|$, is a complete metric space.) If the function F maps X into itself and if there is a Lipschitz constant K such that

$$0\leqslant K<1 \text{ and } d(F(x),F(y))\leqslant Kd(x,y) \text{ for all } x \text{ and } y \text{ in } X,$$

then F is a **contraction**.

The General Contraction Mapping Principle asserts that a contraction from a complete metric space into itself has a unique fixed point which is the limit of every iterative sequence generated by the contraction. The metric is needed for a general definition of contraction and the completeness property assures the convergence of the iterative sequences.

THEOREM 6 (General Contraction Mapping Principle). *If X is a complete metric space and F is a contraction from X into itself, then*

(i) *there is a unique fixed point x^*, such that $F(x^*)=x^*$, and*
(ii) *for every starting value x_0 in X, the iterative sequence $x_0, x_1, x_2, \ldots$, defined by $x_n=F(x_{n-1})$ for $n\geqslant 1$, converges to x^* with*

$$d(x_n,x^*)\leqslant\frac{K}{1-K}d(x_n,x_{n-1})\leqslant\frac{K^n}{1-K}d(x_1,x_0),$$

for $n=1,2,3,\ldots$.

A useful corollary of the GCMP concerns functions $F(x)$ for which some iterated power $F^{(k)}(x)$ is a contraction, where $F^{(1)}(x)=F(x)$, and $F^{(n+1)}(x)=F[F^{(n)}(x)]$ for $n=1,2,\ldots$.

COROLLARY. *If X is a complete metric space, if F is a function from X into X, and if $F^{(k)}$ is a contraction for some positive integer k, then F has a unique fixed point x^* and every sequence $x_0, x_1, x_2, \ldots$, defined by $x_n=F(x_{n-1})$ for $n\geqslant 1$, converges to x^*.*

Since the proofs of the GCMP and its corollary are somewhat technical, we omit them here and include them in the last section for interested readers. The ECMP (Theorem 1) is a specific interpretation of the GCMP for the case where X is a closed interval of $\mathbb{R}$.

The iterative methods of the next three sections involve contractions defined on: (1) the set of m-dimensional probability vectors with the box metric, (2) the set $\mathbb{R}^m$ of m-dimensional real vectors with the box metric, and (3) the set of continuous functions from one closed interval to another with the sup metric. That each of these is a complete metric space can be proven directly (using the completeness property of the real numbers) or indirectly within the general theory of metric spaces [12].

Fixed points for Markov chains

A finite Markov chain with m states is determined by specifying an $m \times m$ transition matrix $\mathbf{P} = (p_{ij})$ where p_{ij} is the probability of passing from state i to state j during any transition period (and where $\Sigma_{j=1}^{m} p_{ij} = 1$ for each integer i from 1 to m).

To illustrate, suppose the annual marketing goal for the automobile insurance division of Farm States Insurance Company is to retain 80% of its policyholders and to capture 10% of the automobile policies of other companies. The auto insurance market can be modelled as a Markov chain with two states: State 1, the customers of Farm States, and State 2, all other customers. Let the initial probability vector $\mathbf{x}_0 = (.15 \ .85)$ indicate that Farm States currently has a 15% market share. Let the transition matrix be

$$\mathbf{P} = \begin{bmatrix} .80 & .20 \\ .10 & .90 \end{bmatrix},$$

and let the auto insurance market after n years be determined by the probability vector $\mathbf{x}_n = \mathbf{x}_{n-1}\mathbf{P} = \mathbf{x}_0\mathbf{P}^n$. For example, the market after one year is given by $\mathbf{x}_1 = (.205 \ .795)$ and after two years by $\mathbf{x}_2 = (.2435 \ .7565)$. What is the long-term market share projection for Farm States?

A remarkable property of this hypothetical auto insurance example, as well as other so-called finite regular Markov chains, is that the ultimate convergence of the chain (that is, the limit of $\mathbf{x}_0\mathbf{P}^n$ as $n \to \infty$) is totally independent of the initial probability vector $\mathbf{x}_0$. This claim is the essence of Theorem 7, for which we need a few definitions. A Markov chain is **regular** if, for some integer k, every state is reachable from every other state in exactly k transitions, or equivalently if the kth power of the transition matrix has only positive entries [13], [20]. For any finite Markov chain having m states, the associated **set of probability vectors** is the set $\mathbf{X}$ defined by

$$\mathbf{X} = \left\{ \mathbf{x} = (x_1, x_2, \ldots, x_m) \colon \sum_{i=1}^{m} x_i = 1 \text{ and } x_i \geqslant 0 \text{ for each } i \right\}.$$

The **distance** between probability vectors $\mathbf{x}$ and $\mathbf{y}$ may be given by the **box metric.**

$$d(\mathbf{x},\mathbf{y}) = \max\{|x_i - y_i| : i = 1,2,\ldots,m\}.$$

The set of probability vectors is a complete metric space with respect to the box metric.

THEOREM 7. *Every finite regular Markov chain has a unique probability vector* $\mathbf{x}^*$ *which is a fixed point of its transition matrix* $\mathbf{P}$, *and all iterative sequences* $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots,$ *determined by* $\mathbf{x}_n = \mathbf{x}_{n-1}\mathbf{P}$ *for* $n \geqslant 1$, *converge to* $\mathbf{x}^*$.

Proof. Let $\mathbf{X}$ denote the associated set of probability vectors for the regular Markov chain with transition matrix $\mathbf{P}$. Define $f\colon \mathbf{X} \to \mathbf{X}$ by $f(\mathbf{x}) = \mathbf{x}\mathbf{P}$.

First suppose that $\mathbf{P}$ has no zero entries. A suitable Lipschitz constant is $K = 1 - \varepsilon$, where ε is the smallest entry of $\mathbf{P}$.

For arbitrary probability vectors $\mathbf{x}$ and $\mathbf{y}$ we have $m_0 = \min(x_i - y_i) \leqslant 0$, $M_0 = \max(x_i - y_i) \geqslant 0$, and $d(\mathbf{x},\mathbf{y}) = \max\{-m_0, M_0\}$. And letting z_i represent the ith coordinate of $\mathbf{xP} - \mathbf{yP}$, we have

$$d(\mathbf{xP},\mathbf{yP}) = \max|z_i|. \tag{4}$$

As shown by Kemeny and Snell [**13**, pp. 69–70],

$$z_i \leqslant M_0 - \varepsilon(M_0 - m_0) = M_0(1-\varepsilon) + \varepsilon m_0 = KM_0 + \varepsilon m_0,$$

and similarly,

$$z_i \geqslant Km_0 + \varepsilon M_0$$

for each i. But $\varepsilon m_0 \leqslant 0$ and $\varepsilon M_0 \geqslant 0$, so that

$$Km_0 \leqslant z_i \leqslant KM_0$$

for each i. From (4) then,

$$d(\mathbf{xP},\mathbf{yP}) \leqslant K\max\{-m_0, M_0\} = Kd(\mathbf{x},\mathbf{y}).$$

For this case, the desired result now follows from the General Contraction Mapping Principle.

For the case where $\mathbf{P}$ has zero entries but $\mathbf{P}^k$ does not, the function $g(\mathbf{x}) = f^{(k)}(\mathbf{x}) = \mathbf{xP}^k$ is a contraction, and the desired result follows from the Corollary of the General Contraction Mapping Principle.

Let's determine the ultimate market share for Farm States Insurance Company in the preceding example. Since the Markov chain is finite and regular, all iterative sequences $\mathbf{x}_0, \mathbf{x}_0\mathbf{P}, \mathbf{x}_0\mathbf{P}^2, \ldots$ converge to a unique probability vector $\mathbf{x}^*$. Because $\mathbf{x}^*$ must satisfy $\mathbf{x} = \mathbf{xP}$, it can be found directly by solving the following redundant system of equations:

$$\begin{aligned} .8x_1 + .1x_2 &= x_1, \\ .2x_1 + .9x_2 &= x_2, \\ x_1 + x_2 &= 1. \end{aligned}$$

The unique solution is $\mathbf{x}^* = (1/3 \;\; 2/3)$, and the Farm States market share converges to $33\frac{1}{3}$ percent, regardless of the initial vector $\mathbf{x}_0$. Also note that $\mathbf{x}^*$ can be approximated using any iterative sequence $\mathbf{x}_0, \mathbf{x}_0\mathbf{P}, \mathbf{x}_0\mathbf{P}^2, \ldots, \mathbf{x}_0\mathbf{P}^n$.

Systems of linear equations

In standard matrix form, a system of m linear equations with m unknowns is given by $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is an $m \times m$ matrix of constants, and $\mathbf{x}$ and $\mathbf{b}$ are $m \times 1$ column vectors of variables and constants. For example, the system

$$\begin{bmatrix} 4 & -1 & 0 & 0 & 0 & -1 \\ -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & -1 & 4 & -1 & 0 & 0 \\ 0 & 0 & -1 & 4 & -1 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 \\ -1 & 0 & 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \\ 20 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

corresponds to a particular finite difference approximation for the following heat distribution problem. A long metal bar with rectangular cross sections of dimensions 3 by 4 is half submerged in water. The water has a constant temperature of 0° and the atmosphere above the water has a constant temperature of 10°. When the distribution of heat within the bar reaches equilibrium, what are the equilibrium temperatures at various points near the middle of the bar?

The finite difference method which leads to the system of equations (5) uses a mesh of sixteen evenly spaced points in a cross section near the middle, with the assumption that cross sections nearby have essentially the same temperature distribution at equilibrium.

Each of the six unknown equilibrium temperatures may be approximated by the average of the equilibrium temperatures at the four closest mesh points in FIGURE 4. For example,

$$x_1 = (x_2 + x_6 + 10 + 10)/4$$

is equivalent to the first equation represented in (5).

The given system of equations is analogous to a continuous Dirichlet problem involving an

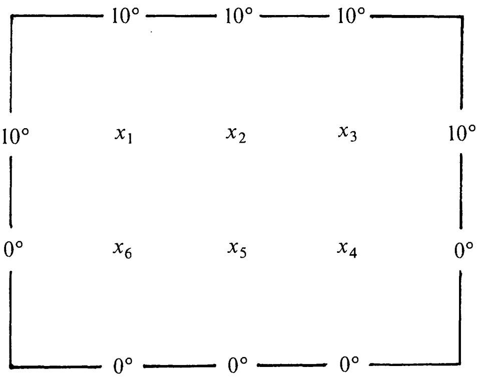


FIGURE 4. **Equilibrium temperatures in a cross section.**

elliptic partial differential equation with boundary conditions [**8**], [**25**]. Although the corresponding continuous solution is obtainable in the case of rectangular regions [**3**], finite difference solutions have wider application.

A common class of linear systems $\mathbf{Ax} = \mathbf{b}$ are those, such as the given system (5), for which the coefficient matrix $\mathbf{A}$ is **strictly diagonally dominant,** which means that

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{m} |a_{ij}|$$

for $i = 1, 2, \ldots, m$. For such systems, there is always exactly one solution. One way of approximating that solution is to generate an iterative sequence using $f(\mathbf{x}) = \mathbf{D}^{-1}[\mathbf{b} - (\mathbf{A} - \mathbf{D})\mathbf{x}]$, where $\mathbf{D}$ is the diagonal matrix

$$\mathbf{D} = \begin{bmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{mm} \end{bmatrix},$$

and where $\mathbf{x}$ is a column vector in $\mathbb{R}^m$. If we measure the distance between two vectors by the box metric, then the set $\mathbb{R}^m$ becomes a complete metric space.

The function $f(\mathbf{x})$ is not a magical discovery, for the matrix equation $\mathbf{Ax} = \mathbf{b}$ is equivalent to $[(\mathbf{A} - \mathbf{D}) + \mathbf{D}]\mathbf{x} = \mathbf{b}$, from which it follows that a solution $\mathbf{x}$ must satisfy

$$\mathbf{Dx} = \mathbf{b} - (\mathbf{A} - \mathbf{D})\mathbf{x}, \text{ or } \mathbf{x} = \mathbf{D}^{-1}[\mathbf{b} - (\mathbf{A} - \mathbf{D})\mathbf{x}] = f(\mathbf{x}).$$

THEOREM 8 (Jacobi Method). *If* $\mathbf{A}$ *is an* $m \times m$ *strictly diagonally dominant matrix*, $\mathbf{D}$ *is the diagonal matrix* (a_{ii}), *and* $\mathbf{b}$ *is a fixed vector in* $\mathbb{R}^m$, *then*

(i) *the function* $f(\mathbf{x}) = \mathbf{D}^{-1}[\mathbf{b} - \mathbf{Ax} + \mathbf{Dx}]$ *is a contraction from* $\mathbb{R}^m$ *into itself with Lipschitz constant*

$$K = \max\left\{ \frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^{m} |a_{ij}| : i = 1, 2, \ldots, m \right\},$$

(ii) *the unique fixed point* $\mathbf{x}^*$ *of* $f(\mathbf{x})$ *is also a unique solution of the system* $\mathbf{Ax} = \mathbf{b}$, *and*

(iii) *every iterative sequence generated by* $f(\mathbf{x})$ *must converge to* $\mathbf{x}^*$.

Proof. The ith coordinate of $f(\mathbf{x}) - f(\mathbf{y})$ has norm

$$\left|\frac{1}{a_{ii}}\left(b_i - \sum_j a_{ij}x_j + a_{ii}x_i\right) - \frac{1}{a_{ii}}\left(b_i - \sum_j a_{ij}y_j + a_{ii}y_i\right)\right|$$

$$= \frac{1}{|a_{ii}|}\left|\sum_{\substack{j=1\\ j\neq i}}^{m} a_{ij}(x_j - y_j)\right| \leqslant \frac{1}{|a_{ii}|}\sum_{\substack{j=1\\ j\neq i}}^{m} |a_{ij}|\, d(\mathbf{x},\mathbf{y}) \leqslant Kd(\mathbf{x},\mathbf{y}).$$

Therefore $d[f(\mathbf{x}), f(\mathbf{y})] \leqslant Kd(\mathbf{x},\mathbf{y})$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^m$, confirming part (i). The other conclusions follow from the Contraction Mapping Principle and our earlier observation that fixed points of $f(\mathbf{x})$ coincide with solutions of $\mathbf{Ax} = \mathbf{b}$.

A similar, but generally more efficient, method for solving many systems of equations is the Gauss-Seidel method based upon iteratively evaluating the function

$$g(\mathbf{x}) = (\mathbf{D} + \mathbf{L})^{-1}(\mathbf{b} - \mathbf{Ux}), \tag{6}$$

where $\mathbf{L} = (\ell_{ij})$ is the **lower part** of $\mathbf{A}$, defined by

$$\ell_{ij} = \begin{cases} a_{ij} & \text{if } i > j, \\ 0 & \text{if } i \leqslant j, \end{cases}$$

and $\mathbf{U}$ is the **upper part** of $\mathbf{A}$, defined by $\mathbf{U} = \mathbf{A} - (\mathbf{D} + \mathbf{L})$. Since $\mathbf{A} = (\mathbf{D} + \mathbf{L}) + \mathbf{U}$, the system $\mathbf{Ax} = \mathbf{b}$ is also equivalent to $(\mathbf{D} + \mathbf{L})\mathbf{x} = \mathbf{b} - \mathbf{Ux}$ or $\mathbf{x} = (\mathbf{D} + \mathbf{L})^{-1}(\mathbf{b} - \mathbf{Ux}) = g(\mathbf{x})$.

THEOREM 9 (Gauss-Seidel Method). *If* $\mathbf{A}$ *is strictly diagonally dominant and* $g(\mathbf{x})$ *is the mapping* (6), *then* $g(\mathbf{x})$ *is a contraction from* $\mathbb{R}^m$ *into itself with the same Lipschitz constant as the Jacobi function* $f(\mathbf{x})$ (Theorem 8, (i)). *Moreover, every iterative sequence generated by* $g(\mathbf{x})$ *converges to the unique solution of* $\mathbf{Ax} = \mathbf{b}$.

Proof. Let z_i be the ith coordinate of $g(\mathbf{x}) - g(\mathbf{y})$. We only need to show that each $|z_i|$ is bounded above by $Kd(\mathbf{x},\mathbf{y})$, since then we have $d[g(\mathbf{x}), g(\mathbf{y})] = \max|z_i| \leqslant Kd(\mathbf{x},\mathbf{y})$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^m$. To show that $|z_i| \leqslant Kd(\mathbf{x},\mathbf{y})$, we note that the individual scalar equations represented by the matrix equation

$$(\mathbf{D} + \mathbf{L})[g(\mathbf{x}) - g(\mathbf{y})] = \mathbf{U}(\mathbf{y} - \mathbf{x})$$

are each of the form

$$\sum_{j=1}^{i} a_{ij}z_j = \sum_{j=i+1}^{m} a_{ij}(y_j - x_j), \text{ where } 1 \leqslant i \leqslant m.$$

Solving for the term with z_i gives

$$a_{ii}z_i = \sum_{j=1}^{i-1} -a_{ij}z_j + \sum_{j=i+1}^{m} a_{ij}(y_j - x_j). \tag{7}$$

We now proceed by induction on i. For $i = 1$, equation (7) leads to

$$|z_1| \leqslant \frac{1}{|a_{11}|}\sum_{j=2}^{m} |a_{1j}||y_j - x_j| \leqslant Kd(\mathbf{x},\mathbf{y}).$$

If we assume that $|z_i| \leqslant Kd(\mathbf{x},\mathbf{y}) \leqslant d(\mathbf{x},\mathbf{y})$ for $i = 1, 2, \ldots, k-1$, then (7) leads to

$$|a_{kk}||z_k| \leqslant \sum_{j=1}^{k-1} |a_{kj}|Kd(\mathbf{x},\mathbf{y}) + \sum_{j=k+1}^{m} |a_{kj}|d(\mathbf{x},\mathbf{y})$$

$$\leqslant d(\mathbf{x},\mathbf{y}) \sum_{\substack{j=1\\ j\neq k}}^{m} |a_{kj}|.$$

From the definition of K, it follows that $|z_k| \leqslant Kd(\mathbf{x},\mathbf{y})$.

The actual two-place solution of (5) is $x_1 = x_3 = 7.39$, $x_2 = 6.96$, $x_4 = x_6 = 2.61$, and $x_5 = 3.04$. If the iterative sequences start with an initial vector having $x_i = 0$ for each i, then the Jacobi sequence attains an accuracy of two significant decimal places in fifteen steps while the Gauss-Seidel requires only eight steps. Indeed the latter method usually converges more quickly, although this is not indicated by Theorems 8 and 9 where the same Lipschitz constant is given for both $f(\mathbf{x})$ and $g(\mathbf{x})$.

The methods just outlined, as well as the SOR method which generalizes the Gauss-Seidel method, are described more fully in many linear algebra and numerical analysis texts (e.g., **[11]**, **[16]**, **[19]**, **[24]**).

Ordinary differential equations

The standard method for proving the existence and uniqueness of solutions to ordinary differential equations is due to Émile Picard **[3]**, **[14]**, **[23]**. Picard's theorem involves a contraction on the set X of all continuous functions from I into J, where I and J are specified closed intervals of $\mathbb{R}$. If y and z are functions in X, their distance can be given by the **sup metric**

$$d(y,z) = \max\{|y(x) - z(x)|: x \in I\}.$$

The set X is a complete metric space with respect to this metric.

THEOREM 10 (Picard's Existence-Uniqueness Theorem). *If the function $f(x, y)$ and its partial derivative $\partial f/\partial y$ are continuous at all points in a nontrivial rectangle $R = \{(x, y): |x - x_0| \leqslant a$ and $|y - y_0| \leqslant b\}$, then for some nontrivial interval $I = [x_0 - h, x_0 + h]$, there exists a unique real valued function y^* which has domain I and solves the initial value problem*

$$\frac{dy}{dx} = f(x, y) \text{ for all } x \text{ in } I, \text{ and}$$

$$y(x_0) = y_0.$$

Proof. Pick any K in $(0, 1)$; K will be a Lipschitz constant for a contraction which we construct. Let

$$M_1 = \max\{|f(x, y)|: (x, y) \text{ is in } R\},$$

and

$$M_2 = \max\{|\partial f/\partial y|: (x, y) \text{ is in } R\}.$$

Choose an $h > 0$ small enough to satisfy

$$h \leqslant a, \quad hM_1 \leqslant b, \quad \text{and} \quad hM_2 \leqslant K.$$

This "pick and choose" process determines the intervals

$$I = [x_0 - h, x_0 + h], \text{ and } J = [y_0 - b, y_0 + b],$$

and the set of functions $X = \{y: y \text{ is a continuous function from } I \text{ into } J\}$.

Define a mapping from X into itself by $F(y) = F_y$ where for each function y in X, F_y is the function defined by

$$F_y(x) = y_0 + \int_{x_0}^{x} f[t, y(t)]\, dt, x \in I.$$

For each y in X, the restrictions on h imply that $F(y)$ is also a continuous function from I into J, and therefore F does indeed map X into itself. Moreover, we can show that the function F is a contraction with respect to the metric for X. If y and z are functions in X, then for each t in I there is some c between $y(t)$ and $z(t)$ for which

$$f[t, y(t)] - f[t, z(t)] = \frac{\partial f(t, c)}{\partial y}[y(t) - z(t)],$$

and therefore,

$$|f[t,y(t)]-f[t,z(t)]| \leqslant M_2|y(t)-z(t)| \leqslant M_2 d(y,z).$$

Bounds for the distance between F_y and F_z involve maxima with respect to all x in I, and

$$\begin{aligned} d[F(y),F(z)] &= \max_x |F_y(x)-F_z(x)| \\ &= \max_x \left| \int_{x_0}^{x} (f[t,y(t)]-f[t,z(t)])\,dt \right| \\ &\leqslant \max_x \left| \int_{x_0}^{x} M_2 d(y,z)\,dt \right| \\ &= M_2 d(y,z)h \leqslant Kd(y,z). \end{aligned}$$

Consequently, $F(y)$ has a unique fixed point y^* and every iterative sequence generated by $F(y)$ converges to y^*. Since $F(y)$ is defined in such a way that its fixed points coincide exactly with solutions of the given initial value problem, the theorem follows (use the Fundamental Theorem of Calculus and its converse).

The restriction in Theorem 10 of the solution function $y^*(x)$ to some interval $I=[x_0-h, x_0+h]$ is not always binding. For if $h<a$ and $|y^*(x_0+h)-y_0|<b$, then $y^*(x)$ can be extended (again uniquely) to a larger interval by solving the given differential equation with the new initial condition $y(x_0+h)=y^*(x_0+h)$, and then joining y^* together with the new right-hand solution. Repeating this extension process as needed on both the right- and left-hand sides leads to a unique function which not only solves the original initial value problem but also possesses a graph that passes from one boundary edge of R to another.

Conclusion

If the fixed point of a contraction is approximated by an iterative sequence and you wish to improve the accuracy of the approximation, follow the advice of a popular rock and roll lyric: [repeat the iterative process] *one more time* [**6**]. If you wish to view iterative techniques with a wide-angle lens, use the Contraction Mapping Principle. And if you want to learn how two versions of this principle are proved, read the following two proofs.

THEOREM 6 (General Contraction Mapping Theorem [**9**],[**14**],[**17**],[**26**]). *If X is a complete metric space and F is a contraction mapping from X into itself with Lipschitz constant K such that $0 \leqslant K < 1$ and*

$$d(F(x),F(y)) \leqslant Kd(x,y) \text{ for all } x \text{ and } y \text{ in } X,$$

then

(i) *there is a unique fixed point x^*, such that $F(x^*)=x^*$, and*
(ii) *for every starting value x_0 in X, the iterative sequence $x_0, x_1, x_2, \ldots$, defined by $x_n = F(x_{n-1})$ for $n \geqslant 1$, converges to x^* with*

$$d(x_n,x^*) \leqslant \frac{K}{1-K} d(x_n,x_{n-1}) \leqslant \frac{K^n}{1-K} d(x_1,x_0),$$

for $n = 1,2,3,\ldots$.

Proof. First consider a single iterative sequence.
(a) For all pairs of positive integers i and j with $i<j$,

$$\begin{aligned} d(x_j,x_{j-1}) &= d\left[F(x_{j-1}),F(x_{j-2})\right] \leqslant Kd(x_{j-1},x_{j-2}) \\ &\leqslant K^2 d(x_{j-2},x_{j-3}) \leqslant \cdots \leqslant K^{j-i} d(x_i,x_{i-1}), \end{aligned}$$

that is,

$$d(x_j,x_{j-1}) \leqslant K^{j-i} d(x_i,x_{i-1}). \tag{8}$$

Repeated use of the triangle inequality and inequality (8) leads to a generalization of (8): for positive integers n and m with $n < m$,

$$\begin{aligned} d(x_m, x_n) &\leqslant d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+2}, x_{n+1}) + d(x_{n+1}, x_n) \\ &\leqslant K^{m-n} d(x_n, x_{n-1}) + K^{m-n-1} d(x_n, x_{n-1}) \\ &\quad + \cdots + K^2 d(x_n, x_{n-1}) + K d(x_n, x_{n-1}) \\ &= \frac{K - K^{m-n+1}}{1-K} d(x_n, x_{n-1}) \leqslant \frac{K}{1-K} d(x_n, x_{n-1}), \end{aligned}$$

that is,

$$d(x_m, x_n) \leqslant \frac{K}{1-K} d(x_n, x_{n-1}). \tag{9}$$

Combining (8) and (9), we have

$$d(x_m, x_n) \leqslant \frac{K}{1-K} d(x_n, x_{n-1}) \leqslant \frac{K^n}{1-K} d(x_1, x_0). \tag{10}$$

For n sufficiently large, the bound $\frac{K^n}{1-K} d(x_1, x_0)$ is arbitrarily close to 0, so the iterative sequence is Cauchy and has a unique limit x^*.

(b) The limit x^* must be a fixed point of F because all contractions are continuous, and therefore

$$F(x^*) = F\Big(\lim_{n\to\infty} x_n\Big) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} x_{n+1} = x^*.$$

(c) For any positive integers n and m with $n < m$, it follows from (10) that

$$d(x^*, x_n) \leqslant d(x^*, x_m) + d(x_m, x_n) \leqslant d(x^*, x_m) + \frac{K}{1-K} d(x_n, x_{n-1}). \tag{11}$$

But $\lim_{m\to\infty} d(x^*, x_m) = 0$, so (11) implies that

$$d(x^*, x_n) \leqslant \frac{K}{1-K} d(x_n, x_{n-1}). \tag{12}$$

The desired bounds for $d(x^*, x_n)$ are given by (12) and (10).

Now consider all possible iterative sequences. As just shown, each such sequence converges to a fixed point. But a contraction can have at most one fixed point since $x^* = F(x^*)$ and $y^* = F(y^*)$ imply that $d(x^*, y^*) \leqslant K d(x^*, y^*)$. Consequently, the unique fixed point x^* is the limit of every iterative sequence.

COROLLARY. (See [4], [15], [17].) *If X is a complete metric space, if F is a function from X into X, and if $F^{(k)}(x)$ is a contraction for some positive integer k, then F has a unique fixed point x^* and every sequence $x_0, x_1, x_2, \ldots,$ defined by $x_n = F(x_{n-1})$ for $n \geqslant 1$, converges to x^*.*

Proof. By Theorem 6 the function $G(x) = F^{(k)}(x)$ has a unique fixed point x^* which is the limit of each of the following sequences:

$$x_0, x_k = G(x_0), x_{2k} = G(x_k), x_{3k}, \ldots$$

$$x_1, x_{k+1} = G(x_1), x_{2k+1} = G(x_{k+1}), x_{3k+1}, \ldots$$

$$x_2, x_{k+2}, x_{2k+2}, x_{3k+2}, \ldots$$

$$\vdots$$

$$x_{k-1}, x_{2k-1}, x_{3k-1}, \ldots .$$

These sequences are subsequences of $x_0, x_1, x_2, \ldots$ which must then also converge to x^*. That is, every iterative sequence generated by F converges to the unique fixed point of G.

The functions G and F share the same set of fixed points, for if $y^* = F(y^*)$, then $G(y^*) = F^{(k)}(y^*) = y^*$; and if $x^* = G(x^*)$, then

$$F(x^*) = F[G(x^*)] = F^{(k+1)}(x^*) = G[F(x^*)].$$

But G has exactly one fixed point and therefore $F(x^*) = x^*$, and so x^* is also a unique fixed point of F.

Bounds to indicate the rate of convergence in the corollary are given by the inequalities

$$d(x_{nk}, x^*) \leqslant \frac{K}{1-K} d(x_{nk}, x_{nk-k}) \leqslant \frac{K^n}{1-K} d(x_k, x_0).$$

References

[1] R. G. D. Allen, Mathematical Economics, 2nd ed., St. Martin's Press, New York, 1966.

[2] S. Banach, Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales, Fund. Math., 3 (1922) 133–181.

[3] M. Braun, Differential Equations and Their Applications, 2nd ed., Springer-Verlag, New York, 1978.

[4] V. W. Bryant, A remark on a fixed-point theorem for iterated mappings, Amer. Math. Monthly, 75 (1968) 399–400.

[5] R. C. Buck, Advanced Calculus, 3rd ed., McGraw-Hill, New York, 1978.

[6] R. Charles, What'd I Say, Atlantic Recording No. 2031, New York, 1959.

[7] R. A. Fisher, The Genetical Theory of Natural Selection, 2nd revised ed., Dover, New York, 1958.

[8] G. E. Forsythe and C. B. Moler, Computer Solution of Linear Algebraic Systems, Prentice-Hall, Englewood Cliffs, New Jersey, 1967.

[9] C. Goffman, Preliminaries to functional analysis, Studies in Modern Analysis, R. C. Buck, ed., Math. Assoc. Amer., Washington, D.C., 1962, pp. 138–180.

[10] P. Henrici, Elements of Numerical Analysis, Wiley, New York, 1964.

[11] E. Isaacson and H. B. Keller, Analysis of Numerical Methods, Wiley, New York, 1966.

[12] I. Kaplansky, Set Theory and Metric Spaces, 2nd ed., Chelsea, New York, 1977.

[13] J. G. Kemeny and J. L. Snell, Finite Markov Chains, Springer-Verlag, New York, 1976.

[14] A. N. Kolmogorov and S. V. Fomin, Elements of the Theory of Functions and Functional Analysis, vol. 1, Graylock, Rochester, New York, 1957.

[15] I. I. Kolodner, On the proof of the contractive mapping theorem, Amer. Math. Monthly, 74 (1967) 1212–1213.

[16] B. Noble and J. W. Daniel, Applied Linear Algebra, 2nd ed., Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[17] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[18] J. H. Pollard, Mathematical Models for the Growth of Human Populations, Cambridge Univ. Press, New York, 1973.

[19] A. Ralston and P. Rabinowitz, A First Course in Numerical Analysis, 2nd ed., McGraw-Hill, New York, 1978.

[20] F. S. Roberts, Discrete Mathematical Models, Prentice-Hall, Englewood Cliffs, New Jersey, 1976.

[21] P. A. Samuelson, Foundations of Economic Analysis, Atheneum, New York, 1965.

[22] P. A. Samuelson, Economics, 11th ed., McGraw-Hill, New York, 1980.

[23] G. F. Simmons, Differential Equations with Applications and Historical Notes, McGraw-Hill, New York, 1972.

[24] G. Strang, Linear Algebra and Its Applications, 2nd ed., Academic Press, New York, 1980.

[25] R. S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

[26] C. O. Wilde, The contraction mapping principle, (UMAP Unit 326), Birkhäuser Boston, Cambridge, Massachusetts, 1978.

[27] A. Wouk, Direct iteration, existence and uniqueness, Nonlinear Integral Equations, P. M. Anselone, ed., Univ. of Wisconsin Press, Madison, Wisconsin, 1964, pp. 3–34.

NOTES

Tooth Tables: Solution of a Dental Problem by Vector Algebra

GARY HOSLER MEISTERS
University of Nebraska—Lincoln
Lincoln, NE 68588

The following problem in dentistry concerning the way in which a tooth is prepared for a gold inlay was posed to me by two instructors at the University of Nebraska Dental College, Linda Dubois and Stanley Kull. They expected that a mathematical analysis of the problem would require heavy use of a computer. However, the problem they posed is nicely amenable to a formulation (and solution) using some elementary facts of three-dimensional vector algebra which are usually included in the standard first course in calculus. The solution we obtained and its implications for dentistry have been described in the paper [**2**]. But, since [**2**] was written for dentists, the mathematics was only briefly summarized and little attempt was made to explain the details of the mathematical formulation and solution of the problem. Consequently, it seemed appropriate to write this paper in order to give a full account of the mathematics involved.

In order to describe the problem we will need a small amount of dental terminology which can be easily learned from FIGURE 1 and the accompanying "Dental Vocabulary" gleaned from Dorland's Medical Dictionary [**1**] and Gray's Anatomy [**3**]. A tooth is prepared for a gold inlay by carving out (from the occlusal surface and from a proximal surface of contact with a neighboring tooth) a cavity shaped more or less as illustrated in FIGURE 1. This cavity is called an **inlay preparation.**

The **proximal box** is that part of the inlay preparation which is bounded in the rear by the **axial wall,** on the sides by the left and right walls, on the bottom by the **seat,** on the top by a portion of the **occlusal** (which has been cut away), and on the front by a portion of the **proximal surface of contact** (which has also been cut away).

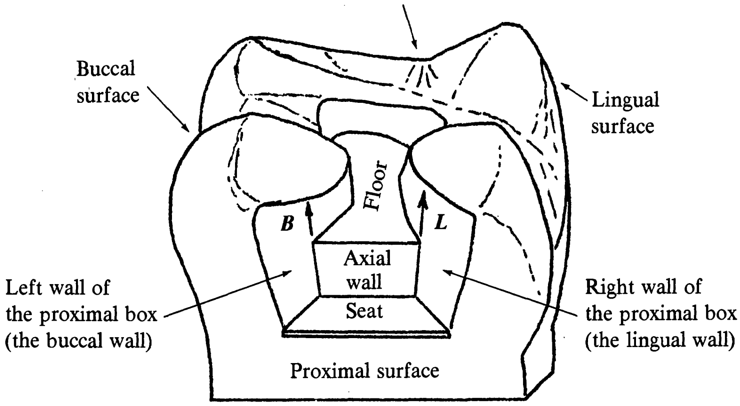


FIGURE 1. **A tooth (molar) showing a gold-inlay preparation carved out of it.**

A Brief Dental Vocabulary
Buccal surface is the cheek side of tooth.
Lingual surface is the tongue side of the tooth.
Occlusal surface is the chewing surface of the tooth.
Preparation (or **inlay preparation**) is the shaped cavity cut out of a tooth to receive a gold inlay.
Proximal is next to the neighboring tooth.
Proximal box is that portion of the inlay preparation nearest to the neighboring tooth.
Proximal surface is the surface of the tooth which is in contact with the neighboring tooth.
Seat (or **gingival seat**) is the seat of the proximal box nearest the gums.
Floor of inlay preparation is parallel to the seat but is closer to the occlusal surface. It is more-or-less perpendicular to the axial wall and extends from the axial wall into the center of the tooth. The problem discussed in this paper is not concerned with the part of the inlay preparation which lies between the floor and the occlusal surface.

It is required to cut the cavity for the gold inlay in such a manner that a wax mold of the cavity can be drawn out toward the occlusal and the gold inlay itself can be dropped into the cavity from the occlusal side of the tooth. This is called an **occlusal draw**. In order to guarantee a good occlusal draw it is both necessary and sufficient that all of the nearly-vertical walls of the inlay preparation should be tilted slightly back so that they all face slightly toward the occlusal surface of the tooth.

The lines of intersection of the axial wall with the left and right walls of the proximal box will be called the **axial corner lines**. Their continuations are marked in FIGURE 1 by arrows **B** and **L** (for "buccal" and "lingual"). Since these axial corner lines lie in a common plane (namely, the plane of the axial wall) they are either parallel or else they intersect. If they intersect on the occlusal side of the seat, then we say that these lines "converge in the direction of the occlusal" or simply "**converge**." On the other hand, if they intersect on the root side of the seat, then we say that these lines "diverge in the direction of the occlusal" or simply "**diverge**."

It has been customary for dentists to cut inlay preparations in such a way that the axial corner lines diverge in the occlusal direction as illustrated by the two arrows **B** and **L** shown in FIGURE 1. Evidently this divergence of **B** and **L** has been generally regarded by dentists (see [**2**] and the references cited there) as both necessary and sufficient in order to avoid an undercut by the proximal box walls in an inlay preparation. Indeed, divergence is frequently used as a visual criterion for a good occlusal draw. It is the main purpose of this paper to prove by means of a mathematically rigorous argument that this divergence is neither necessary nor sufficient for an occlusal draw. In fact we show that

(i) *Convergence of* **B** *and* **L** *toward the occlusal can be consistent with a good occlusal draw*;
(ii) *Divergence is no guarantee of a good occlusal draw.*

One consequence of this result which may have important significance for dentistry is that when an inlay preparation is cut with *convergence* (consistant with draw) it *allows less good tooth material to be cut away* than when one requires divergence.

It remains of course for dental instructors to perfect a procedure which, on the one hand, allows convergence and guarantees a good draw and, on the other hand, is simple enough to be practical for dentists to learn and to be able to put into practice. In this paper we have merely shown that it is geometrically possible.

The mathematical formulation and solution of the problem

We are concerned with the geometry of the proximal box, and it suffices to consider the right-hand side since the left-hand side could be treated in a completely analogous manner. We introduce a three-dimensional rectangular coordinate system oriented so that the seat of the proximal box lies in the xy-plane, the z-axis coincides with the central axis of the tooth (directed from the root out through the occlusal surface in the direction of draw), the y-axis is parallel to the plane of the axial wall and passes out through the right (lingual) surface, and the x-axis pierces at right angles the line of intersection of the axial wall with the seat and then exits the tooth through the proximal surface. This rectangular coordinate system is illustrated in FIGURE 2. Instead of comparing the right **L** and left **B** axial corner lines (for convergence or divergence toward the occlusal), we will compare the right (or lingual) axial corner line **L** with the line of intersection of the axial wall with the xz-plane. This latter line we call the "**medial**" **line**.

As usual, we use the letters $\mathbf{i},\mathbf{j},\mathbf{k}$ to denote the unit vectors pointing in the x, y, z directions, respectively. Recall that the orientation of a plane in three-dimensional space can be conveniently described mathematically by means of a unit normal vector (a unit vector perpendicular to the surface of the plane). In FIGURE 2, the vector labeled **A** is the unit normal vector to the plane of the axial wall, and the vector labeled **P** is the unit normal to the plane of the right (lingual) wall of the proximal box. Both of these vectors are chosen so that their directions point *into* the proximal box. Also in FIGURE 2, the vector labeled **M** is parallel to the medial line and the vector labeled **L** is parallel to the right (lingual) axial corner line.

For the formulation and solution of our problem we shall need to use the following well-known properties of the "cross" and "dot" products of the two vectors **U** and **V**, which can be found in any good calculus text.

(1) Definitions. If $\mathbf{U}=u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k}$ and $\mathbf{V}=v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k}$, then the **cross product** of **U** and **V** is defined as

$$\mathbf{U}\times\mathbf{V}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{vmatrix},$$

or

$$\mathbf{U}\times\mathbf{V}=(u_2v_3-u_3v_2)\mathbf{i}-(u_1v_3-u_3v_1)\mathbf{j}+(u_1v_2-u_2v_1)\mathbf{k}.$$

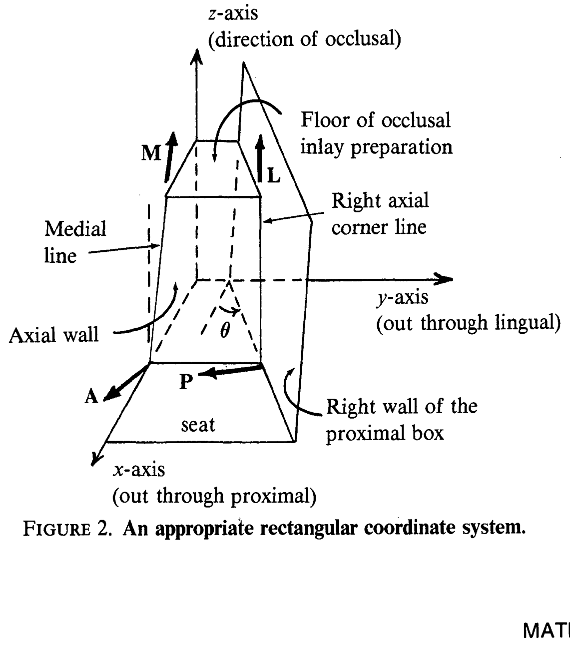


FIGURE 2. **An appropriate rectangular coordinate system.**

The **dot product** of $\mathbf{U}$ and $\mathbf{V}$ is defined as $\mathbf{U}\cdot\mathbf{V}=u_1v_1+u_2v_2+u_3v_3$. The **length** $|\mathbf{U}|$ is defined to be $[(u_1)^2+(u_2)^2+(u_3)^2]^{1/2}$.

(2) $\mathbf{U}\times\mathbf{V}$ is perpendicular to both $\mathbf{U}$ and $\mathbf{V}$.

(3) $\mathbf{U}\times\mathbf{V}=-\mathbf{V}\times\mathbf{U}$.

(4) The "right-hand rule." If $0\leqslant\phi\leqslant\pi$ is the (smaller) angle between $\mathbf{U}$ and $\mathbf{V}$, then $\mathbf{U}\times\mathbf{V}$ points in the direction of the thumb when the right-hand fingers are folded (through the angle ϕ) from the direction of $\mathbf{U}$ into the direction of $\mathbf{V}$.

(5) $|\mathbf{U}\times\mathbf{V}|=|\mathbf{U}||\mathbf{V}|\sin\phi$.

(6) $\mathbf{U}\cdot\mathbf{V}=|\mathbf{U}||\mathbf{V}|\cos\phi$.

(7) $\mathbf{U}\times(\mathbf{V}\times\mathbf{W})=(\mathbf{U}\cdot\mathbf{W})\mathbf{V}-(\mathbf{U}\cdot\mathbf{V})\mathbf{W}$.

(8) $\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W})=(\mathbf{U}\times\mathbf{V})\cdot\mathbf{W}=\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$. If this product is positive (negative), the triple $\mathbf{U},\mathbf{V},\mathbf{W}$ is called dextral (sinistral).

Now note that since $\mathbf{M}$ is parallel to the xz-plane and also to the axial plane (see FIGURE 2), $\mathbf{M}$ must be perpendicular to both of the vectors $\mathbf{j}$ and $\mathbf{A}$. Therefore (since the length of $\mathbf{M}$ is immaterial), we may write

$$\mathbf{M}=\mathbf{A}\times\mathbf{j}, \tag{9}$$

and similarly,

$$\mathbf{L}=\mathbf{P}\times\mathbf{A}. \tag{10}$$

Since $\mathbf{M}$ and $\mathbf{L}$ are both parallel to the axial wall, it follows that either $\mathbf{M}\times\mathbf{L}=\mathbf{0}$ (when they are parallel to each other) or $\mathbf{M}\times\mathbf{L}$ is a nonzero scalar multiple of the vector $\mathbf{A}$. Consideration of the right-hand rule (4) and the algebraic sign of this scalar coefficient of $\mathbf{A}$ yields the following criteria:

I. *The right (lingual) corner line converges toward the medial line in the direction of the occlusal precisely when*

$$\mathbf{M}\times\mathbf{L}=\lambda\mathbf{A},\ \textit{with}\ \lambda>0.$$

II. *The right (lingual) corner line diverges from the medial line in the direction of the occlusal precisely when*

$$\mathbf{M}\times\mathbf{L}=-\lambda\mathbf{A},\ \textit{with}\ \lambda>0.$$

III. *The condition for a good occlusal draw is met (as far as the proximal box is concerned) precisely when the normal vectors* $\mathbf{A}$ *and* $\mathbf{P}$ *(as well as the normal vector to the left wall) each have a positive z-component.*

The question then is this: Is occlusal convergence of $\mathbf{L}$ toward $\mathbf{M}$ consistent with a good occlusal draw? That is, can conditions I and III be simultaneously met, rather than (as traditionally believed by dentists) only II and III?

In order to answer this question, we will need the x, y, and z components of the unit normal vectors $\mathbf{A}$ and $\mathbf{P}$. These can be obtained from FIGURES 3, 4, and 5 which are two-dimensional figures extracted from FIGURE 2.

From FIGURE 3 we see that

$$\mathbf{A}=(\cos\alpha)\mathbf{i}+(\sin\alpha)\mathbf{k}, \tag{11}$$

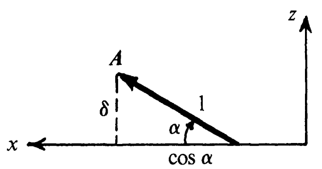


FIGURE 3. **Components of the vector A; here** $\delta=\sin\alpha$.

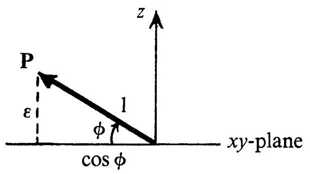


FIGURE 4. **The z-component ε of the vector P; here $\varepsilon = \sin\phi$.**

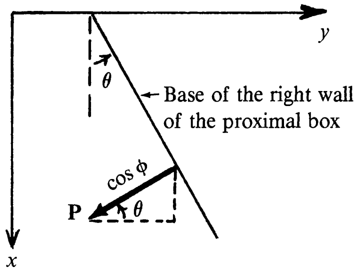


FIGURE 5. **Projection of P onto the xy-plane.**

and from FIGURES 4 and 5 we see that

$$\mathbf{P} = (\cos\phi\sin\theta)\mathbf{i} - (\cos\phi\cos\theta)\mathbf{j} + (\sin\phi)\mathbf{k}. \tag{12}$$

Extracting the z-components from (11) and (12), criterion III says that the condition for good occlusal draw is met when $\delta = \sin\alpha > 0$ and $\varepsilon = \sin\phi > 0$. The axial plane will be undercut if $\delta < 0$, and the right wall of the proximal box will be undercut if $\varepsilon < 0$.

We can apply some of the properties of cross product to rewrite $\mathbf{M}\times\mathbf{L}$ in criteria I and II:

$$\mathbf{M}\times\mathbf{L} = \mathbf{M}\times(\mathbf{P}\times\mathbf{A}) = (\mathbf{M}\cdot\mathbf{A})\mathbf{P} - (\mathbf{M}\cdot\mathbf{P})\mathbf{A} = -(\mathbf{M}\cdot\mathbf{P})\mathbf{A}.$$

Thus criteria I and II can be rephrased as follows:

I′, II′. $\mathbf{M}\cdot\mathbf{P}$ *should be negative for occlusal convergence and positive for occlusal divergence.*

But by formulas (8), (9), (11), and (12) we obtain

$$\mathbf{M}\cdot\mathbf{P} = (\mathbf{A}\times\mathbf{j})\cdot\mathbf{P} = \begin{vmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ \cos\phi\sin\theta & -\cos\phi\cos\theta & \sin\phi \end{vmatrix}$$

$$= \cos\alpha\sin\phi - \sin\alpha\cos\phi\sin\theta.$$

Therefore, the criterion for occlusal convergence of the right axial corner line **L** toward the medial line **M** becomes

$$\tan\phi < \tan\alpha\sin\theta, \tag{13}$$

while the criterion for a good occlusal draw may be expressed as

$$\tan\phi > 0 \text{ and } \tan\alpha > 0, \tag{14}$$

because $\cos\phi$ and $\cos\alpha$ are positive in the realistic range $|\phi| < \pi/2$ and $|\alpha| < \pi/2$.

Conditions (13) and (14) may easily hold simultaneously for a continuum of values of α, ϕ, and θ. That is, *convergence is definitely consistent with a good occlusal draw.*

The reverse inequalities

$$\tan\phi > \tan\alpha\sin\theta, \tag{15}$$

and

$$\tan\phi < 0 \text{ and/or } \tan\alpha < 0, \tag{16}$$

are also easily satisfied simultaneously for a continuum of values of ϕ, α and θ with $0 < \theta < \pi/2$. But (15) is the condition for *occlusal divergence* of the right axial corner line **L** away from the medial line $\boldsymbol{M}$, and (16) is the condition for failure of occlusal draw, because of an undercut of either the right proximal wall ($\tan\phi < 0$) or the axial wall ($\tan\alpha < 0$). Of course (15) implies that $\tan\phi$ can be negative only if $\tan\alpha$ is also negative, because $\sin\theta$ is always positive ($0 < \theta < \pi/2$). Thus *occlusal divergence is not a guarantee of a good occlusal draw*, and thus should not be used as a visual test for such. Note that for simultaneous convergence and draw it must be true that $\tan\phi < \tan\alpha$ because $0 < \sin\theta < 1$.

The criterion (I′) that $\mathbf{M}\cdot\mathbf{P}$ should be negative for convergence has important geometric implications. The dot product $\mathbf{M}\cdot\mathbf{P}$ negative means that the component of $\mathbf{P}$ in the direction of $\mathbf{M}$ is negative, i.e., the projection of $\mathbf{P}$ along the line through $\mathbf{M}$ has a direction opposite to that of $\mathbf{M}$. In addition, $\mathbf{M}\cdot\mathbf{P}=(\mathbf{A}\times\mathbf{j})\cdot\mathbf{P}$ is *negative* if and only if the three vectors $\{\mathbf{A},\mathbf{j},\mathbf{P}\}$ form a *left-handed* (sinistral) coordinate system.

A table of values for the angles that allow simultaneous convergence and draw

In addition to the angles α, ϕ, and θ introduced in FIGURES 3, 4, and 5, we can introduce the angle μ between $\mathbf{M}$ and $\mathbf{L}$ and call it the "angle of convergence" (taken positive for convergence and negative for divergence). We defined $\mathbf{M}$ in (9) as $\mathbf{M}=\mathbf{A}\times\mathbf{j}$; using equation (11), we find that

$$\mathbf{M}=\mathbf{A}\times\mathbf{j}=(-\sin\alpha)\mathbf{i}+(\cos\alpha)\mathbf{k},$$

so that $\mathbf{M}$ is a unit vector. From (6) it follows that $\mathbf{M}\cdot\mathbf{L}=|\mathbf{L}|\cos\mu$. Also, from (11) and (12) we have

$$\mathbf{M}\cdot\mathbf{L}=(\mathbf{A}\times\mathbf{j})\cdot(\mathbf{P}\times\mathbf{A})=\cos\phi\cos\theta,$$

so that

$$\cos\mu=(\cos\phi\cos\theta)/|\mathbf{L}|. \tag{17}$$

Properties (3), (5), and (8) give $\mathbf{M}\times\mathbf{L}=|\mathbf{L}|\mathbf{A}\sin\mu=-(\mathbf{M}\cdot\mathbf{P})\mathbf{A}$, so that

$$\sin\mu=(\sin\alpha\cos\phi\sin\theta-\cos\alpha\sin\phi)/|\mathbf{L}|. \tag{18}$$

Combining (17) and (18) gives

$$\tan\mu=\sin\alpha\tan\theta-\cos\alpha\tan\phi\sec\theta. \tag{19}$$

Thus the angle of convergence μ is completely determined by the angles α, ϕ, and θ.

By using standard tables of the trigonometric functions (such as those found in [4]), or a calculator with trig function capability, one can easily construct a table of acceptable values of the angles α, ϕ, and θ and compute the corresponding values of μ by means of the formula (19). We give an example of such a table in TABLE 1.

α	ϕ	θ	μ
6°	3°	30°	0.01°
"	"	35°	0.55°
"	"	40°	1.13°
"	"	45°	1.77°
"	"	50°	2.49°
"	"	55°	3.34°
"	"	60°	4.39°
"	"	65°	5.76°
"	"	70°	7.68°
7°	3°	26°	0.09°
"	"	30°	0.59°
"	"	35°	1.25°
"	"	40°	1.97°
"	"	45°	2.77°
"	"	50°	3.68°
"	"	55°	4.77°
"	"	60°	6.11°
"	"	65°	7.87°
"	"	70°	10.36°

α = angle of tilt of axial wall back from vertical.

ϕ = angle of tilt of proximal wall back from vertical.

θ = angle of (tangent plane) of right wall of proximal box.

μ = angle of convergence between vectors $\mathbf{M}$ and $\mathbf{L}$.

$$0<\tan\phi<\tan\alpha\sin\theta$$

Draw ↑ ↑ Convergence

TABLE 1

I would like to thank Paul Campbell for encouraging me to expand and write up this problem, which was the topic of a talk I presented to the Nebraska-South Dakota Section of the MAA at its meeting in Vermillion, South Dakota in April of 1981.

References

[1] Dorland's Illustrated Medical Dictionary, Twenty-fifth ed., W. B. Saunders, Philadelphia, 1974.

[2] Linda Dubois, Gary H. Meisters, and Stanley Kull, The relationship of wall orientation to proximal box design in an inlay preparation, (submitted to the Journal of Dental Research).

[3] Henry Gray, Anatomy of the Human Body, Twenty-ninth American Edition, C. M. Goss, ed., Lea & Febiger, Philadelphia, 1975.

[4] S. M. Selby, ed., Standard Mathematical Tables, Twenty-first ed., The Chemical Rubber Co., Cleveland, 1973.

Periodic Sums of Functions with Incommensurable Periods

Regina Caveny
James Caveny
University of Southern Mississippi
Hattiesburg, MS 39406

If f and g are real functions with fundamental periods α and β, respectively, when will the sum $f+g$ be periodic? Variations of this question arise naturally as soon as the concept of periodicity is introduced, perhaps in a trigonometry course, and surface repeatedly in more advanced mathematical treatments which involve periodic functions and their applications to various physical phenomena. If we admit only continuous functions, then the sum $f+g$ is periodic if and only if the ratio α/β is rational. But even for continuous functions the proofs are elusive. This paper deals mainly with the question of whether or not the irrationality of α/β insures that the sum $f+g$ is not periodic. We give some examples to show that the sum can be periodic when α/β is irrational and can even have a fundamental period. Then two conditions are given to establish that this does not happen for the functions which are normally encountered in applications. The arguments are, for the most part, self-contained and accessible to any serious undergraduate student of mathematics. We make no claim of originality of results, which must surely be fairly well known, but which seem to have escaped the popular literature. Our intent is simply to assemble a reasonably complete investigation which will be accessible to the general reader of this *Magazine*.

For preciseness we agree on the following terminology and notation. Throughout the paper, $\mathbb{J}$ denotes the set of integers and $\mathbb{R}$ the set of real numbers. A function f defined on $\mathbb{R}$ is **periodic with period α** if $f(x+\alpha)=f(x)$ for all x in $\mathbb{R}$. If there is a positive period α such that $\alpha \leqslant \beta$ for all positive periods β of f, then α is called the **fundamental period** of f. Two nonzero real numbers α and β are **commensurable** if their quotient is rational (i.e., there exist nonzero integers i and j such that $i\alpha + j\beta = 0$) and are **incommensurable** if their quotient is irrational (i.e., for all integers i and j, $i\alpha + j\beta = 0$ implies that $i=j=0$). If A and B are subsets of $\mathbb{R}$, then we define

$$A+B=\{x\in\mathbb{R}: x=a+b \text{ with } a\in A \text{ and } b\in B\}$$

$$A^c=\{x\in\mathbb{R}: x\notin A\}$$

and, for $r\in\mathbb{R}$,

$$rA=\{x\in\mathbb{R}: x=ra \text{ with } a\in A\}.$$

I would like to thank Paul Campbell for encouraging me to expand and write up this problem, which was the topic of a talk I presented to the Nebraska-South Dakota Section of the MAA at its meeting in Vermillion, South Dakota in April of 1981.

References

[1] Dorland's Illustrated Medical Dictionary, Twenty-fifth ed., W. B. Saunders, Philadelphia, 1974.
[2] Linda Dubois, Gary H. Meisters, and Stanley Kull, The relationship of wall orientation to proximal box design in an inlay preparation, (submitted to the Journal of Dental Research).
[3] Henry Gray, Anatomy of the Human Body, Twenty-ninth American Edition, C. M. Goss, ed., Lea & Febiger, Philadelphia, 1975.
[4] S. M. Selby, ed., Standard Mathematical Tables, Twenty-first ed., The Chemical Rubber Co., Cleveland, 1973.

Periodic Sums of Functions with Incommensurable Periods

REGINA CAVENY
JAMES CAVENY
University of Southern Mississippi
Hattiesburg, MS 39406

If f and g are real functions with fundamental periods α and β, respectively, when will the sum $f+g$ be periodic? Variations of this question arise naturally as soon as the concept of periodicity is introduced, perhaps in a trigonometry course, and surface repeatedly in more advanced mathematical treatments which involve periodic functions and their applications to various physical phenomena. If we admit only continuous functions, then the sum $f+g$ is periodic if and only if the ratio α/β is rational. But even for continuous functions the proofs are elusive. This paper deals mainly with the question of whether or not the irrationality of α/β insures that the sum $f+g$ is not periodic. We give some examples to show that the sum can be periodic when α/β is irrational and can even have a fundamental period. Then two conditions are given to establish that this does not happen for the functions which are normally encountered in applications. The arguments are, for the most part, self-contained and accessible to any serious undergraduate student of mathematics. We make no claim of originality of results, which must surely be fairly well known, but which seem to have escaped the popular literature. Our intent is simply to assemble a reasonably complete investigation which will be accessible to the general reader of this *Magazine*.

For preciseness we agree on the following terminology and notation. Throughout the paper, $\mathbb{J}$ denotes the set of integers and $\mathbb{R}$ the set of real numbers. A function f defined on $\mathbb{R}$ is **periodic with period α** if $f(x+\alpha)=f(x)$ for all x in $\mathbb{R}$. If there is a positive period α such that $\alpha \leqslant \beta$ for all positive periods β of f, then α is called the **fundamental period** of f. Two nonzero real numbers α and β are **commensurable** if their quotient is rational (i.e., there exist nonzero integers i and j such that $i\alpha+j\beta=0$) and are **incommensurable** if their quotient is irrational (i.e., for all integers i and j, $i\alpha+j\beta=0$ implies that $i=j=0$). If A and B are subsets of $\mathbb{R}$, then we define

$$A+B=\{x\in\mathbb{R}: x=a+b \text{ with } a\in A \text{ and } b\in B\}$$

$$A^c=\{x\in\mathbb{R}: x\notin A\}$$

and, for $r\in\mathbb{R}$,

$$rA=\{x\in\mathbb{R}: x=ra \text{ with } a\in A\}.$$

In case α and β are commensurable, there is a common multiple $\gamma = i\alpha = -j\beta$ with $i, j \in \mathbb{J}$. Then the sum $s = f + g$ satisfies

$$s(x+\gamma) = f(x+\gamma) + g(x+\gamma) = f(x+i\alpha) + g(x-j\beta) = f(x) + g(x) = s(x),$$

so that having commensurable fundamental periods for the terms f and g guarantees a periodic sum without further conditions on f and g. The least common multiple of α and β will be a period of the sum, but it need not be the fundamental period. In the special case where f and g are nonzero multiples of simple sine and cosine functions, respectively, the least common multiple of α and β is the fundamental period of the sum. This fact is fairly well known, but proofs are so elusive that it has been referred to as mathematical folklore [**11**]. One noncalculus proof was recently given in [**1**].

We now turn to the case where α and β are incommensurable. It is often stated, for certain types of functions f and g, that the sum is not periodic [**2**, p. ix]; [**3**, p. 31]; [**8**, pp. 9–10]. This is certainly true for continuous functions, but here again, the proofs are elusive. Example 1 below demonstrates that it is possible for the sum to be periodic when the periods α and β are incommensurable. The Theorem and Examples 2 and 3 shed some light on the role played by the property of boundedness of the functions f and g. Conditions 1 and 2 establish criteria for the terms f and g which will insure that the sum $f+g$ is not periodic when their periods are incommensurable.

To facilitate the discussion of the examples, we introduce two subsets X and Z of $\mathbb{R}$. Let α and β be any two incommensurable real numbers, and define the set X by

$$X = \{x \in \mathbb{R} : x = i\alpha + j\beta \text{ with } i, j \in \mathbb{J}\}.$$

The following three properties of X are easily verified:

(1) if $x \in X$, the integers i, j such that $x = i\alpha + j\beta$, are unique;

(2) $X + X \subset X$;

(3) $X^c + X \subset X^c$.

Let γ be a real number such that $k\gamma \notin X$ for all $k \in \mathbb{J}$ with $k \neq 0$. (There exist such γ since the set $\{(i/k)\alpha + (j/k)\beta \colon i, j, k \in \mathbb{J}, k \neq 0\}$ is countable.) If Z is the subset of $\mathbb{R}$ defined by

$$Z = \{z \in \mathbb{R} : z = i\alpha + j\beta + k\gamma \text{ with } i, j, k \in \mathbb{J}\},$$

then

(1′) if $z \in Z$, the integers i, j, k, such that $z = i\alpha + j\beta + k\gamma$, are unique;

(2′) $Z + Z \subset Z$;

(3′) $Z^c + Z \subset Z^c$.

EXAMPLE 1. We construct two bounded functions f and g with fundamental periods α and β, respectively, such that the sum has fundamental period γ. If Z is the set defined as above, put

$$f(z) = \begin{cases} 2^{-|j|} + 2^{-|k|}, & \text{if } z \in Z \text{ and } z = i\alpha + j\beta + k\gamma \\ 0, & \text{if } z \in Z^c \end{cases}$$

and

$$g(z) = \begin{cases} 2^{-|i|} - 2^{-|k|}, & \text{if } z \in Z \text{ and } z = i\alpha + j\beta + k\gamma \\ 0, & \text{if } z \in Z^c, \end{cases}$$

so that

$$s(z) = f(z) + g(z) = \begin{cases} 2^{-|i|} + 2^{-|j|}, & \text{if } z \in Z \text{ and } z = i\alpha + j\beta + k\gamma \\ 0, & \text{if } z \in Z^c. \end{cases}$$

Because of the similar forms of f, g, and s, we verify only the properties of f. To see that α is a period of f, note that if $z \in Z^c$, then $z + \alpha \in Z^c + Z \subset Z^c$, so that $f(z) = 0 = f(z+\alpha)$. On the other hand, if $z \in Z$, say $z = i\alpha + j\beta + k\gamma$, then $f(z+\alpha) = f((i+1)\alpha + j\beta + k\gamma) = 2^{-|j|} + 2^{-|k|} = f(i\alpha + j\beta + k\gamma) = f(z)$. To verify that α is the fundamental period of f, we show that every period of f is an integer multiple of α. If $\delta \in Z^c$, then δ cannot be a period of f since for $z \in Z$, $z+\delta \in Z^c$ and $f(z+\delta)=0$ but $f(z) \neq 0$. Thus, if δ is a period of f, then $\delta \in Z$, so $\delta = l\alpha + m\beta + n\gamma$ where l, m, $n \in \mathbb{J}$, and $f(\delta)=f(0)$ gives $2^{-|m|}+2^{-|n|}=2^0+2^0=2$. But the sum of the exponentials can achieve its strict absolute maximum only for $m=n=0$, and hence $\delta = l\alpha$ where $l \in J$.

We will see in the theorem below that the assumption that $k\gamma \in X^c$ for all nonzero $k \in \mathbb{J}$ was essential for the construction of the bounded functions of Example 1. We will need the following well-known fact about periodic functions.

PROPOSITION 1. *If f is a periodic function with fundamental period α and τ is any period of f, then τ is an integer multiple of α.*

Proof. We may assume $0<\alpha<|\tau|$ so there is a unique integer n and a unique real number r such that $|\tau|=n\alpha+r$, where $n>0$ and $0 \leqslant r<\alpha$. Then, $f(x)=f(x+|\tau|)=f(x+n\alpha+r)=f(x+r)$ for all x. Thus, if $r \neq 0$, f would have a positive period r smaller than α. Hence, $r=0$ and $|\tau|=n\alpha$.

THEOREM. *Let f and g have incommensurable fundamental periods α and β, respectively, and suppose that either f or g is bounded. If γ is any real number such that $k\gamma \in X$ for some integer $k \neq 0$, then the sum $s=f+g$ cannot have period γ.*

Proof. We may assume that f is bounded. We will show that the assumption that γ is a period of $s=f+g$ leads to a contradiction. If γ is a period of s, then so is $k\gamma$, and thus for all $y \in \mathbb{R}$ and $x=i\alpha+j\beta \in X$ we have

$$\begin{aligned} f(y+x+k\gamma)-f(y+x) &= f(y+j\beta+k\gamma)-f(y+j\beta) \\ &= g(y+j\beta)-g(y+j\beta+k\gamma) \\ &= g(y)-g(y+k\gamma). \end{aligned}$$

Taking successively $x=n(k\gamma) \in X$ for $n=0,1,2,\ldots$ leads to

$$f(y+nk\gamma)=f(y)+n[g(y)-g(y+k\gamma)].$$

Since f is bounded, we must have $g(y)-g(y+k\gamma)=0$ for all $y \in \mathbb{R}$. But then $f(y+k\gamma)-f(y)=0$ for all $y \in \mathbb{R}$ so that $k\gamma$ is a period of both f and g. Hence, $k\gamma$ is an integer multiple of both α and β by Proposition 1, contrary to the hypothesis that α and β are incommensurable.

By allowing f and g to be unbounded, it is possible to have a sum with a period in the set X.

EXAMPLE 2. If f and g are defined by

$$f(x)=\begin{cases} j & \text{for } x=i\alpha+j\beta \in X \\ 0 & \text{for } x \in X^c \end{cases}$$

and

$$g(x)=\begin{cases} -i & \text{for } x=i\alpha+j\beta \in X \\ 0 & \text{for } x \in X^c, \end{cases}$$

then

$$s(x)=f(x)+g(x)=\begin{cases} j-i & \text{for } x=i\alpha+j\beta \in X \\ 0 & \text{for } x \in X^c. \end{cases}$$

The (unbounded) functions f and g have fundamental periods α and β, respectively, and their sum s has fundamental period $\gamma=\alpha+\beta \in X$.

For the sake of completeness we note that if α and β are incommensurable, then the set $X=\{x \in \mathbb{R}\colon x=i\alpha+j\beta \text{ with } i, j \in \mathbb{J}\}$ is dense in $\mathbb{R}$. A proof of this can be based on the pigeonhole principle (Dirichlet drawer principle) from combinatorics. In [**10**, p. 126] this concept is generalized to the complex setting, and the result is used to show that the only triply-periodic analytic functions are constants (a result due to Jacobi).

For the functions which one normally encounters in applications, a periodic sum will insure that the periods of the terms cannot be incommensurable. Conditions 1 and 2 below each give criteria for f and g which will guarantee that the pathology of Examples 1 and 2 is not present.

CONDITION 1 (Strict absolute maxima). *Suppose that f and g have positive periods α and β, respectively, and let f and g both have strict absolute maxima at the origin, i.e.,*

$$f(0) > f(x) \quad \text{for all } x \in (0, \alpha)$$

and

$$g(0) > g(x) \quad \text{for all } x \in (0, \beta).$$

If $s = f + g$ has positive period γ, then α and β are commensurable.

Proof. $f(0) + g(0) = s(0) = s(\gamma) = f(\gamma) + g(\gamma)$, and since f and g can achieve their respective maxima only at multiples of their respective periods, we must have $m\alpha = \gamma = n\beta$ where m and n are integers.

This simple condition suffices to show, for example, that $\cos \pi x + \cos\sqrt{2}\,\pi x$ is not periodic. It is also worth noting that Condition 1 allows f and g to have randomly bad behavior in the interiors of their fundamental period intervals. Next, we establish a fairly universal condition that will be applicable to a large percentage of the functions which one encounters in practice. For this purpose, the following modification of Proposition 1 will be useful.

PROPOSITION 2. *If D is a periodic function (with a nonzero period) and*

$$\bar{\delta} = \inf\{\delta : \delta > 0 \quad \text{and} \quad D(x+\delta) = D(x) \quad \text{for all } x \in \mathbb{R}\},$$

then either $\bar{\delta} = 0$ or every period τ of D is an integer multiple of $\bar{\delta}$.

Proof. Assume that $\bar{\delta} \neq 0$ and let τ be any nonzero period of D. (Obviously $\tau = 0$ is an integer multiple of $\bar{\delta}$.) If $\bar{\delta}$ is a period of D, then τ is a multiple of $\bar{\delta}$ by Proposition 1. If $\bar{\delta}$ is not a period of D, there exist arbitrarily small positive numbers ε such that $\bar{\delta} + \varepsilon$ is a period of D. Since $0 < \bar{\delta} \leq |\tau|$, there exist unique numbers n and r such that $|\tau| = n\bar{\delta} + r$, where $n > 0$ is an integer and $0 \leq r < \bar{\delta}$. We want to show that $r = 0$. Assume to the contrary that $r > 0$ and let ε be a positive real number such that $\bar{\delta} + \varepsilon$ is a period of D and $r - n\varepsilon > 0$. Then $D(x) = D(x + |\tau|) = D(x + n\bar{\delta} + r) = D(x + n(\bar{\delta} + \varepsilon) + (r - n\varepsilon)) = D(x + (r - n\varepsilon))$ for all $x \in \mathbb{R}$. Thus, $r - n\varepsilon$ is a period of f. But $0 < r - n\varepsilon < r < \bar{\delta}$ contradicts the definition of $\bar{\delta}$. Hence $r = 0$ and τ is an integer multiple of $\bar{\delta}$.

Although the divisor property of the infimum of Proposition 2 is all we require for the verification of our next condition, it is true that the infimum is actually a period of D. We include this result for completeness.

COROLLARY. *The infimum $\bar{\delta}$ of the positive periods of a periodic function D is a period of D.*

Proof. If $\bar{\delta}$ is not a period of D, then $\bar{\delta} > 0$ and there exists a small positive number ε, say $0 < \varepsilon < \bar{\delta}$, such that $\bar{\delta} + \varepsilon$ is a period of D. But then Proposition 2 implies that $\bar{\delta} + \varepsilon$ is an integer multiple of $\bar{\delta}$. This is impossible since $\bar{\delta} < \bar{\delta} + \varepsilon < 2\bar{\delta}$.

Proposition 2 has a group-theoretic counterpart which we mention here. If f is any periodic function on $\mathbb{R}$, then the set F of all periods of f is an additive subgroup of $\mathbb{R}$. Conversely, if G is any additive subgroup of $\mathbb{R}$, then there exists a function g whose periods are precisely the elements of G (for example, the characteristic function of G). With this relationship between periodic functions on $\mathbb{R}$ and additive subgroups of $\mathbb{R}$, it is not hard to show that Proposition 2 is equivalent to the following theorem:

Every proper closed additive subgroup of $\mathbb{R}$ is a discrete group of the form aJ, where $a > 0$ [**4**, Chap. V, Prop. 1.1, p. 7].

The following condition is applicable to a broad class of functions, including sectionally continuous functions and functions of bounded variation.

CONDITION 2 (Finite one-sided limits). *Let f and g be periodic functions with fundamental periods α and β, respectively, and let $s=f+g$ have a positive period γ. If either f or g has finite one-sided limits at each point $x \in \mathbb{R}$, then α and β are commensurable.*

Proof. Put $D(x)=f(x+\gamma)-f(x)=g(x)-g(x+\gamma)$ so that both α and β are periods of D. Put $\bar{\delta}=\inf\{\delta : \delta>0$ and δ is a period of $D\}$. If $\bar{\delta}>0$, then both α and β are multiples of $\bar{\delta}$, and the conclusion follows. If $\bar{\delta}=0$, then D has arbitrarily small positive periods, and hence D takes on every value in its range in every interval $[x_0, x_0+\varepsilon]$ for any $\varepsilon>0$. This, together with the fact that

$$D(x_0^+)=\lim_{x\to x_0^+}[f(x+\gamma)-f(x)]=\lim_{x\to x_0^+}[g(x)-g(x+\gamma)]$$

exists at each point $x_0 \in \mathbb{R}$, implies that D must be constant, say $D(x)=K$ for all $x \in \mathbb{R}$. Taking $x=n\gamma$ successively for $n=0,1,2,\ldots$ leads to

$$f(n\gamma)=f(0)+nK \quad \text{and} \quad g(n\gamma)=g(0)-nK.$$

Let us assume for definiteness that $f(x^+)$ and $f(x^-)$ exist and are finite for each $x \in \mathbb{R}$. Then the Bolzano-Weierstrass Theorem insures that f must be bounded, and so we must have $K=0$. Thus, $f(x+\gamma)=f(x)$ and $g(x+\gamma)=g(x)$ for all $x \in \mathbb{R}$. Hence, since γ is a period of both f and g, by Proposition 1, γ must be a multiple of each of the fundamental periods α and β, and the conclusion follows.

We note that Conditions 1 and 2 do not encompass all of the situations which the reader is likely to encounter, even in practical applications. We give an example in which neither of the conditions is fulfilled.

EXAMPLE 3. If we define $f(x)$ and $g(x)$ for $x \in \mathbb{R}$ by

$$f(x)=\begin{cases} 0, & \text{if } x=\dfrac{2n+1}{2} \quad \text{for some } n \in \mathbb{J}, \\ \tan \pi x, & \text{otherwise,} \end{cases}$$

and

$$g(x)=\begin{cases} 0, & \text{if } x=\dfrac{2n+1}{2\sqrt{2}} \quad \text{for some } n \in \mathbb{J}, \\ \tan\sqrt{2}\,\pi x, & \text{otherwise,} \end{cases}$$

then Conditions 1 and 2 are not applicable. However, part of the proof of Condition 2 is salvageable. Suppose we assume, contrary to what will be shown, that the sum $f+g$ is periodic with a period γ. The first part of the proof of Condition 2 can be modified to conclude that D takes on every value in its range in every interval. One can use this fact to show that γ must be an integer multiple of 1, the period of f, by considering the behavior of D near the asymptotes of f. Similarly, it can be shown that γ is an integer multiple of $1/\sqrt{2}$, the period of g, by considering the behavior of D near the asymptotes of g. This is absurd, so the sum is not periodic.

The following theorem which is closely related to some of the arguments used in this paper was given in 1915 by C. Burstin: *A real valued Lebesgue measurable function f on $\mathbb{R}$ having arbitrarily small periods is constant almost everywhere.* The literature relating to Burstin's Theorem is fairly extensive. For three recent proofs, see [**5**], [**6**], and [**9**].

The spectra (Fourier transforms) of functions are utilized extensively in the engineering and physical sciences. In this connection the following quote is of interest here. "...Periodic functions have discrete spectra (line spectra) ... the spectra of nonperiodic functions are not discrete..." [**8**, p. 113]. Thus, for functions satisfying Condition 2, the spectrum of the sum of two periodic functions is discrete if and only if their periods are commensurable.

Finally, we mention an interesting application of periodic functions due to Gale [7]. The empirical assertion of that paper is that a listener should be able to distinguish between simple tones (where the listener is aware of only one pitch) and composite tones (where the listener is aware of more than one pitch). Musical tones are represented mathematically by periodic functions. Gale's paper contains a mathematical development of decomposable functions (which represent composite tones) and indecomposable functions (which represent simple tones). The major results of the paper are a uniqueness theorem regarding the expansion of functions into sums of indecomposable terms and an existence theorem which establishes the indecomposability of the characteristic function of the integers. The argument in [7] that the function $\cos x$ is indecomposable contains an error, and to our knowledge a correct proof of this has not been given.

References

[1] R. Beigel, Solution to Problem 1095, this MAGAZINE, 54 (1981) 142–143.

[2] A. S. Besicovitch, Almost Periodic Functions, Dover, New York, 1954.

[3] H. Bohr, Almost Periodic Functions, Chelsea, 1951.

[4] N. Bourbaki, Elements of Mathematics, General Topology, Part 2, Hermann, Paris, 1966.

[5] H. Burkill, The periods of a periodic function, this MAGAZINE, 47 (1974) 206–210.

[6] R. Cignoli and J. Hounie, Functions with arbitrarily small periods, Amer. Math. Monthly, 86 (1978) 582–584.

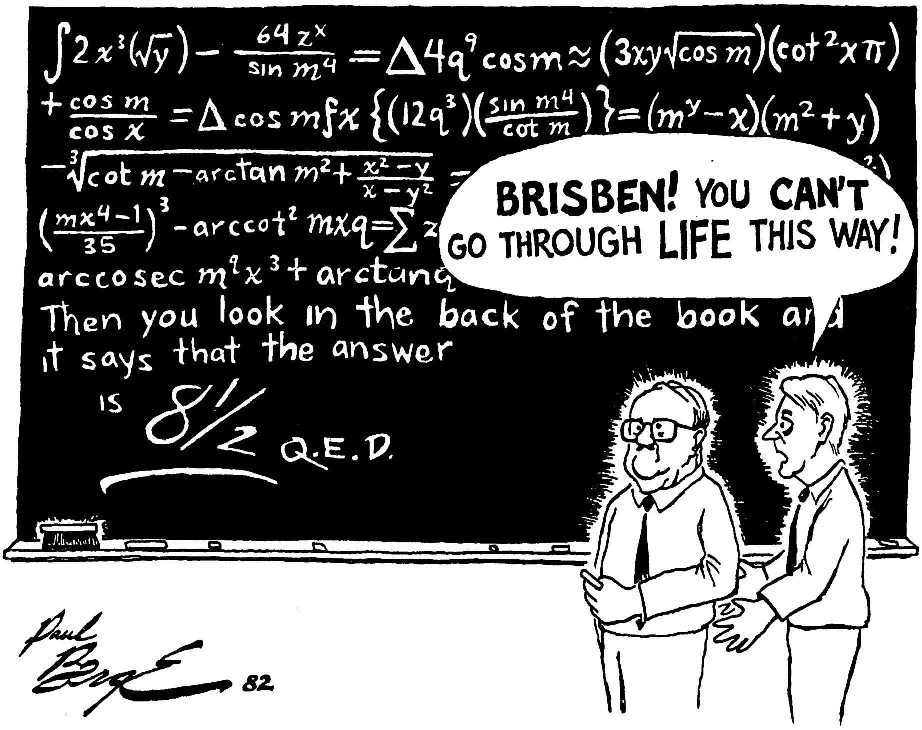
[7] D. Gale, Tone perception and decomposition of periodic functions, Amer. Math. Monthly, 86 (1979) 36–43.

[8] J. D. Gaskill, Linear Systems, Fourier Transforms, and Optics, Wiley, New York, 1978.

[9] J. M. Henle, Functions with arbitrarily small periods, Amer. Math. Monthly, 87 (1980) 816.

[10] E. Hille, Analytic Function Theory, vol. II, Ginn, 1962.

[11] S. Yeshurun, Problem 1095, this MAGAZINE, 53 (1980) 112.



Tiling Rectangles with Rectangles

F. R. K. Chung
E. N. Gilbert
R. L. Graham
Bell Laboratories
Murray Hill, NJ 07974

J. B. Shearer
Massachusetts Institute of Technology
Cambridge, MA 02139

J. H. van Lint
The Technical University of Eindhoven
Eindhoven, The Netherlands

Rectangles are among the most fundamental geometrical shapes. It is therefore only natural that some of the most common types of tilings involve rectangular pieces (or tiles), arranged in various (nonoverlapping) patterns so as to form some specified geometrical shape, often also a rectangle. In this note we will investigate certain basic kinds of tilings of rectangles by rectangles, namely, the so-called simple tilings. At the end of the paper we list for the interested reader a variety of additional references which treat various rectangular tiling problems, including the famous "squared square."

Definition. A tiling of a rectangle R by rectangles (called the **elements** of the tiling) is called **simple** if no connected set of two or more elements form a rectangle strictly inside of R.

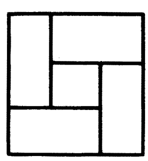


Figure 1

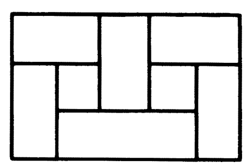


Figure 2

Figure 1 and Figure 2 show simple tilings of rectangles (which will be particularly useful in our later discussion). The fact that even two elements in a simple tiling cannot form a larger rectangle means that two adjacent elements in a simple tiling cannot have as their intersection the (whole) side of each. It is also not hard to see that no generality is lost by assuming (as we will from now on) that all elements have *integral* side lengths. Thus, an $m \times n$ rectangle can have a simple tiling only if $m \geqslant 3$ and $n \geqslant 3$. This shows that Figure 1 actually gives the smallest rectangle which has a simple tiling.

How many elements can be in a simple tiling?

An easy variation of the tiling in Figure 1, shown in Figure 3, (formed by adjusting the side lengths of some of the elements) provides a simple tiling with five elements for any $m \times n$ rectangle with $m \geqslant 3$, $n \geqslant 3$.

A natural problem to pose is:

Problem 1. *For which integers n does there exist a simple tiling with n elements?*

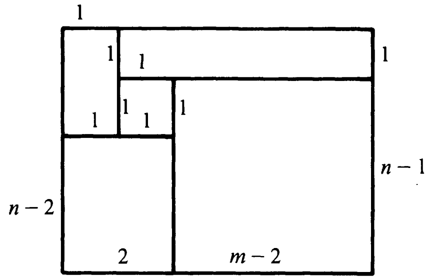


FIGURE 3

At this point the reader is encouraged to experiment in constructing some simple tilings in order to develop some intuition on simple tilings. It is quickly discovered that no simple tilings with n elements exist when $n \leqslant 4$ (no fair counting $n = 1$!). FIGURE 1 gives one with $n = 5$. What about $n = 6$? It turns out (although we will not prove it here) that no simple tiling can have exactly six elements. However, all *larger* values of n admit a tiling with n elements as the following result shows.

THEOREM 1. *For all $n \geqslant 7$, a simple tiling with n elements exists.*

Proof. In FIGURE 4 we show two "left blocks," four "right blocks" and one "middle block" of elements which can be combined to form simple tilings. The general construction pattern is given by taking one "left block," some number m (possibly zero) of copies of the "middle block," and one "right block" to form the tiling.

Since the "left blocks" have five or six elements, the "middle blocks" have seven elements, and the "right blocks" have two, four, six or eight elements, then a general tiling constructed this way can have $\begin{Bmatrix}5\\6\end{Bmatrix} + 7m + \begin{Bmatrix}2\\4\\6\\8\end{Bmatrix}$ elements, $m \geqslant 0$. Since these are *all* the integers $\geqslant 7$, the theorem is proved.

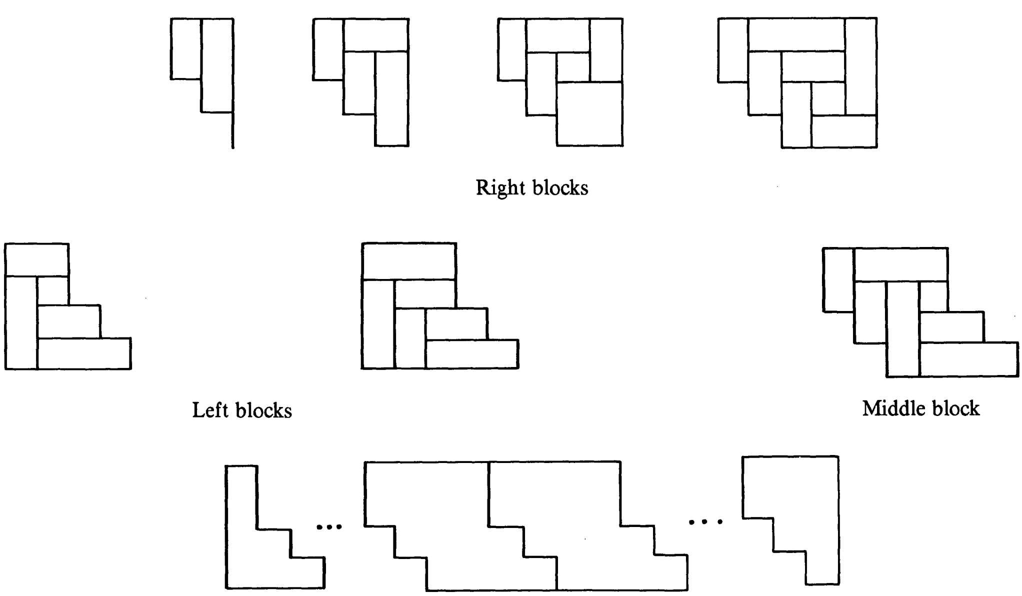


FIGURE 4

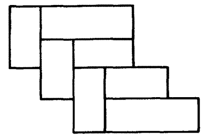


FIGURE 5

We point out here that there is a different middle block of seven elements which would have worked equally well, namely, the one shown in FIGURE 5. Since by forming simple tilings with large values of m middle blocks we could use any (ordered) mixture of the two types, it follows that for a suitable $c > 0$, there are in fact at least $c \cdot 2^{n/7}$ essentially different simple tilings using n elements. Of course, to make precise statements concerning the number of different simple tilings with n elements, we must say just what we mean by two tilings being different. For example, we would not regard the (infinitely many) tilings shown in FIGURE 3 as different. We will not pursue this question any further except to remark that under a rather natural definition of equivalence, it can be shown that there are no more than 20000^n simple tilings with n elements. This bound is not meant to be particularly accurate but rather only to show that the number of tilings is bounded by a simple exponential function of the form C^n as opposed to more rapidly growing functions such as $n!$ or 2^{n^2}.

A related question which suggests itself at this point is the following. What is the *maximum* number of elements possible in a simple tiling of an $m \times n$ rectangle?

This is exactly the problem which motivates the next section.

The average size of elements in a simple tiling

Given a simple tiling of a rectangle R, we define the **average area** of the elements of the tiling to be the ratio $\frac{\text{area } R}{n}$ where n is the number of elements in the tiling. For example, this average area for the tiling in FIGURE 1 is $9/5$, while that for the tiling in FIGURE 2 is $15/8$. Of course, for a fixed rectangle R, the average area of the elements of a tiling of R is minimized exactly when the number of elements in a tiling of R is made as large as possible.

PROBLEM 2. *Find a lower bound for the average area of the elements in any simple tiling of a rectangle.*

The reader may have discovered by now that almost all simple tilings of rectangles discovered by trial and error have an average area of elements of at least 2. This is true roughly because only a small fraction of the elements can have area 1. Thus, the tilings in FIGURE 1 and FIGURE 2 are especially "tight." The next result answers Problem 2.

THEOREM 2. (a) *The average area of the elements in a simple tiling of a* 3×3 *rectangle is* $9/5$. *For all other simple tilings of rectangles it is strictly greater than* $11/6$. (b) *There is a unique (up to rotation, translation and reflection) simple tiling of the plane for which the average area of the elements is* $11/6$.

Proof. To begin with, we examine what the elements adjacent to a unit square U in a simple tiling of a rectangle R can look like. It is not hard to see that U can never be on the boundary of the tiled rectangle and, in fact, one of the two situations shown in FIGURE 6 must hold.

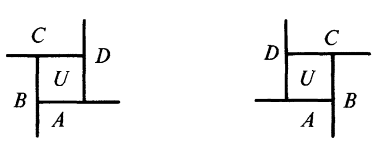


FIGURE 6

In FIGURE 6, elements A, B, C and D are called the **neighbors** of U. Also, we call B the **next neighbor** to A, C the next neighbor to B, etc. We want to define some measure of the "influence" of the various unit squares which may be in the simple tiling of R. To do this, we let A denote a generic element which has a unit square U as a neighbor and an element B as its next neighbor. We assign the value $\nu_U(A)$ to the pair A and U according to the following rules:

(i) If area $A = 2$ then $\nu_U(A) = \frac{1}{2}$;

(ii) If area $A > 2$ and area $B = 2$ then $\nu_U(A) = \frac{7}{2}$;

(iii) If area $A > 2$ and area $B > 2$ then $\nu_U(A) = \frac{5}{4}$.

Also, we assign to A the value $\nu(A)$, defined to be the sum $\nu_U(A)$ over all unit squares U neighboring A. Since the tiling given in FIGURE 1 cannot be a subconfiguration of our given simple tiling of R, at least one of the neighbors of *any* unit square U must have area greater than 2. Hence, every unit square U with neighbors A, B, C and D satisfies

$$\nu_U(A) + \nu_U(B) + \nu_U(C) + \nu_U(D) \geqslant 5. \tag{1}$$

Before continuing the proof, we need the following Lemma.

LEMMA.

(*a*) *If area* $A = 2$ *then* $\nu(A) \leqslant 1$.

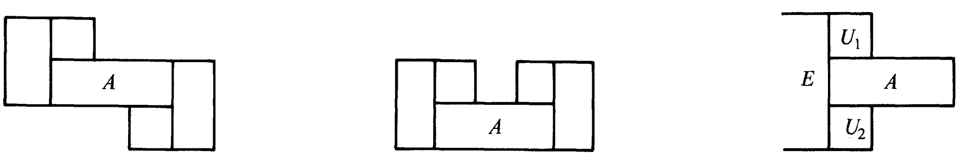
(b) *If area* $A = 3$ *then* $\nu(A) \leqslant 7$ *with equality only if* A *has* 1×1 *neighbors as shown in* FIGURE 7(a) *or* 7(b).

(c) *If area* $A = 4$ *then* $\nu(A) \leqslant 7$.

(d) *If area* $A > 4$ *then* $\nu(A) \leqslant 14$.

Proof (*of Lemma*). (a) This follows from the observation that in this case A can have at most two unit squares as neighbors.

(b) Since in this case A has size 1×3 then A has at most four 1×1 neighbors, i.e., unit squares.



(a) (b) (c)

FIGURE 7

If A has two 1×1 neighbors, then they must be placed as shown in FIGURES 7(a), 7(b), or 7(c). In FIGURE 7(c), the neighbor E of A must have area at least 3. Thus, $\nu_{U_1}(A) \leqslant \frac{5}{4}$ and $\nu_{U_2}(A) \leqslant \frac{5}{4}$. It follows that in this situation, $\nu(A) \leqslant \frac{5}{2}$. In FIGURES 7(a) and 7(b), it is easy to see that $\nu(A) = 7$. A similar argument shows that if A has three 1×1 neighbors then $\nu(A) \leqslant \frac{5}{2} + \frac{7}{2} = 6$ and if A has four 1×1 neighbors then $\nu(A) \leqslant 5$.

(c), (d) The proofs here are nearly the same as in (b). The only thing one needs to observe is that two neighbors of A which are unit squares cannot have a corner of A in common. (The details are left to the reader.)

Proof of Theorem 2 (continued). Let n_i denote the number of elements with area i in a simple tiling of a rectangle R where area $R > 9$. From inequality (1) and the Lemma we have

$$5n_1 \leqslant \sum_A \nu(A) \leqslant n_2 + 7n_3 + 7n_4 + 14 \sum_{i>4} n_i$$
$$\leqslant \sum_{i \geqslant 2} (6i - 11) n_i. \tag{2}$$

From (2) it follows that

$$\text{area } R = \sum_{i \geqslant 1} i n_i \geqslant \frac{11}{6} \sum_{i \geqslant 1} n_i. \tag{3}$$

Since $\Sigma_{i \geqslant 1} n_i$ is just the number of elements of the tiling then (3) implies that the average area of the elements in any simple tiling of R is at least $11/6$.

Let us now focus on the case of equality in (3). This implies that equality must also hold in (2), and in particular,

$$n_2 + 7n_3 + 7n_4 + 14 \sum_{i>4} n_i = \sum_{i \geqslant 2} (6i - 11) n_i,$$

i.e.,

$$0 = 6n_4 + 5n_5 + 11n_6 + 17n_7 + \cdots.$$

This can hold with $n_i \geqslant 0$ only if $n_i = 0$ for $i \geqslant 4$.

A further consequence of equality in (3) is that every element A of the tiling with area 2 has two 1×1 neighbors and every element A with area 3 has neighbors as shown in FIGURE 7(a) and 7(b). Since equality in (3) also implies that equality must hold for (1) as well, then any unit square U in the tiling must be surrounded by neighbors as shown in FIGURE 8. It now follows that a

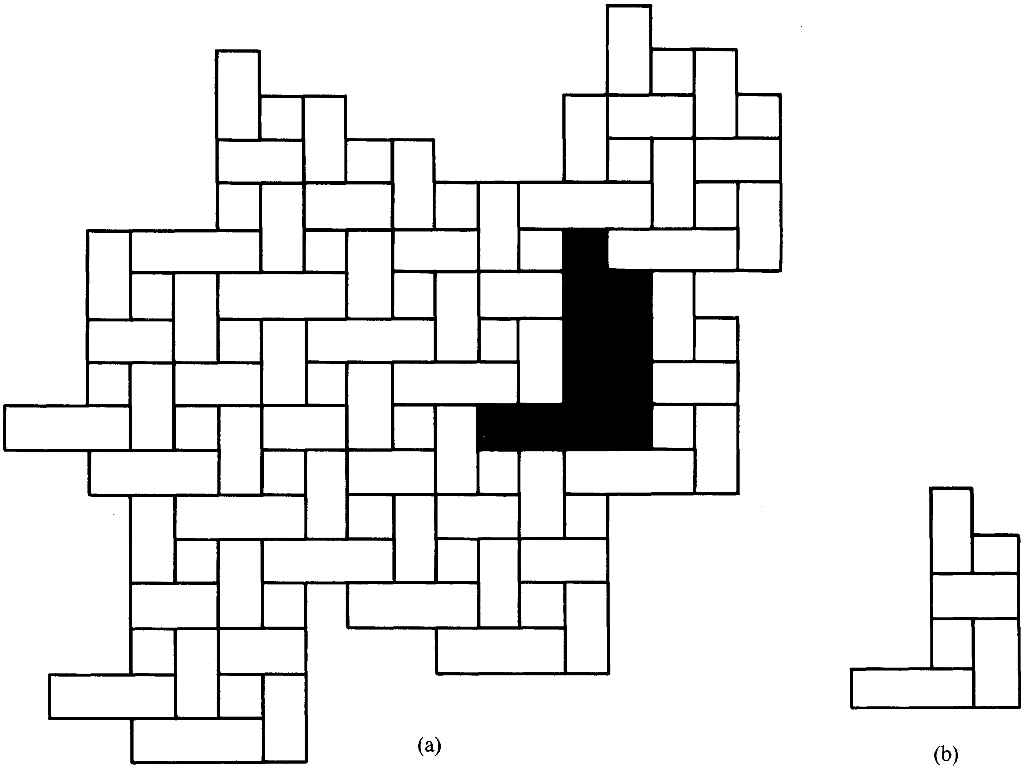
FIGURE 8

1×2 element A of the type shown in FIGURE 7(b) must be part of the configuration shown in FIGURE 2. However, for a *simple* tiling, this is impossible. Consequently, for elements A of area 3, all must occur as shown in FIGURE 7(a). From this we deduce that every 1×2 element must have its two 1×1 neighbors diagonally opposite, i.e., as shown in FIGURE 7(a) (or the mirror image of this).

Finally, if we start with the configuration shown in FIGURE 7(a), apply FIGURE 8 for the two unit squares and use the preceding remark for all 1×2 elements, a unique (up to reflection) tiling of the plane is forced. A portion of it is shown in FIGURE 9(a). Note that it can be generated by tiling with translates of the pattern B shown in FIGURE 9(b). It is also not difficult to check that this process can never terminate with an exact rectangle being simply tiled. Thus, any simply tiled *rectangle* must have the area of its elements strictly greater than $11/6$. However, if we appropriately tile a large nearly rectangular region S of the plane with translates of the pattern B in FIGURE 9(b), then it is not too hard to show that the jagged border of S can be completed by relatively few elements (possibly with areas greater than 3) to form a simply tiled rectangle which has an average element area as close to $11/6$ as desired (depending on the size of S; see the cover of this MAGAZINE). We leave the details of this construction to the energetic reader. This completes the proof of Theorem 2.

There is an active literature on the topic of tilings of rectangles by rectangles. The interested reader should consult some of the references ([**1**], [**14**], and especially [**6**]) for a sample.

As is natural with many geometrical problems, one might ask the analogous questions in three (or more) dimensions. To the best of our knowledge, almost nothing is known in these cases although it would certainly be an interesting area for exploration.



(a) (b)

FIGURE 9

The authors gratefully acknowledge the incisive comments of Doris Schattschneider in preparing the final form of this note.

References

[1] C. J. Bouwkamp, On the dissection of rectangles into squares, I, Koninkl. Nederl. Akad. Wetensch. Proc. Ser. A, 49 (1946) 1172–1178; II, III, 50 (1947) 58–71, 72–78.

[2] C. J. Bouwkamp, A. J. W. Duijvestijn, and P. Medema, Tables relating to simple squared rectangles of orders 9 through 15, Technische Hogeschool, Eindhoven, Netherlands, 1960.

[3] C. J. Bouwkamp, A. J. W. Duijvestijn and J. Haubrich, Catalogue of simple perfect rectangles of orders 9 through 18, Philips Research Laboratories, Eindhoven, Netherlands, 1964; unpublished.

[4] R. L. Brooks, C. A. B. Smith, A. H. Stone, and W. T. Tutte, The dissection of rectangles into squares, Duke Math. J., 7 (1940) 312–340.

[5] A. J. W. Duijvestijn, Electronic computation of squared rectangles, dissertation, Technische Hogeschool, Eindhoven, Netherlands; also in Philips Res. Rep., 17 (1962) 523–612.

[6] ______, Simple perfect squared squares of lowest order, J. Comb. Theory, B25 (1978) 260–263.

[7] A. J. W. Duijvestijn, P. J. Federico and P. Leeuw, Compound perfect squares, Amer. Math. Monthly, 89 (1982) 15–32.

[8] P. J. Federico, Note on some low-order perfect squares, Can. J. Math., 15 (1963) 350–362.

[9] ______, Squaring rectangles and squares: A historical review with annotated bibliography, in Graph Theory and Related Topics, J. A. Bondy and U. S. R. Murty, eds., Academic Press, 1979, pp. 173–196.

[10] Martin Gardner, Mystery tiles at Murray Hill, Isaac Asimov's Science Fiction Mag., 6 (1982) 40–42.

[11] R. L. Graham, Fault-free tilings of rectangles, in The Mathematical Gardner, D. A. Klarner, ed., Wadsworth Int'l., Belmont, California, 1981, pp. 120–126.

[12] N. D. Kazarinoff and R. Weitzenkamp, On existence of compound perfect squares of small order, J. Comb. Theory, B14 (1973) 163–179.

[13] ______, Squaring rectangles and squares, Amer. Math. Monthly, 80 (1973) 877–888.

[14] P. Leeuw, Electronic computation of compound squared squares, Bachelor's thesis, Twente Technical University, Enschede, Netherlands, 1979.

Reflections on a Pair of Theorems by Budan and Fourier

Alkiviadis G. Akritas
University of Kansas
Lawrence, KS 66045

Isolation of the real roots of a polynomial equation is the process of finding real, disjoint intervals such that each contains exactly one real root and every real root is contained in some interval. This process is quite important because, as J. B. J. Fourier pointed out, it constitutes the first step toward the solution of general equations of degree greater than four, the second step being the approximation of roots to any desired degree of accuracy.

In the beginning of the 19th century F. D. Budan and J. B. J. Fourier presented two different (but equivalent) theorems which enable us to determine the maximum possible number of real roots that an equation has within a given interval.

Budan's theorem appeared in 1807 in the memoir "Nouvelle méthode pour la résolution des équations numériques" [**10**, p. 219], whereas Fourier's theorem was first published in 1820 in "Le Bulletin des sciences par la Société Philomatique de Paris," pp. 156, 181 [**10**, p. 223]. Due to the importance of these two theorems, there was a great controversy regarding priority rights. In his book (1859) "Biographies of distinguished scientific men," p. 383, F. Arago informs us that Fourier "deemed it necessary to have recourse to the certificates of early students of the Polytechnic School or Professors of the University" in order to prove that he had taught his theorem in 1796, 1797 and 1803 [**10**]. Based on Fourier's proposition, C. Sturm presented in 1829 an improved theorem whose application yields the exact number of real roots which a polynomial equation without multiple zeros has within a real interval; thus he solved the real root isolation problem. Since 1830 Sturm's method has been the only one widely known and used, and consequently Budan's theorem was pushed into oblivion. To our knowledge, Budan's theorem can be found only in [**16**] and [**6**] whereas Fourier's proposition appears in almost all texts on the theory of equations. We feel that Budan's theorem merits special attention because it constitutes the basis of Vincent's forgotten theorem of 1836 which, in turn, is the foundation of our method for the isolation of the real roots of an equation [**1**], a method which far surpasses Sturm's in efficiency [**2**], [**3**].

In the discussion which follows we present separately, and without proofs, the classical theorems by Fourier and Budan and we indicate how they lead to the corresponding real root isolation methods. Some empirical results are also presented for comparison.

Fourier's theorem

Fourier's theorem, first published in 1820, was also included in his *Analyse des Equations*, published posthumously by C. L. M. N. Navier in 1831. Found in almost all texts on the theory of equations, it is sometimes given under the name Budan-Fourier or even Budan [**9**], [**17**]. Hurwitz [**12**] presents it as a special case of a more general theorem and Obreschkoff [**13**, pp. 76–87] generalizes it for complex roots. The statement given below is the way it is rendered by Vincent [**16**, p. 342]. We must first define the notion of sign variation.

Definition. We say that a sign variation exists between two nonzero numbers c_p and c_q $(p<q)$ of a finite or infinite sequence of real numbers $c_1, c_2, c_3, \ldots,$ if the following holds:

for $q=p+1$, c_p and c_q have opposite signs;
for $q \geqslant p+2$, the numbers $c_{p+1}, \ldots, c_{q-1}$ are all zero and c_p and c_q have opposite signs.

Theorem 1 (Fourier 1820). *If in the sequence of the $m+1$ functions $P(x), P^{(1)}(x), \ldots, P^{(m)}(x)$ (where $P^{(i)}=$ the ith derivative), we replace x by any two real numbers p, q $(p<q)$ and if we represent the two resulting sequences of numbers by $\tilde{P}$ and $\tilde{Q}$, then*

(i) *the sequence* $\tilde{P}$ *cannot present fewer sign variations than the sequence* $\tilde{Q}$;
(ii) *the number of real roots of the equation* $P(x)=0$, *located between* p *and* q, *can never be more than the number of sign variations lost in passing from the substitution* $x=p$ *to the substitution* $x=q$;
(iii) *when the first number is less than the second, the difference is an even number.*

The sequence of the $m+1$ derivatives is called **Fourier's sequence.** In (iii) the "first number" means the number of the real roots of $P(x)=0$ located between p and q; the "second number," on the other hand, refers to the number of sign variations lost in passing from the substitution $x=p$ to the substitution $x=q$. Obviously, Fourier's theorem gives an upper bound on the number of real roots which the equation $P(x)=0$ (of degree m) has inside the interval (p, q).

We remind the reader that the two main subjects of Fourier's life work were the theory of heat and the theory of the solution of numerical equations. Both of these subjects were carried forward by Sturm, who had personal and scientific relations with Fourier [**8**]. The manuscript of Fourier's treatise on the solution of numerical equations was by 1829 communicated to several persons including Sturm, who mentions explicitly what a great influence it had on his own work.

What Sturm did was to replace Fourier's sequence by

$$P(x), P^{(1)}(x), R_1(x), \ldots, R_k(x)$$

which is called **Sturm's sequence** or **chain**. This new sequence is obtained by applying the Euclidean algorithm to the polynomials $P(x)$ and $P^{(1)}(x)$, and taking $R_i(x), i=1,\ldots,k$ as the negative of the remainder polynomial; that is, the sequence is defined by the following relations:

$$\begin{aligned} P(x) &= P^{(1)}(x)Q_1(x) - R_1(x), \\ P^{(1)}(x) &= R_1(x)Q_2(x) - R_2(x), \\ &\vdots \\ R_{k-2}(x) &= R_{k-1}(x)Q_k(x) - R_k(x). \end{aligned}$$

The advantage of Sturm's sequence is that we can now obtain the exact number of real roots which the equation $P(x)=0$ has within a given interval. This is formally stated as follows:

THEOREM 2 (Sturm 1829). *If the equation* $P(x)=0$ *has only simple roots, then the number of its real roots in the interval* (p, q) *is equal to the difference*

$$v(p) - v(q),$$

where $v(\xi)$ *denotes the number of sign variations in Sturm's sequence for* $x=\xi$.

Sturm himself tells us [**8**] that the above theorem was merely a by-product of his extensive investigations on the subject of linear difference equations of the second order. The requirement that $P(x)=0$ has only simple roots is no restriction of the generality because we can first apply square-free factorization [**4**], [**15**] and then use Sturm's theorem.

Clearly Sturm's theorem can be used in the isolation of the real roots of an equation. The process itself is quite simple because all we have to do, once Sturm's sequence has been obtained, is to compute an absolute upper root bound b so that all the roots lie within the interval $(-b, b)$. We then subdivide this interval until in each subinterval there is at most one root; that is, Sturm's method is actually a bisection method. Quite recently, this method was implemented within a computer algebra system [**11**] using exact integer arithmetic algorithms and its computing time was thoroughly analyzed. (Computer algebra systems usually deal only with integer (rational) numbers, so that the user does not have to worry about round off and truncation errors. For a survey of such systems see [**14**].) It was shown that if $P(x)=0$ is an integral-coefficient univariate polynomial equation of degree $n>0$ without multiple roots, then the computing time of Sturm's method is

$$O\left(n^{13}L(|P|_\infty)^3\right)$$

where $L(|P|_\infty)$ is the length, in bits, of the maximum of the absolute values of the coefficients of P. This lengthy computing time shows Sturm's method leaves a lot to be desired; it has been determined that its slowness is due to the computation of the Sturm sequence.

Budan's theorem

Although Budan's theorem appeared much earlier than Fourier's, it seems to have been ignored; as far as we have been able to determine it does not appear in any of the standard texts on the theory of equations. The following statement of the theorem is from Vincent's paper [**16**, p. 342].

THEOREM 3 (Budan 1807). *If in an equation in* $x, P(x)=0$, *we make two transformations,* $x=p+x'$ *and* $x=q+x''$, *where* p *and* q *are real numbers such that* $p<q$, *then*

(i) *the transformed equation in* $x'=x-p$ *cannot have fewer sign variations than the transformed equation in* $x''=x-q$;
(ii) *the number of real roots of the equation* $P(x)=0$, *located between* p *and* q, *can never be more than the number of sign variations lost in passing from the transformed equation in* $x'=x-p$ *to the transformed equation in* $x''=x-q$;
(iii) *when the first number is less than the second, the difference is always an even number.*

Like Theorem 1, Budan's theorem also gives us an upper bound on the number of real roots of the equation $P(x)=0$ inside the interval (p,q). However, it only makes use of the transformations $x=p+x'$ and $x=q+x''$ and does not depend on any sequence of polynomials.

Theorems 1 and 3 are equivalent; this fact can be easily seen if in Fourier's sequence we replace x by any real number α. The $m+1$ resulting numbers are proportional to the corresponding coefficients of the transformed polynomial equation $P(x+\alpha)=0$, obtained by Taylor's expansion theorem.

Budan's theorem constitutes the basis of the following statement [**16**], [**3**].

THEOREM 4. *Let* $P(x)=0$ *be a polynomial equation of degree* $n>1$, *with rational coefficients and without multiple roots, and let* $\Delta>0$ *be the smallest distance between any two of its roots. Let* m *be the smallest index such that*

$$\frac{1}{2}F_{m-1}\Delta>1 \text{ and } F_{m-1}F_m\Delta>1+\frac{1}{\varepsilon_n},$$

where F_k *is the kth member of the Fibonacci sequence*

$$1,1,2,3,5,8,13,21,\ldots,$$

and

$$\varepsilon_n=\left(1+\frac{1}{n}\right)^{\frac{1}{n-1}}-1.$$

Let $a_1,a_2,\ldots,a_m$ *be arbitrary positive integers. Then the transformation*

$$x=a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+\cfrac{\ddots}{\ddots+\cfrac{1}{a_m+\cfrac{1}{y}}}}} \tag{1}$$

(which is equivalent to the series of successive transformations of the form $x = a_i + 1/\xi$, $i = 1, 2, \ldots, m$) transforms the equation $P(x) = 0$ into the equation $\tilde{P}(y) = 0$, which has not more than one sign variation.

This theorem is an extended version of the one originally presented by Vincent [**16**], [**4**]. The latter was first hinted by Fourier and, in his paper, Vincent indicates his surprise that Fourier did not try to go further and prove the proposition that was the main subject of Vincent's article. He states, however, the belief that such a proof may exist in other manuscripts which were not published because of the untimely death of Navier.

Theorem 4 can also be used in the isolation of the real roots of an equation. To see roughly why it is true and also how it is applied, observe the following:

(i) The continued fraction transformation (1) can be also written as

$$x = \frac{P_m y + P_{m-1}}{Q_m y + Q_{m-1}}, \tag{2}$$

where P_k/Q_k is the kth convergent to the continued fraction

$$a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}$$

and, as we recall,

$$P_{k+1} = a_{k+1}P_k + P_{k-1},$$
$$Q_{k+1} = a_{k+1}Q_k + Q_{k-1}.$$

(ii) The distance between two consecutive convergents is

$$\left| \frac{P_{m-1}}{Q_{m-1}} - \frac{P_m}{Q_m} \right| = \frac{1}{Q_{m-1}Q_m}.$$

It can be proven that the smallest values of the Q_i occur when all of the $a_i = 1$. Then $Q_m = F_m$, the mth Fibonacci number. This explains why there is a relation between the Fibonacci numbers and the distance Δ in Theorem 4.

(iii) Let $\tilde{P}(y) = 0$ be the equation obtained from $P(x) = 0$ after a transformation of the form (2). Observe that (2) maps the interval $0 < y < \infty$ onto the x-interval whose unordered endpoints are the consecutive convergents P_{m-1}/Q_{m-1} and P_m/Q_m. If this x-interval has length less than Δ, then it contains at most one root of $P(x) = 0$, and the corresponding equation $\tilde{P}(y) = 0$ has at most one root in $(0, \infty)$.

(iv) If $\tilde{y}$ was this positive root, then the corresponding root $\tilde{x}$ of $P(x) = 0$ could be easily obtained from (2). We only know though, that $\tilde{y}$ lies in the interval $(0, \infty)$; therefore, substituting y in (2) once by 0 and once by ∞, we obtain for the positive root $\tilde{x}$ its isolating interval whose unordered endpoints are P_{m-1}/Q_{m-1} and P_m/Q_m. To each positive root there corresponds a different continued fraction; at most m partial quotients have to be computed for the isolation of any positive root. (Negative roots can be isolated if we replace x by $-x$ in the original equation.)

REMARK. It is clear that if we knew the value of Δ, we could compute m from the inequalities of Theorem 4. Then, without any tests, we could obtain $\tilde{P}(y) = 0$. However, in our algorithmic procedure (to be described below), we do not initially know Δ. Thus we need the stronger conclusion that $\tilde{P}(y) = 0$ has at most one sign variation in order to have an effective test for root isolation. This is what requires the additional complexities in our theorem. For details see [3].

From the above discussion it is obvious that the calculation of the partial quotients $a_1, a_2, \ldots, a_m$ (for each positive root) constitutes the real root isolation procedure. (From Budan's theorem we know that the value of a particular partial quotient a_i has been computed if $P(x+a_i)=0$ has more sign variations in the sequence of its coefficients than $P(x+a_i+1)=0$.) There are two methods, Vincent's and ours, corresponding to the two different ways in which the computation of the a_i's may be performed. As we will see, the difference between these two methods can be thought of as being analogous to the difference between the integrals of Riemann and Lebesgue. That is, it is well known that the sum $1+1+1+1+1$ can be computed in the following two ways: (a) $1+1=2$, $2+1=3$, $3+1=4$, $4+1=5$ (Riemann) and (b) $5\cdot 1=5$ (Lebesgue).

Vincent's method basically consists of computing a particular a_i by a series of unit incrementations $a_i \leftarrow a_i+1$ (replace a_i by a_i+1), with each one of which we have to perform the translation $\tilde{P}(x) \leftarrow \tilde{P}(x+1)$ (for some polynomial equation $\tilde{P}(x)=0$) and check for a change in the number of sign variations. This "brute force" approach results in a method with exponential behavior and hence is of little practical importance. As an example, let us isolate the roots of the polynomial equation

$$P(x)=(x-\alpha)(x-\beta)=0$$

where $\alpha = 5\cdot 10^9+\varepsilon$ and $\beta=\alpha+1$. Consider $a_1^{(\alpha)}$, the first partial quotient for α, which is $5\cdot 10^9$. Using Vincent's method we set $a_1^{(\alpha)} \leftarrow 1$, $\tilde{P}(x) \leftarrow P(x)$ and compute $\tilde{P}(x) \leftarrow \tilde{P}(x+1)$. Since the number of sign variations in the sequence of coefficients of the transformed polynomial $\tilde{P}(x)$ has not changed, we set $a_1^{(\alpha)} \leftarrow a_1^{(\alpha)}+1$ and compute a new $\tilde{P}(x) \leftarrow \tilde{P}(x+1)$, checking again the number of sign variations. This process is repeated $5\cdot 10^9$ times and, on the fastest computer available, it would take about six years! (Note, however, that Vincent's method can be quite efficient when the values of the partial quotients are small; for examples see [**15**].)

Our method, on the contrary, basically consists of computing a particular a_i as the lower bound b on the values of the positive roots of a polynomial equation. (It is assumed that $b=\lfloor \alpha_s \rfloor$ (the floor function or greatest integer function), where α_s is the smallest positive root.) This is achieved with the help of

Cauchy's Rule. *Let $P(x)=x^n+c_{n-1}x^{n-1}+\cdots+c_1x+c_0=0$ be a polynomial equation of degree n with integral coefficients, at least one of which is negative. If λ is the number of negative coefficients of $P(x)$, then*

$$b=\max_{\substack{1\leq k\leq n\\ c_{n-k}<0}} |\lambda c_{n-k}|^{1/k}$$

is an upper bound on the values of the positive roots of $P(x)=0$.

Proof. From the way b is defined we conclude that

$$b^k \geqslant \lambda |c_{n-k}|$$

for each k such that $c_{n-k}<0$; for these k's the last inequality can also be written as

$$b^n \geqslant \lambda |c_{n-k}| b^{n-k}.$$

Summing over all the appropriate k's we obtain

$$\lambda b^n \geqslant \lambda \sum_{\substack{k=1\\ c_{n-k}<0}}^{n} |c_{n-k}| b^{n-k}$$

or

$$b^n \geqslant \sum_{\substack{k=1\\ c_{n-k}<0}}^{n} |c_{n-k}| b^{n-k}.$$

From the last inequality we conclude that if we substitute b for x in $P(x)=0$, the first term, i.e.,

b^n, will be greater than or equal to the sum of the absolute values of all the negative coefficients. Therefore, $P(x) > 0$ for all $x > b$.

Observe that computing the lower bound b of $P(x)=0$ is equivalent to computing the upper bound on the values of the positive roots of $P(1/x)=0$. It might be thought that Cauchy's rule requires a great amount of computation, since it seems that the calculation of kth roots is needed. This, however, is not true because instead of computing each kth root we compute, very efficiently, the smallest integer $m(k)$ such that

$$|\lambda c_{n-k}|^{1/k} \leq 2^{m(k)},$$

and then we set $b = 2^{K+1}$, where K is the maximum of the $m(k)$'s. For details see [**5**].

Once we have computed $a_i \leftarrow b, b \geq 1$, we need to perform only one translation, namely, $\tilde{P}(x) \leftarrow \tilde{P}(x+b)$ which takes the same amount of time as $\tilde{P}(x) \leftarrow \tilde{P}(x+1)$ [**7**]; therefore, with our method we have enormous savings of computing time, and the previous example is solved in a matter of a few seconds. In what follows we present a recursive definition of our method as found in [**3**]:

Let

$$P(x) = 0 \tag{3}$$

be a polynomial equation without multiple roots and with v sign variations in the sequence of its integer coefficients.

If $v=0$ *or* $v=1$: From the Cardano-Descartes rule of signs we know that $v=0$ implies that (3) has no positive roots, whereas $v=1$ indicates that (3) has exactly one positive root, in which case $(0,\infty)$ is its isolating interval; in either case, no transformation of (3) is necessary, and the method terminates.

If $v > 1$: In this case (3) has to be further investigated. We first compute the lower bound b on the values of the positive roots and then we obtain the translated equation $P_b(x) = P(x+b) = 0$, which also has v sign variations provided $P(b) \neq 0$ (if $P(b)=0$, we have found an integer root of the original equation and v is decreased). The equation $P_b(x)=0$ is now transformed by the substitutions $x \leftarrow x+1$ and $x \leftarrow 1/(x+1)$, and the procedure is applied again twice, once with $P_b(1/(x+1))=0$ in place of (3) and once with $P_b(x+1)=0$.

We have implemented our method in a computer algebra system (for a detailed description of the algorithms see [**2**]) and have been able to show that its computing time bound is

$$O\left(n^5 L(|P|_\infty)^3\right),$$

which is the fastest obtained so far when exact integer arithmetic algorithms are used.

TABLES 1 and 2 show the observed computing times for the methods of Sturm, Vincent, and ours for certain classes of polynomials. All times are in seconds and were obtained using the SAC-1 computer algebra system on the IBM S/370 computer, located at the Triangle Universities Computation Center (North Carolina), where a subroutine CCLOCK is available which reads the computer clock [**3**]. TABLE 1 clearly indicates that, for this class of polynomials, Sturm's method is completely out of the race, whereas TABLE 2 makes clear the exponential nature of Vincent's method.

Polynomials with Randomly Generated Coefficients		
	Computation Time	
Degree	Sturm	Our Method
5	2.05	.26
10	33.28	.48
15	156.40	.94
20	524.42	2.36

TABLE 1

Polynomials of Degree 5 with Randomly Generated Roots		
	Computation Time	
Roots are in the Interval	Vincent	Our Method
$(0,10^2)$	.45	.16
$(0,10^3)$	1.61	.71
$(0,10^4)$	16.43	2.01
$(0,10^5)$	175.62	4.81

TABLE 2

References

[1] A. G. Akritas, A new method for polynomial real root isolation, Proceedings of the 16th Annual Southeast Regional ACM Conference, Atlanta, GA (1978) 39–43.

[2] ______, The fastest exact algorithms for the isolation of the real roots of a polynomial equation, Computing, 24 (1980) 299–313.

[3] ______, An implementation of Vincent's theorem, Numer. Math., 36 (1980) 53–62.

[4] ______, Vincent's forgotten theorem, its extension and application, Internat. J. Comput. Math. with Applications, 7 (1981) 309–317.

[5] ______, Exact algorithms for the implementation of Cauchy's rule, Internat. J. Comput. Math., 9 (1981) 323–333.

[6] A. G. Akritas and S. D. Danielopoulos, On the forgotten theorem of Mr. Vincent, Historia Math., 5 (1978) 427–435.

[7] ______, On the complexity of algorithms for the translation of polynomials, Computing, 24 (1980) 51–60.

[8] M. Bôcher, The published and unpublished works of Charles Sturm on algebraic and differential equations, Bull. Amer. Math. Soc., 18 (1911) 1–18.

[9] L. E. Dickson, First Course in the Theory of Equations, Wiley, New York, 1922.

[10] F. Cajori, A history of the arithmetical methods of approximation to the roots of numerical equations of one unknown quantity, Colorado College Publications, General Series No. 51, Science Series vol. XII, no. 7 (1910) 171–215, Colorado Springs, CO.

[11] L. E. Heindel, Integer arithmetic algorithms for polynomial real zero determination, J. Assoc. Comput. Mach., 18 (1971) 533–548.

[12] A. Hurwitz, Über den Satz von Budan-Fourier, Math. Ann., 71 (1912) 584–591.

[13] N. Obreschkoff, Verteilung und Berechnung der Nullstellen reeller Polynome, VEB Deutscher Verlag der Wissenschaften, Berlin, 1963.

[14] S. R. Petricle, ed., Proceedings of the 2nd symposium on symbolic and algebraic manipulation, ACM, 1971.

[15] J. V. Uspensky, Theory of Equations, McGraw-Hill, New York, 1948.

[16] A. J. H. Vincent, Sur la résolution des équations numériques, J. Math. Pures Appl., 1 (1836) 341–372.

[17] L. Weisner, Introduction to the Theory of Equations, MacMillan, New York, 1938.

The Last Duel

Hector's paces

were vector spaces.

—RICHARD MOORE

Polynomials of Degree 5 with Randomly Generated Roots		
	Computation Time	
Roots are in the Interval	Vincent	Our Method
$(0,10^2)$	.45	.16
$(0,10^3)$	1.61	.71
$(0,10^4)$	16.43	2.01
$(0,10^5)$	175.62	4.81

TABLE 2

References

[1] A. G. Akritas, A new method for polynomial real root isolation, Proceedings of the 16th Annual Southeast Regional ACM Conference, Atlanta, GA (1978) 39–43.

[2] ______, The fastest exact algorithms for the isolation of the real roots of a polynomial equation, Computing, 24 (1980) 299–313.

[3] ______, An implementation of Vincent's theorem, Numer. Math., 36 (1980) 53–62.

[4] ______, Vincent's forgotten theorem, its extension and application, Internat. J. Comput. Math. with Applications, 7 (1981) 309–317.

[5] ______, Exact algorithms for the implementation of Cauchy's rule, Internat. J. Comput. Math., 9 (1981) 323–333.

[6] A. G. Akritas and S. D. Danielopoulos, On the forgotten theorem of Mr. Vincent, Historia Math., 5 (1978) 427–435.

[7] ______, On the complexity of algorithms for the translation of polynomials, Computing, 24 (1980) 51–60.

[8] M. Bôcher, The published and unpublished works of Charles Sturm on algebraic and differential equations, Bull. Amer. Math. Soc., 18 (1911) 1–18.

[9] L. E. Dickson, First Course in the Theory of Equations, Wiley, New York, 1922.

[10] F. Cajori, A history of the arithmetical methods of approximation to the roots of numerical equations of one unknown quantity, Colorado College Publications, General Series No. 51, Science Series vol. XII, no. 7 (1910) 171–215, Colorado Springs, CO.

[11] L. E. Heindel, Integer arithmetic algorithms for polynomial real zero determination, J. Assoc. Comput. Mach., 18 (1971) 533–548.

[12] A. Hurwitz, Über den Satz von Budan-Fourier, Math. Ann., 71 (1912) 584–591.

[13] N. Obreschkoff, Verteilung und Berechnung der Nullstellen reeller Polynome, VEB Deutscher Verlag der Wissenschaften, Berlin, 1963.

[14] S. R. Petricle, ed., Proceedings of the 2nd symposium on symbolic and algebraic manipulation, ACM, 1971.

[15] J. V. Uspensky, Theory of Equations, McGraw-Hill, New York, 1948.

[16] A. J. H. Vincent, Sur la résolution des équations numériques, J. Math. Pures Appl., 1 (1836) 341–372.

[17] L. Weisner, Introduction to the Theory of Equations, MacMillan, New York, 1938.

The Last Duel

Hector's paces

were vector spaces.

—RICHARD MOORE

PROBLEMS

LEROY F. MEYERS, Editor
G. A. EDGAR, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before April 1, 1983.

1154. Prove the combinatorial identity

$$\sum_{k=0}^{n-1}\binom{n-1}{k}(-1)^k\frac{1}{(k+1)^2}=\frac{1}{n}H_n,\quad n\geqslant 1,$$

where

$$H_n=\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}.$$

[*Chico Problem Group, California State University.*]

1155. A plane intersects a sphere forming two spherical segments. Let S be one of these segments and let A be the point of the sphere furthest from the segment S. Prove that the length of the tangent from A to a variable sphere inscribed in the segment S is a constant. [*Stanley Rabinowitz, Merrimack, New Hampshire.*]

1156. Let a, b, c, and F be the three sides and the area of triangle ABC, and let a', b', c', and F' be the corresponding quantities for triangle $A'B'C'$. Show that

$$a'(-a+b+c)+b'(a-b+c)+c'(a+b-c)\geqslant\sqrt{48FF'},$$

with equality if and only if both triangles are equilateral. [*Gao Ling, Chongqing, China.*]

1157. The interior surface of a wine glass is a right circular cone. The glass contains some wine and is tilted so that the wine-to-air interface is an ellipse of eccentricity e and is at right angles to a generator of the cone.

Prove that the area of the ellipse is e times the area of that part of the curved surface of the cone which is in contact with the wine. [*R. C. Lyness, Southwold, Suffolk, England.*]

ASSISTANT EDITORS: DANIEL B. SHAPIRO and 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 Leroy F. Meyers, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.*

1158. Set $a_0 = 1$ and for $n > 1$, $a_n = a_{n'} + a_{n''}$, where $n' = [n/2]$ and $n'' = [n/3]$. Find $\lim_{n\to\infty} a_n/n$. [*Anon, Erewhon-upon-Spanish River.*]

1159. Phone books, n in number, are kept in a stack. The probability that the book numbered i (where $1 \leqslant i \leqslant n$) is consulted for a given phone call is $p_i > 0$, where $\Sigma_{i=1}^n p_i = 1$. After a book is used, it is placed at the top of the stack. Assuming that the calls are independent and evenly spaced, and that the system has been employed indefinitely far into the past, let P be the probability that, right now, each book is in its proper place, the book numbered i being ith from the top for $1 \leqslant i \leqslant n$. Unfixing the p_i's, find the least upper bound of P. [*James Propp, student, Cambridge University.*]

1160. In their "completion" of my solution to problem 1129 [pp. 304–305, this issue], the editors claim: "Elementary but tedious calculation shows that the function f defined by

$$f(x) = \frac{1}{\sin x} - \frac{1}{x}, \quad 0 < x \leqslant \frac{\pi}{2},$$

is positive and increasing." Justify the claim. [*Anon, Erewhon-upon-Spanish River.*]

Quickies

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

Q677. Let P be a polynomial of degree $n \geqslant 2$ with real coefficients:

$$P(x) = ax^n + bx^{n-1} + cx^{n-2} + \cdots, \text{ where } a \neq 0.$$

Show that P has at most $n-2$ distinct real zeros if $b^2 - \frac{2n}{n-1}ac < 0$. [*Herbert L. Holden, Gonzaga University.*]

Q678. It is well known that if f is continuous and of period T, then

$$\int_a^{a+T} f(t)\, dt = \int_0^T f(t)\, dt.$$

Prove this without the usual device of splitting the interval into parts. [*Henry E. Fettis, Mountain View, California.*]

Q679. Let S_n be the sum of the digits of 2^n. Prove or disprove that $S_{n+1} = S_n$ for some positive integer n. [*M. S. Klamkin, University of Alberta & M. R. Spiegel, East Hartford, Connecticut.*]

Solutions

Summing Units **November 1980**

1111. Let R be a finite ring. Evaluate $\Sigma_{r\in R^*} r$ and $\Sigma_{r\in R} r$, where R^* is the group of units of R. [*Douglas Lewan, Brown University.*]

Solution: (a) Let F_2 be the field with two elements.

PROPOSITION 1. *Let R be a finite (associative) ring (possibly without 1). If there exists a unique $\bar{r} \in R$ such that $\bar{r} \neq 0$ and $\bar{r} + \bar{r} = 0$, then $\Sigma_{r \in R} r = \bar{r}$. Otherwise $\Sigma_{r \in R} r = 0$.*

Proof. It is easy to see that if V is a finite dimensional vector space over F_2, then $\Sigma_{v \in V} v = 0$ if $\dim V \neq 1$, and $\Sigma_{v \in V} v = \bar{v}$, where $\bar{v}$ is the unique nonzero element of V, if $\dim V = 1$ (in fact $\Sigma_{v \in V} v$ is a vector of V fixed by every automorphism of V). Now let $G = \{r \in R | r + r = 0\}$. Then R is the disjoint union of G and the sets $\{r, -r\}$ where r ranges over $R \setminus G$. Hence $\Sigma_{r \in R} r = \Sigma_{r \in G} r$. But G is an additive subgroup of R (it is a vector space over F_2!). The conclusion immediately follows.

ALBERTO FACCHINI
Università di Padova, Italy

Solution: (b) Let R be a finite (associative) ring with 1 and let R^* be its group of units. Define $s(R) = \Sigma_{r \in R^*} r$. We begin with three examples where $s(R) \neq 0$.

(i) $R = F_2$. Here $|R| = 2$ and $s(R) = 1$.
(ii) $R = F_2[x]/(x^2)$. Here $|R| = 4$ and $s(R) = x + (x^2)$.
(iii) $R = T_2$, the ring of upper triangular 2×2 matrices over F_2. Here $|R| = 8$ and $s(R) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

PROPOSITION 2. *Let R be a finite associative ring with 1. Then $s(R) \neq 0$ if and only if R is isomorphic to a direct product $R_1 \times K_2 \times \cdots \times K_m$ for some $m \geqslant 1$, where each K_i is a finite field of characteristic 2, and R_1 is one of the three examples listed above.*

Before beginning the proof, we establish several lemmas. We will use classical results about rings with minimum condition, as presented in Chapters 4 and 6 of [**1**], for example. Note that an *ideal* of R is a two-sided ideal.

LEMMA 1. *If R has characteristic $\neq 2$, then $s(R) = 0$.*

Proof. Since $2 \neq 0$, we have $2r \neq 0$ for each $r \in R^*$, and hence $r \neq -r$. Then R^* is a disjoint union of subsets $\{r, -r\}$, so that $s(R) = 0$. ■

From now on we will always assume that the ring R has characteristic 2, so that R is an F_2-algebra.

A nonzero ideal $A \subseteq R$ is said to be *indecomposable* if it cannot be expressed as a direct product $A_1 \times A_2$ of nonzero ideals A_1, A_2. It follows that our ring R is a direct product of certain indecomposable ideals: $R = R_1 \times \cdots \times R_m$ for some $m \geqslant 1$.

LEMMA 2. *Suppose that $R = R_1 \times \cdots \times R_m$ is a direct product of $m \geqslant 2$ rings. Then $s(R) \neq 0$ if and only if for some k, $s(R_k) \neq 0$, but $|R_i^*|$ is odd for every $i \neq k$.*

Proof. From the definition we see that

$$s(R) = (n_1 \cdot s(R_1), \ldots, n_m \cdot s(R_m)),$$

where $n_i = |R^*|/|R_i^*|$. Then $s(R) \neq 0$ if and only if for some k, $n_k \cdot s(R_k) \neq 0$, i.e., $s(R_k) \neq 0$ and n_k is odd. ■

LEMMA 3. *If R is a simple ring and $s(R) \neq 0$, then $R \cong F_2$.*

Proof. By Wedderburn's structure theorem, $R \cong M_n(K)$ is a matrix ring over a skew field K. Wedderburn's theorem on finite skew fields implies that K is a field. If $n \geqslant 2$, let $e_{ij} \in R$ be the matrix with 1 as the (i, j) entry and zeros elsewhere. Generally, if $r \in R^*$, then $r \cdot s(R) = s(R) = s(R) \cdot r$. Since $1 + e_{ij} \in R^*$ whenever $i \neq j$, it follows that

$$e_{ij} \cdot s(R) = 0 = s(R) \cdot e_{ij}.$$

But if $s(R) \neq 0$, there exists some $i \neq j$ such that $e_{ij} \cdot s(R) \neq 0$. Hence $n = 1$ and $R \cong K$. If $r \in K^*$, then $r \cdot s(R) = s(R)$ implies $r = 1$. Hence $K \cong F_2$. ∎

Let J be the radical of R. Then J is the maximal nilpotent (two-sided) ideal, and the quotient ring $\bar{R} = R/J$ is semisimple. Now if $j \in J$, then $j^k = 0$ for some k and

$$(1-j)(1+j+\cdots+j^{k-1}) = 1,$$

so that $1 + j = 1 - j \in R^*$. Hence if also $r \in R^*$, then

$$r + j = r(1 + r^{-1}j) \in R^*.$$

LEMMA 4. *If $J \neq 0$, then $|R^*|$ is even. If in addition $s(R) \neq 0$, then $|J| = 2$ and $s(R)$ is the unique nonzero element of J.*

Proof. Let j be a fixed nonzero element of J. Then $J_0 = \{0, j\}$ is an F_2-subspace of R, and R breaks into a disjoint union of congruence classes mod J_0. Since $R^* = R^* + j$, we know that R^* is the disjoint union of some n of these congruence classes, each class of the form $\{r, r+j\}$. Then $|R^*| = 2n$ is even, and $s(R) = n \cdot j$. If $s(R) \neq 0$, then n must be odd and $s(R) = j$. Since j was any nonzero element of J, it follows that j is unique and $|J| = 2$. ∎

LEMMA 5. *If $|R^*|$ is odd and R is indecomposable, then R is a field.*

Proof. By Lemma 4, $J = 0$, so that R is semisimple. Since R is indecomposable, R is simple and hence by Wedderburn's theorems, $R \cong M_n(K)$ for some field K. If $n \geqslant 2$, there exist nonscalar matrices $u \in R$ with $u^2 = 1$. Then R^* has elements of order 2, so that $|R^*|$ is even. Hence $n = 1$ and $R \cong K$. ∎

The radical J of R can also be characterized in terms of the maximal (two-sided) ideals. We state this formally here, since it is used later.

LEMMA 6. *J is the intersection of the maximal ideals of R.*

Proof. Since a maximal ideal is prime and $J^2 = 0$, it follows that J is contained in each maximal ideal. Hence it suffices to pass to R/J and note that in a semisimple ring the intersection of the maximal ideals is zero. This is immediate from Wedderburn's structure theorem. ∎

Proof of Proposition 2. From the three examples and Lemma 2, we find that the rings R listed in the proposition do have $s(R) \neq 0$. Conversely, suppose $s(R) \neq 0$. Then R has characteristic 2, by Lemma 1. Express R as a direct product of its indecomposable (two-sided) ideals: $R = R_1 \times \cdots \times R_m$. By Lemma 2 we may re-index to assume $s(R_1) \neq 0$ and $|R_i^*|$ is odd for every $i > 1$. Then Lemma 5 says that each $R_i = K_i$ is a field for $i > 1$. It remains to show that R_1 is isomorphic to one of the three examples.

Dropping the subscript, we have an indecomposable ring R with $s(R) \neq 0$. If $J = 0$, then R is simple, and Lemma 3 says that $R \cong F_2$. Suppose that $J \neq 0$, so that $J = \{0, j\}$ as in Lemma 4. Let A and B be the left and right annihilators of J, respectively. Since J is an ideal, it follows that A and B are ideals. The action of R on the left R-module J yields a ring homomorphism $R \to \mathrm{End}_{F_2}(J)$ with kernel A. Since J is a 1-dimensional vector space over F_2, we have $\mathrm{End}_{F_2}(J) \cong F_2$. Therefore $R/A \cong F_2$ as rings, and A is a maximal ideal. A similar result holds for B. We will prove next that A and B are the only maximal ideals of R.

Suppose that $M_1, \ldots, M_t$ are the other maximal ideals, and let $P = M_1 \cdots M_t$ be their product. By Lemma 6 we have

$$J = M_1 \cap \cdots \cap M_t \cap A \cap B.$$

Then $PAB \subseteq J$ and $ABP \subseteq J$, and therefore $PABAB \subseteq JB = 0$ and $ABABP \subseteq AJ = 0$. Setting $Q = ABAB$, we have $PQ = QP = 0$ and $P + Q = R$ (for $P + Q$ cannot be inside any maximal ideal). Then

$$P \cap Q = (P \cap Q)(P + Q) \subseteq QP + PQ = 0.$$

The Chinese remainder theorem now implies that

$$R \cong R/(P \cap Q) \cong (R/P) \times (R/Q).$$

Since R is indecomposable and $Q = ABAB \neq R$, we must have $P = R$, i.e., no other maximal ideals M_i can exist.

Now Lemma 6 says $A \cap B = J$. If $A = B$, then $A = B = J$ is the unique maximal ideal of R, and has index 2. Since $|J| = 2$, we have $|R| = 4$ and

$$R = \{0, 1, j, 1 + j\} \cong F_2[x]/(x^2),$$

as in example (ii). Suppose $A \neq B$. Then $A + B = R$ and the Chinese remainder theorem implies

$$R/J \cong (R/A) \times (R/B) \cong F_2 \times F_2.$$

Hence $|R| = 8$, so that $|A| = |B| = 4$. The action of R on the left R-module A yields a ring homomorphism ϕ: $R \to \mathrm{End}_{F_2}(A)$. Since A is 2-dimensional over F_2 and contains J, we can choose an F_2-basis $\{j, a\}$ of A. Relative to this basis we can view ϕ as ϕ: $R \to M_2(F_2)$ and note that the image $\phi(R)$ is a subring of T_2, the ring of upper triangular matrices in $M_2(F_2)$. (This follows since $J = \{0, j\}$ is an ideal and $aj = 0$.)

Let $I = \ker \phi$, so that I is an ideal of R with $IA = 0$. Since $J \subseteq A$, we see that $IJ = 0$ and hence $I \subseteq A$. Then $I^2 \subseteq IA = 0$, so that I is a nilpotent ideal and therefore $I \subseteq J$. If $I = J$, then $JA = 0$, forcing $A \subseteq B$ and hence $A = B$, contrary to hypothesis. Therefore $I = 0$ and ϕ is injective. Then $|\phi(R)| = |R| = 8$, and we conclude that $\phi(R) = T_2$. Thus ϕ provides the isomorphism $R \cong T_2$, as in example (iii). ■ ■

REMARK: Some steps in the argument above can be shortened by invoking the Wedderburn Principal Theorem. Thanks are due to A. Wadsworth for simplifying the proof, which is the completion of a partial solution submitted by A. Facchini.

Reference

[1] C. Curtis and I. Reiner, Representation Theory of Finite Groups and Associative Algebras, Wiley, 1962.

DANIEL B. SHAPIRO
The Ohio State University

Also solved (part (a)) by Roger Cuculière (France), F. J. Flanigan, G. A. Heuer & Karl Heuer, and the proposer. Partial solutions to part (b) were submitted by Cuculière, Alberto Facchini (Italy), Flanigan, and the Heuers.

Radial Shift **September 1981**

1128. Show that for each $\varepsilon > 0$ there exists a one-to-one mapping σ from the open disk $D = \{z: |z| < 1\}$ onto its closure $\bar{D}$ such that $|z - \sigma(z)| < \varepsilon$ for all z in D. [*Carl P. McCarty & Loretta McCarty, LaSalle College.*]

Solution: Let $(r_i)_{i=1,2,\ldots}$ be a decreasing sequence of positive real numbers satisfying

$$r_1 = 1, r_i - r_{i+1} < \varepsilon \text{ for all } i, \text{ and } \lim_{i \to \infty} r_i = 0.$$

Then the function σ defined by

$$\sigma(re^{i\theta}) = \begin{cases} r_{i-1}e^{i\theta} & \text{if } r = r_i,\ i = 2, 3, \ldots, \\ re^{i\theta} & \text{otherwise,} \end{cases}$$

satisfies the requirements.

HAROLD REITER
The University of North Carolina at Charlotte

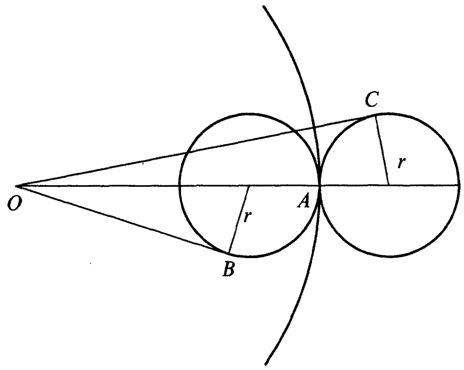
Also solved by F. S. Cater, W. M. Causey, Chico Problem Group, Jim Griffith, Victor Hernandez (Spain), G. A. Heuer, J. Pfaendtner (West Germany), Daniel A. Rawsthorne, Harry Sedinger, J. M. Stark, J. Suck (West Germany), David Vetterlein, Michael Woltermann, and the proposers.

Several solvers mentioned that the dimension of the space is irrelevant. Causey noted that no such function can be continuous.

Roller Bearings **September 1981**

1129. Find radii r and R, with $r < R$, so that m circles of radius r form a closed ring with each circle externally tangent to a circle of radius R, and n circles of radius r form a closed ring with each circle internally tangent to the circle of radius R. [*Anon, Erewhon-upon-Spanish River.*]

Solution (completed by the editors):



$OA = R$
$\angle AOB = \frac{\pi}{n}$
$\angle AOC = \frac{\pi}{m}$

From the figure it is clear that

$$\sin\frac{\pi}{n} = \frac{r}{R-r} \quad \text{and} \quad \sin\frac{\pi}{m} = \frac{r}{R+r},$$

so that

$$\frac{1}{\sin\frac{\pi}{m}} - \frac{1}{\sin\frac{\pi}{n}} = 2, \tag{1}$$

where m and n are integers such that $m > n \geqslant 2$.

For $n = 2$ we have $m \approx 9.244$, not an integer. For $m \geqslant 3$, numerical evidence seems to indicate that $m - n < 7$ and that $m - n$ decreases to 2π as $n \to \infty$.

The conjecture based on numerical evidence is in fact true, so that there is no solution in integers m, n with $2 \leqslant n < m$.

Assume $n \geqslant 3$, and consider the function f defined by

$$f(x) = \frac{1}{\sin x} - \frac{1}{x}, \quad 0 < x \leqslant \frac{\pi}{2}.$$

Elementary but tedious calculation shows that f is positive and increasing. (Ed. note: See proposal 1160, this issue, p. 300.) Now if m, n satisfy (1), then

$$f\left(\frac{\pi}{n}\right) > f\left(\frac{\pi}{n}\right) - f\left(\frac{\pi}{m}\right) = \frac{m-n}{\pi} - 2 > 0,$$

or

$$2\pi < m - n < 2\pi + \pi f\left(\frac{\pi}{n}\right).$$

Then

$$6 < 2\pi < m - n \leqslant 2\pi + \pi f\left(\frac{\pi}{3}\right) < 7,$$

so $m - n$ is not an integer. Since $\lim_{x\to 0} f(x) = 0$, this also shows $m - n \to 2\pi$ as $n \to \infty$.

ANON
Erewhon-upon-Spanish River

Also solved incompletely by Milton P. Eisner, Thomas E. Elsner, and J. M. Stark.

Sums of Abundant Numbers **September 1981**

1130. A positive integer is *abundant* if the sum of its proper divisors exceeds n (i.e., $\sigma(n) > 2n$). Show that every integer greater than 89×315 is the sum of two abundant numbers. [*J. L. Selfridge, Mathematical Reviews.*]

Solution: First note that if p is a prime not dividing 315, then

$$\sigma(315p) = (p+1)\sigma(315) = 624(p+1),$$

and so $315p$ is abundant if and only if $p < 104$. Also, $315p$ is abundant if p is 3, 5, or 7. Since every multiple of an abundant number is abundant, it follows that $315y$ is abundant for $1 < y < 107$.

Suppose $n > 89 \times 315$. Let $n = 315a + b$ with $0 \leqslant b \leqslant 314$. Then $a \geqslant 89$. Since 88 and 315 are relatively prime, let x be the unique solution of $b + 315x \equiv 0 \pmod{88}$ satisfying $0 \leqslant x < 88$. Then $a - x > 1$. Choose k so that $1 < a - x - 88k < 107$. Then

$$n = 315a + b = 315(a - x - 88k) + (b + 315x + 315 \cdot 88k).$$

Now $315(a - x - 88k)$ is abundant since $1 < a - x - 88k < 107$, and $b + 315x + 315 \cdot 88k$ is abundant since it is a multiple of the abundant number 88. Hence n is the sum of two abundant numbers.

Actually, it can be shown that $20161 = 315 \times 64 + 1$ is not the sum of two abundant numbers, but every integer larger than it is such a sum.

CHARLES R. WALL
Trident Technical College

Also solved by Lorraine L. Foster, John P. Robertson, J. M. Stark, J. Suck (West Germany), and the proposer (two solutions).

The proposer noted that this problem is an improvement on Leo Moser's estimate $n > 100000$, which appeared as problem E848 in the *American Mathematical Monthly*, vol. 56 (1949), pp. 31, 478. Bob Prielipp noted that in *Excursions in Number Theory*, by Olgivy and Anderson (1972), pp. 23–24, it is stated that T. R. Parkin and L. J. Lander had solved the problem completely in their report *Abundant Numbers*, Aerospace Corporation, California (1964). The problem is proposed in *Tomorrow's Math*, by Olgivy, first ed. (1962), pp. 93–94; the second edition (1972), p. 181, merely refers to *Excursions*. J. Suck noted a similar problem on pp. 125-129 of Ross Honsberger, *Ingenuity in Mathematics* (NML vol. 23), wherein it is shown that every even integer exceeding 46 is the sum of two abundant numbers. Robertson also proved this, and confirmed with computer printout the last claim in Wall's solution above.

Quadratic Residues **September 1981**

1131. Every odd prime is of the form $p = 4n \pm 1$.

(a) Show that n is a quadratic residue $\pmod{p}$.

(b) Calculate the value of $n^n \pmod{p}$. [*Oren N. Dalton, El Paso, Texas.*]

Solution I: Write $p=4n+e$, where $e=\pm 1$. Let $\left(\frac{-}{p}\right)$ denote the Legendre symbol.

(a)
$$\left(\frac{n}{p}\right)=\left(\frac{4n}{p}\right)=\left(\frac{-e}{p}\right)=\left(\frac{-1}{p}\right)\left(\frac{e}{p}\right)$$
$$=\begin{cases}\left(\frac{-1}{p}\right)=1 & \text{if } e=1,\\ \left(\frac{-1}{p}\right)^2=1 & \text{if } e=-1.\end{cases}$$

Thus n is a quadratic residue (mod p).

(b)
$$(-e)^n\equiv(4n)^n\equiv 2^{2n}n^n\ (\mathrm{mod}\ p).$$

CASE 1: $e=1$. Then $2n=(p-1)/2$ and
$$2^{2n}\equiv\left(\frac{2}{p}\right)\equiv(-1)^{(p^2-1)/8}\equiv(-1)^{n(2n+1)}\equiv(-1)^n\ (\mathrm{mod}\ p).$$

Hence $n^n\equiv 1$ (mod p).

CASE 2: $e=-1$. Then
$$2^{2n}n^n\equiv 1\equiv p+1\ (\mathrm{mod}\ p)$$
and
$$\left(\frac{2}{p}\right)\equiv 2^{(p-1)/2}\equiv 2^{2n-1}\ (\mathrm{mod}\ p).$$
Therefore
$$\left(\frac{2}{p}\right)n^n\equiv 2^{2n-1}n^n\equiv(p+1)/2\ (\mathrm{mod}\ p)$$
and $n^n\equiv\left(\frac{2}{p}\right)(p+1)/2$ (mod p).

If $p\equiv 3$ (mod 8), then $\left(\frac{2}{p}\right)=-1$ and
$$n^n\equiv(p-1)/2\equiv 2n-1\ (\mathrm{mod}\ p).$$
If $p\equiv 7$ (mod 8), then $\left(\frac{2}{p}\right)=1$ and
$$n^n\equiv(p+1)/2\equiv 2n\ (\mathrm{mod}\ p).$$

STEVE GALOVICH
Carleton College

Solution II: Generalization of part (a). Let $p=4n+e$ be prime, where $e=\pm 1$, and let k be a divisor of n. Then k is a quadratic residue (mod p).

Proof. Let $k=2^s m$, where m is odd and $s\geqslant 0$. If $s\geqslant 1$, then $p\equiv e$ (mod 8) and
$$\left(\frac{2}{p}\right)=(-1)^{(p^2-1)/8}=1.$$
Hence by the properties of the Legendre and Jacobi symbols, if $s\geqslant 0$, then
$$\left(\frac{k}{p}\right)=\left(\frac{2}{p}\right)^s\cdot\left(\frac{m}{p}\right)=\left(\frac{m}{p}\right)=(-1)^{\frac{m-1}{2}\cdot\frac{p-1}{2}}\left(\frac{e}{m}\right).$$

If $e = 1$, then $(p-1)/2$ is even and $\left(\frac{k}{p}\right) = 1$. If $e = -1$, then $(p-1)/2$ is odd and

$$\left(\frac{k}{p}\right) = (-1)^{(m-1)/2} \cdot (-1)^{(m-1)/2} = 1.$$

In either case, k is a quadratic residue (mod p).

LORRAINE L. FOSTER
California State University
Northridge

Also solved by Kenneth A. Brown, Jr., Chico Problem Group (2 solutions), Peter Flusser, Lorraine L. Foster (second solution of both parts), Enzo R. Gentile (Argentina), Reinaldo E. Giudici (Venezuela), W. C. Igips, Robert E. Kennedy & Curtis N. Cooper, L. Kuipers (Switzerland), Daniel A. Rawsthorne, John P. Robertson, Sahib Singh, Lawrence Somer, J. M. Stark, L. Van Hamme (Belgium), B. Viswanathan (Canada), Michael Woltermann, and the proposer.

For related material, see Lawrence Somer, The residues of n^n modulo p, The Fibonacci Quarterly, vol. 19 (1981), pp. 110-117.

Answers

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

Q677. For $n = 2$, the result is known from elementary algebra. Suppose that the result holds for polynomials of degree $n - 1$, and that P is a polynomial of degree n such that

$$\Delta(P) = b^2 - \frac{2n}{n-1}ac < 0.$$

Then

$$\Delta(P') = (n-1)^2 b^2 - \frac{2(n-1)}{n-2} na(n-2)c = (n-1)^2 \Delta(P) < 0,$$

so that P' has at most $n - 3$ distinct real zeros, by the induction hypothesis. Then by Rolle's theorem, P has at most $n - 2$ distinct real zeros.

It is interesting that a sufficient condition for the "loss" of two real zeros can be obtained from only the first three coefficients of the polynomial.

Q678. Differentiate with respect to a:

$$\frac{d}{da}\int_a^{a+T} f(t)\,dt = f(a+T) - f(a) = 0.$$

Hence $\int_a^{a+T} f(t)\,dt$ is independent of a.

Q679. Let $S(k)$ be the sum of the digits in the base-ten representation of the positive integer k. Then $k \equiv S(k)$ (mod 9). Hence if $S(2^{n+1}) = S(2^n)$, then

$$2^n = 2^{n+1} - 2^n \equiv S(2^{n+1}) - S(2^n) \equiv 0 \pmod 9,$$

which is impossible, since $9 \nmid 2^n$.

REVIEWS

Paul J. Campbell, Editor
Beloit College

Pierre J. Malraison, Jr., Editor
MDSI, Ann Arbor

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

Hofstadter, Douglas R., *Metamagical themas: beyond Rubik's Cube: spheres, pyramids, dodecahedrons and God knows what else*, Scientific American 247:1 (July 1982) 16-29, 154.

Descriptions of other puzzles motivated by Rubik's original, with an attempt by Hofstadter to characterize the qualities of the cube that make it so popular.

Thomsen, Dietrick E., *A place in the sun for fractals*, Science News 121 (9 January 1982) 28-30.

Fractals started out as a geometrical curiosity; their latest areas of application, plasma physics and astrophysics, are informally described here.

Abelson, Philip H., and Dorfman, Mary (eds.), *Computers and electronics*, 215 (12 February 1982) 749-873.

Special issue of *Science* devoted to 21 articles on the topics of computers and electronics, graphics and software, scientific research, business and industry, communications and personal services, and information storage and retrieval.

Rothstein, Edward, *Math and Music: The Deeper Links*, The New York Times (29 August 1982, Section 2) 1, 22.

Mathematical elements of music and musical elements of mathematics are surveyed. Contains some non-trivial examples and incisive quotes by musicians and mathematicians.

Bradley, R., *et al.*, (eds.), Case Studies in Mathematical Modeling: A Course Book for Scientists and Engineers, Wiley, 1981; ii + 167.

Results of a working conference bringing together industrialists and academics. After an introduction to modeling of a blast furnace, groups worked on problems to produce the reports published here. Each study is reported in a very satisfying format; introduction by the leader (background, useful data), written report for the group (commentary, the group's model), and the leader's report (commentary on the group's process and progress, the leader's model). Topics include minimization of sound distortion in a record player, performance of hydraulic buffers, grinding of particulate solids, roller spacing in a glass process, bond graph models, and the freezing and thawing of meat. The importance of background in physics emerges from *every* case.

Jacobs, Harold R., Mathematics: A Human Endeavor: A Book for Those Who Think They Don't Like the Subject, 2nd ed., Freeman, 1982; xiii + 649 pp, $17.95.

Although the overall structure of Jacobs' best-selling come-on to mathematics is unchanged, additional examples, graphics, and exercises have expanded the volume by almost 25%. The book has been reworked sentence-by-sentence, and a second color is used to highlight most figures. May the new edition have even greater success than the first in combating fear and ignorance of mathematics!

Pólya, George, Mathematical Discovery: On Understanding, Learning, and Teaching Problem Solving, Combined Edition, Wiley, 1981; xxv + 432 pp, $20.95 (P).

Timely reissue in one volume of Pólya's marvelous two-volume primer on teaching problem-solving. After 20 years it retains its freshness and compelling insight. This edition features a foreword by Peter Hilton, an updated bibliography, and an expanded index.

Gardner, Martin, aha! Gotcha: Paradoxes to Puzzle and Delight, Freeman, 1982; vii + 164 pp.

Humorous and engaging collection of dozens of puzzles and paradoxes from logic, probability, numbers, geometry, time, and statistics. Similar to the author's *aha! Insight* (1978), this volume is derived from *The Paradox Box* (1975), a set of filmstrips, cassettes, and teacher's guides.

Gardner, Martin, Science: Good, Bad and Bogus, Prometheus, 1981; xvii + 412 pp, $18.95.

Even though there is only one piece of this book touching directly on mathematics, mathematician fans of Gardner will likely also enjoy his skeptical examination of the many varieties of pseudoscience abroad in the world. The book contains 18 articles and 20 reviews, most of them reprinted and accompanied by further comments or correspondence; it may be regarded as a sequel to Gardner's famous *Fads and Fallacies in the Name of Science*.

Davis, Chandler, *et al.*, The Geometric Vein: The Coxeter Festschrift, Springer-Verlag; viii + 598 pp.

Collection of papers to honor H. S. M. Coxeter, given at Toronto in 1979. Includes list of Coxeter's published works to date, and 39 papers grouped into: polytopes and honeycombs, extremal problems, geometric transformations, groups and presentations of groups, and "the combinatorial side."

Wilder, Raymond L., Mathematics as a Cultural System, Pergamon, 1981; xii + 182 pp, $23.

Extension and elaboration of the author's *Evolution of Mathematical Concepts* (1968), in a work intended for philosophers and social scientists as well as mathematicians. Mathematics is looked at from a "culturological" point of view, and examples given of cultural patterns and forces observable in its history. The book culminates in the formulation of 23 "laws" governing the evolution of mathematics.

Muzzio, Douglas, Watergate Games: Strategies, Choices, Outcomes, NYU Pr., 1982; xii + 205 pp, $17.95.

Extended analysis of events of the Watergate scandal, setting out options and their consequences for the principals at key points. At these points interactions produce games: the conspiracy breakdown game, the Saturday night massacre game, and the White House tapes game. The author's conclusions that Nixon acted rationally in these circumstances, and that chance played a large part in his fall, are (unfortunately) not very consoling.

Bestgen, Barbara J., and Reys, Robert E., Films in the Mathematics Classroom, NCTM, 1982; vi + 90 pp, $6.75 (P).

Reviews of over 230 mathematics films, including some at the college level.

Anderssen, Robert S., and deHoog, Frank R., The Application of Mathematics in Industry, Martinus Nijhoff, 1982; xiv + 202 pp.

Descriptions of 13 instances of applied mathematical modeling in Australia, in agriculture (acceptance sampling of peanuts), heavy industry (blast furnaces), city planning (stormwater drainage), optics (contact lens grinding, optical fiber fabrication), life insurance (reserves), and other areas.

Hilton, Peter J., and Young, Gail S., New Directions in Applied Mathematics, Springer-Verlag, 1982; vii + 163 pp.

"My basic thesis is that the spectacular developments in pure mathematics in the last two decades, coupled with the immense powers of the computer, will revolutionize applications, and soon." For support for this claim, Young points to the seven papers in this book, which range from trends in combinatorics, through control theory and nonlinear analysis, to discrete applied mathematics. Most striking, perhaps, is Zeeman's account of a successful new treatment for hyperthyroidism suggested by elementary catastrophe theory. Hilton closes the volume with a plea for teaching abstraction, generalization and simplification as key processes against a false dichotomy between problem-solving and "theory."

Pless, Vera, Introduction to the Theory of Error-Correcting Codes, Wiley, 1982; xi + 169 pp.

Practical background, basic theory, and many examples of error-correcting codes. Suitable for study after a course in linear algebra.

Erdös, P., and Graham, R. L., Old and New Problems and Results in Combinatorial Number Theory, L'Enseignment Mathematique, 1980; 128 pp.

Another in the long series of number-theory-problem surveys from Erdös, this one centering on problems with a combinatorial flavor.

American Statistical Association, 1981 Proceedings of the Section on Statistical Education, v + 158 pp, $11 (P).

Papers presented at the 1981 ASA meeting, offering readable papers on teaching tips, the training needed for government statistical work, what to teach to health professionals, how to train consultants, strategies in teaching statistical literacy, and use of audio-visual and computer assistance.

Kolz, Samuel, and Johnson, Norman L. (eds.), Encyclopedia of Statistical Science, Vol. 1: A - Circular Probable Error, Wiley, 1982; x + 480 pp.

First volume of an eight-volume broadly definitive encyclopedia of statistics, with subsequent volumes to appear two per year. Articles are far greater in number, but shorter, than those of *The International Encyclopedia of Statistics* (1978).

Tarwater, Dalton, *et al.*, (eds.), American Mathematical Heritage: Algebra and Applied Mathematics, Texas Tech University Mathematics Series No. 13, 1981.

Although these eight papers date back to conferences in 1975 and 1976, the caliber of the authors and the talks merits attention: S. MacLane on the history of abstract algebra, W. Feit on theory of finite groups in the twentieth century, Peter Lax on applied mathematics 1945-1975, and others.

NEWS & LETTERS

LESTER R. FORD AWARDS

Two authors will be presented with Lester R. Ford awards at the business meeting of the MAA on January 9, 1983, in Denver. The awards of $200 each are in recognition of excellence in writing expository articles published in 1981 in the *American Mathematical Monthly*.
The recipients are:

Philip Davis, "Are there coincidences in mathematics?" *Monthly* 88 (1981) 311-320.

Arthur Knoebel, "Exponentials reiterated," *Monthly* 88 (1981) 235-252.

AUTHORS SOUGHT BY NCTM

Potential authors are invited to submit manuscripts for the National Council of Teachers of Mathematics 1985 Yearbook, which is tentatively titled *The Secondary School Mathematics Curriculum*. The thrust of the yearbook will be to reexamine the mathematics curriculum of grades 9 through 12, in light of the major changes in society's uses of and need for mathematics that have occurred in the last twenty years.

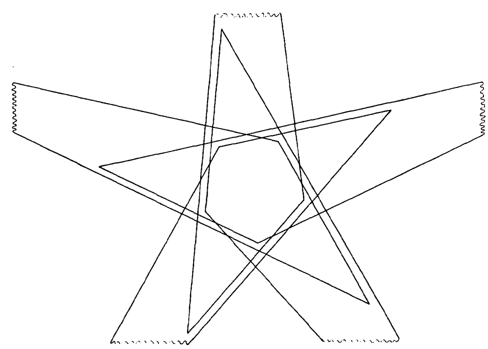
The yearbook editor is Christian R. Hirsch, Western Michigan University, Kalamazoo, Michigan, and the deadline for manuscripts is March 1, 1983.

Guidelines, which include a more complete description of the topics to be addressed and directions for preparing manuscripts, may be obtained from

Marilyn J. Zweng
General Yearbook Editor
N297 Lindquist Center
The University of Iowa
Iowa City, IA 52242

A VENN DIAGRAM OF 5 TRIANGLES

One of the unsolved problems mentioned in B. Grünbaum's "Venn diagrams and independent families of sets" (this *Magazine*, January 1975, 12-22) asked about the existence of a Venn diagram, formed by five triangles, which is *simple* (that is, in which no point is on the boundary of any three of the triangles). As it turns out, such a Venn diagram was constructed--already in 1970, long before the problem was posed--by P. Winkler; he used the diagram to decorate his offices at Ft. Meade, Stanford and Emory. In recent correspondence between Grünbaum and Winkler, the former constructed an even "better" simple Venn diagram. It consists of five triangles which are not only congruent, but are equivalent under symmetries of the diagram (see the Figure, where to show intersections more clearly, the triangles have been truncated with wavy lines.)



Remarkably, it has the same structure as the symmetric and simple Venn diagram formed by five ellipses, described in the earlier article. It is still not known whether there exist simple Venn diagrams consisting of six quadrangles. Another unsolved problem, first formulated by Winkler during a recent graph theory meeting at Oberwolfach, is the following: is every simple Venn diagram formed by n curves extendible to a simple Venn diagram formed by $n + 1$ curves?

Branko Grünbaum
University of Washington
Seattle, WA 98191

Peter Winkler
Emory University
Atlanta, GA 30322

TIED RACES

Readers of Mendelson's "Races with ties" (this *Magazine*, May 1982, pp. 170-175) who want to pursue the subject further will appreciate also I. J. Good's "The number of orderings of n candidates when ties are permitted," *The Fibonacci Quarterly* 13 (1975) 11-18, including some open questions.

J. Suck
Rahmstrasse 140
D 4300 Essen
W. Germany

THE THREE-CIRCLE THEOREM

John McCleary's Note "An Application of Desargues' Theorem" (this *Magazine*, Sept. 1982, pp. 233-235) has the merit of being simple and self-contained and yet thought-provoking. Other readers may like to share some of my thoughts.

(i) The figure in the article illustrates only one possible case. If the radius of the circle centred at A is sufficiently large, and if we still form the triangle $A'B'C'$ by using the tangents exterior to triangle ABC, then $B'B$ and $C'C$ will be external bisectors. Also in a special case A' can be at infinity. How do we show rigorously that we never get two internal bisectors and one external, or three external?

(ii) The circles can be allowed to intersect; and two of the circles can have equal radii.

(iii) If the circles do not intersect, we have also points of intersection of pairs of internal common tangents, X_1, Y_1, Z_1 say, and similar proofs show that X, Y_1, Z_1 are collinear, as are X_1, Y, Z_1 and X_1, Y_1, Z.

(iv) The points X, Y, Z, X_1, Y_1, Z_1 are the *centres of similitude* of the pairs of circles. These centres exist even if the circles intersect or lie one inside the other; McCleary's proof no longer applies, but the centres of similitude are still collinear, as can be proved using Menelaus' theorem.

(v) The results used in McCleary's proof seem to all be valid in *metric geometry*; hence the theorem remains true in the hyperbolic and elliptic planes, where the concept of centres of similitude is no longer valid and Menelaus' theorem is not true. What happens to Coxeter's proof in the hyperbolic and elliptic planes?

John Rigby
University College
Cardiff, CF1 1XL, Wales

XXIII INTERNATIONAL MATH OLYMPIAD

Teams of talented high school students from 30 nations competed in the XXIII IMO held July 9-10, 1982, in Budapest, Hungary. The ten high scoring teams are listed below with their team scores; the U.S. team tied for third place with East Germany. The four U.S. team members all won individual prizes: Noam D. Elkies won a first prize for a score of 40 (out of a possible 42), Brian Hunt and Washington Taylor IV won second prizes, and Douglas Jungreis won a third prize.

Team	*Score*
W. Germany	145
USSR	137
USA	136
E. Germany	136
Vietnam	133
Hungary	125
Czechoslovakia	115
Finland	113
Bulgaria	108
Great Britain	103

Three problems were given each day, with 4½ hours allotted time to solve them. The questions are given below; solutions will appear in our next issue.

PROBLEMS

1. The function $f(n)$ is defined for all positive integers n and takes on non-negative integer values. Also, for all m,n,

$$f(m+n) - f(m) - f(n) = 0 \text{ or } 1;$$
$$f(2) = 0,\ f(3) > 0,$$
$$\text{and } f(9999) = 3333.$$

Determine $f(1982)$.

2. A non-isosceles triangle $A_1A_2A_3$ is given with sides a_1, a_2, a_3 (a_i is the side opposite A_i). For all $i = 1,2,3$, M_i is the midpoint

of side a_i, T_i is the point where the incircle touches side a_i, and the reflection of T_i in the interior bisector of A_i yields the point S_i. Prove that the lines M_1S_1, M_2S_2, and M_3S_3 are concurrent.

3. Consider the infinite sequences $\{x_n\}$ of positive real numbers with the following properties:

$$x_0 = 1 \text{ and for all } i \geq 0, \quad x_{i+1} \leq x_i .$$

(a) Prove that for every such sequence, there is an $n \geq 1$ such that

$$\frac{x_0^2}{x_1} + \frac{x_1^2}{x_2} + \ldots + \frac{x_{n-1}^2}{x_n} \geq 3.999.$$

(b) Find such a sequence for which

$$\frac{x_0^2}{x_1} + \frac{x_1^2}{x_2} + \ldots + \frac{x_{n-1}^2}{x_n} < 4 \text{ for all } n.$$

4. Prove that if n is a positive integer such that the equation

$$x^3 - 3xy^2 + y^3 = n$$

has a solution in integers (x,y), then it has at least three such solutions.

Show that the equation has no solution in integers when $n = 2891$.

5. The diagonals AC and CE of the regular hexagon $ABCDEF$ are divided by the inner points M and N, respectively, so that

$$\frac{AM}{AC} = \frac{CN}{CE} = r.$$

Determine r if B, M, and N are collinear.

6. Let S be a square with sides of length 100 and let L be a path within S which does not meet itself and which is composed of linear segments A_0A_1, A_1A_2, ... ,$A_{n-1}A_n$ with $A_0 \neq A_n$.

Suppose that for every point P of the boundary of S there is a point of L at a distance from P not greater than 1/2. Prove that there are two points X and Y in L such that the distance between X and Y is not greater than 1 and the length of that part of L which lies between X and Y is not smaller than 198.

SOLUTIONS TO U.S.A. OLYMPIAD

We appreciate the work of Loren Larson, St. Olaf College, who prepared the solutions to the 1982 U.S.A. Olympiad for publication here. A pamphlet with the problems and solutions for the 1982 USA and International Olympiads, prepared for the MAA Committee on High School contests, is available for 50¢ from Dr. Walter E. Mientka, 917 Oldfather Hall, U. of Nebraska, Lincoln, NE, 68588-0322. Solutions to the 1982 Canadian Olympiad will appear in our January 1983 issue.

1. In a party with 1982 persons, among any group of four there is at least one person who knows each of the other three. What is the minimum number of people in the party who know everyone else?

Sol. First, assume that "knowing" is a reciprocal relation. Suppose that A, B, C are individuals none of whom know each of the other two. The conditions of the problem guarantee that each of the other 1979 individuals in the group know A, B, and C. We claim that these 1979 persons know everyone else also. For, suppose there is an individual D who does not know everyone. Since D knows A, B, and C, there is a person E among the 1977 others who D does not know. If, without loss of generality, A does not know B, then no one among A, B, D, E knows each of the other three. This contradiction shows that the minimum number is 1979.

Now, assume that "knowing" is not a symmetric relation. Then it is possible that no one knows everyone else, so the minimum number is zero. To see that this is possible, consider the following example. Arrange the people in a circle and consider the case in which each person knows everybody else except the person to the right. It is easy to check that the conditions of the problem are met.

2. Let $S_r = x^r + y^r + z^r$ with x, y and z real. It is known that if $S_1 = 0$,

$$\frac{S_{m+n}}{m+n} = \frac{S_m}{m} \cdot \frac{S_n}{n} \qquad (*)$$

for $(m,n) = (2,3),(3,2),(2,5)$ or $(5,2)$. Determine *all* other pairs of integers (m,n), if any, satisfying (*).

Sol. We shall show that (*) does not hold for any other pair of integers (m,n). In what follows we will assume that $m \geq 2$ and $n \geq 2$ since these are easily seen as necessary conditions.

Condition (*) cannot hold if both m and n are odd. To see this, consider the special case $x = 0$, $y = 1$, $z = -1$.

Condition (*) cannot hold if both m and n are even. To see this, take $x = 0$, $y = 1$, $z = -1$ to get

$$\frac{S_m}{m} \cdot \frac{S_n}{n} = \frac{4}{mn} \text{ and}$$

$$\frac{S_{m+n}}{m+n} = \frac{2}{m+n} .$$

These are equal only if $(m-2)(n-2) = 4$, which, under the conditions, happens only if $m = n = 4$. But $S_8/8 \neq (S_4/2)^2$ as can be checked by taking $x = -1$, $y = -1$, $z = 2$.

Condition (*) cannot hold if n is even and greater than 2 and m is odd. For, take $x = -1$, $y = -1$, $z = 2$. It is easy to check that

$$\frac{S_n}{n} \cdot \frac{S_3}{3} = \frac{6S_n}{3n} < \frac{6S_n}{n+3} < \frac{S_{n+3}}{n+3} ,$$

and, for $m > 3$,

$$\frac{S_n}{n} \cdot \frac{S_m}{m} < \frac{S_n S_m}{2(m+n)} < \frac{S_{n+m}}{n+m} .$$

Condition (*) cannot hold if $n = 2$ and m is odd and greater than 5. For, take $x = -1$, $y = -1$, $z = 2$. It is easy to check that

$$\frac{S_2}{2} \cdot \frac{S_m}{m} = 3\frac{S_m}{m} < \frac{S_{m+n}}{m+2} .$$

By symmetry, the only remaining cases are (2,3), (2,5), and (3,2), (5,2), and the proof is complete.

3. If point A_1 is in the interior of an equilateral triangle ABC and point A_2 is in the interior of triangle A_1BC, prove that

$$I.Q.(A_1BC) > I.Q.(A_2BC)$$

where the *isoperimetric quotient* of a figure F is defined by

$$I.Q.(F) = \frac{\text{Area } F}{(\text{Perimeter } F)^2} .$$

Sol. Lemma. If triangles ABC and $A'B'C'$ are such that $\angle ABC = \angle A'B'C'$ and if

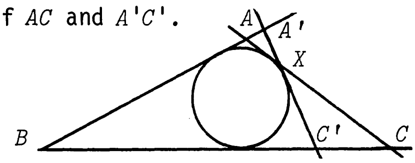
$$\frac{AB}{BC} < \frac{A'B'}{B'C'} < 1 ,$$

then $I.Q.(ABC) < I.Q.(A'B'C')$.

Proof. Similar figures have equal $I.Q.$ values, so we may assume henceforth that the incircles of each of the triangles are of radius one.

If F is a triangle, Area $F = \frac{r}{2}$ Per F, where Per F is the perimeter of F and r is the inradius of F. It follows that $I.Q.(F) = \frac{r^2}{4 \text{ Area } F}$, so to prove the Lemma we must show that Area $A'B'C'$ < Area ABC.

Since $\frac{AB}{BC} < \frac{A'B'}{B'C'} < 1$, the incircles are congruent, and $\angle A'B'C' = \angle ABC$, we may assume that B and B' coincide, B, A, A' are collinear with $AB < A'B'$ and B, C, C' are collinear with $B'C'$ < BC. Let X denote the intersection of AC and $A'C'$.



Then Area AXA' < Area CXC', and it follows that Area $A'B'C'$ < Area ABC. This completes the proof of the Lemma.

Now for the problem, extend BA_2 to intersect A_1C in A_3. By the Lemma, $I.Q.(A_2BC) < I.Q.(A_3BC) < I.Q.(A_1BC)$.

4. Prove that there exists a positive integer k such that $k2^n+1$ is composite for every positive integer n.

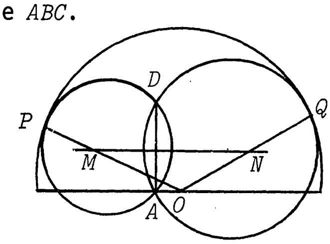
Sol. For each n, there exist unique integers s and t, $0 \leq t < 24$, such that $n = 24s + t$. Let $N = k2^n + 1$. It is straightforward to check the following.

i) $t \equiv 0 \bmod 2$ implies $N \equiv k+1 \bmod 3$;
ii) $t \equiv 1 \bmod 4$ implies $N \equiv 2k+1 \bmod 5$;
iii) $t \equiv 0 \bmod 3$ implies $N \equiv k+1 \bmod 7$;
iv) $t \equiv 7 \bmod 12$ implies $N \equiv 11k+1 \bmod 13$;
v) $t \equiv 11 \bmod 24$ implies $N \equiv 8k+1 \bmod 17$;
vi) $t \equiv 23 \bmod 24$ implies $N \equiv 121k+1 \bmod 241$.

Since one of *i*) - *vi*) must hold for a given n, it suffices to find a k such that each of the following congruences hold simultaneously: $k \equiv 2 \bmod 3$, $k \equiv 2 \bmod 5$, $k \equiv 6 \bmod 7$, $k \equiv 7 \bmod 13$, $k \equiv 2 \bmod 17$, $k \equiv -2 \bmod 241$. Such a k exists by the Chinese Remainder Theorem, or can be found explicitly (e.g., one such k is 1,624,097). (Note: 3, 5, 7, 13, 17, 241 are the prime factors of $2^{24} - 1$, so that modulo these primes, 2^{24} is congruent to 1.)

5. Given that A, B and C are three interior points of a sphere S such that AB and AC are perpendicular to the diameter of S through A. Through A, B and C, two spheres can be constructed which are both tangent to S. Prove that the sum of their radii is equal to the radius of S.

Sol. Let O be the center of S, let S_1 and S_2 denote the two spheres through A, B, and C which are tangent to S, let their centers be M and N respectively, and their respective points of tangency be P and Q; let D be such that AD is a diameter of circle ABC.



From the given information we know that MN passes through the center of circle ABC and is parallel to the given diameter of S through A. Also, P, M, O are collinear, as are Q, N, and O.

Let PQ intersect S_1 in E and S_2 in F; let G denote the intersection of ME and NF. Since $\angle PEM = \angle EPM = \angle EPO = \angle PQO = \angle PQN = \angle QFN$, it must be the case that $OMGN$ is a parallelogram. From this fact it follows that $\Delta MNG = \Delta MNO$, and therefore the distance from G to MN is equal to the distance from O to MN (which in turn is equal to the distance of D from MN). The only way this can happen is for $G = D$; that is, P, Q, D are collinear and $OMDN$ is a parallelogram. Thus, $PO = PM + MO = PM + ND$, and the proof is complete.

ACKNOWLEDGMENTS

In addition to our associate editors, the following have assisted the Magazine *by refereeing papers during the past year. We appreciate their special efforts.*

Abeles, Francine, *Kean College of N.J.*
Alexanderson, Gerald, *University of Santa Clara.*
Anderson, Sabra, *University of Minnesota, Duluth.*
Anton, Howard, *Drexel University.*
Archibald, Thomas, *University of Toronto.*
Banchoff, Thomas F., *Brown University.*
Bankoff, Leon, *Los Angeles, CA.*
Baxley, John V., *Wake Forest University.*
Beaver, Olga R., *Williams College.*
Berggren, John L., *Simon Fraser Univ.*
Bieberich, Richard A., *Ball State Univ.*
Boas, Mary L., *DePaul University.*
Boas, Ralph P., *Northwestern University.*
Brenneman, Frank, *Tabor College.*
Buck, R. Creighton, *University of Wisconsin, Madison.*
Burckel, Robert B., *Kansas State Univ.*
Burr, Stefan A., *City College (CUNY).*
Chakerian, Don, *University of California, Davis.*
Chung, Fan R. K., *Bell Laboratories.*
Cochran, Jane, *West Chester, PA.*
Coleman, Roger, *Ursinus College.*
Connelly, Robert, *Cornell University.*
Coxeter, H. S. M., *University of Toronto.*
Crowe, Michael J., *University of Notre Dame.*
Cummings, Larry J., *Univ. of Waterloo.*
Cunningham, R. S., *CSC - Stanislaus.*
Davis, Donald, *Lehigh University.*
Davis, Morton, *Teaneck, NJ.*
Davison, T. M. K., *McMaster University.*
Dershem, Herbert L., *Hope College.*
Diamond, George, *Moravian College.*
Dixon, Crist, *York College (CUNY).*
Eisenberg, Bennett, *Lehigh University.*
Emerson, John D., *Middlebury College.*
Firey, William J., *Oregon State Univ.*
Fisher, J. Chris, *Univ. of Regina.*
Flesner, David E., *Gettysburg College.*
Foster, Lorraine L., *California State University, Northridge.*
Fraser, Craig, *Yale University.*
Galovich, Steve, *Carleton College.*
Garner, Cyril, *Carleton University.*
Geist, Robert, *Duke University.*
Getchell, Charles L., *Lycoming College.*
Goldberg, Michael, *Washington, DC.*
Goldfeather, Jack, *Carleton College.*

Goldstein, Jerome A., *Tulane University.*
Graham, Ronald L., *Bell Laboratories.*
Grosshans, Frank D., *West Chester State College.*
Grosswald, Emil, *Temple University.*
Growney, JoAnne, *Bloomsburg State Col.*
Grünbaum, Branko, *Univ. of Washington.*
Guy, Richard, *University of Calgary.*
Hailperin, Ruth, *Moravian College.*
Hailperin, Theodore, *Lehigh Univ.*
Hargis, Charles, *Moravian College.*
Heuer, Gerald A., *Concordia College.*
Hilton, Peter, *Case Western University.*
Hutchinson, Joan, *Smith College.*
Infante, Ronald P., *Seton Hall Univ.*
Jones, Charles V., *Univ. of Toronto.*
Kasube, Herbert E., *Bradley University.*
Katz, Victor J., *University of DC.*
Klamkin, Murray S., *Univ. of Alberta.*
Klarner, David, *SUNY at Binghamton.*
Kline, Morris, *NYU, Courant Inst.*
Kolman, Bernard, *Drexel University.*
Konhauser, Joseph D. E., *Macalester College.*
Kullman, David E., *Miami University.*
Laatsch, Richard, *Miami University.*
Landau, Martin D., *Lafayette College.*
Larson, Loren C., *St. Olaf College.*
Leinbach, Carl, *Gettysburg College.*
Lenz, Gerald E., *St. John's University.*
Lewis, Jesse C., *Jackson State Univ.*
L'Heureux, James E., *West Chester State College.*
Liu, C. L., *Univ. of Illinois, Urbana.*
Loengard, Janet, *Moravian College.*
Logothetti, David, *Univ. of Santa Clara.*
Loud, Warren S., *Univ. of Minnesota.*
Mahoney, Michael S., *Princeton Univ.*
Martens, Susan, *York College (SUNY).*
Maurer, Stephen B., *Swarthmore College.*
McAllister, Marialuisa, *Moravian Col.*
Monson, Barry, *Univ. of New Brunswick.*
Moore, L. C., Jr., *Duke University.*
Mugler, Dale H., *Univ. of Santa Clara.*
Murphy, Michael G., *University of Houston, Downtown College.*
Niven, Ivan, *University of Oregon.*
Notz, William I., *Purdue University.*
Olinick, Michael, *Middlebury College.*
Page, Warren, *New York City Technical College.*
Pedoe, Daniel, *Univ. of Minnesota.*
Porter, Gerald J., *U. of Pennsylvania.*
Queen, Clifford, *Lehigh University.*
Raimi, Ralph, *University of Rochester.*
Rhoades, B. E., *Indiana University, Bloomington.*
Rigby, John F., *University College, Cardiff.*
Robertson, Timothy J., *Univ. of Iowa.*
Robinson, Raphael M., *University of California, Berkeley.*
Rosen, David, *Swarthmore College.*
Rusin, David, *University of Chicago.*
Ryden, Roy, *Humboldt State University.*
Sagan, Hans, *North Carolina State University, Raleigh.*
Sawka, John, , *Univ. of Santa Clara.*
Schattschneider, David, *Moravian Col.*
Seebach, J. Arthur, *St. Olaf College.*
Senechal, Marjorie, *Smith College.*
Shephard, Geoffrey C., *University of East Anglia, Norwich.*
Sieber, James L., *Shippensburg State College.*
Siegel, Martha J., *Towson State Univ.*
Small, Lynne B., *University of San Diego, Alcala Park.*
Smogor, Louis E., *DePauw University.*
Somers, Kay, *Moravian College.*
Spencer, Joel, *SUNY at Stony Brook.*
Steen, Lynn A., *St. Olaf College.*
Stephens, Stanley L., *Anderson College.*
Stolarsky, Kenneth B., *University of Illinois, Urbana.*
Stone, Alexander P., *University of New Mexico.*
Stonesifer, Randolph J., *Lafayette Col.*
Straffin, Philip D., Jr., *Beloit Col.*
Tanis, Elliot, *Hope College.*
Thompson, Maynard, *Indiana University, Bloomington.*
Thompson, Robert C., *University of California, Santa Barbara.*
Tingley, Daryl, *Michigan State Univ.*
Trivedi, Kishor, *Duke University.*
Tucker, Alan, *SUNY at Stony Brook.*
Ward, James E., III, *Bowdoin College.*
Warner, Hoyt D., *Hartford Grad. Center.*
White, Arthur T., *Western Michigan University.*
Whitt, Lee, *Daniel H. Wagner Associates.*
Wilansky, Albert, *Lehigh University.*
Wilf, Herbert S., *Univ. of Pennsylvania.*
Wright, Tim, *University of Missouri, Rolla-Rolla.*

SOME INFORMATIVE AND USEFUL BOOKS FROM THE MAA...

Browse through this list and see if your library is missing some of these important books published by the Association.

APPLICATIONS OF UNDERGRADUATE MATHEMATICS IN ENGINEERING—written and edited by Ben Noble. A collection of articles prepared by engineers for the Committee on the Undergraduate Program in Mathematics. 364 pp. Hardbound.
List: $19.00. MAA Member: $13.00.

ANNOTATED BIBLIOGRAPHY OF EXPOSITORY WRITING IN THE MATHEMATICAL SCIENCES, prepared by M. P. Gaffney and L. A. Steen. An invaluable reference source of expository articles in mathematics. 282 pp. Paperbound.
List: $12.00. MAA Member: $8.00.

AN ANNOTATED BIBLIOGRAPHY OF FILMS AND VIDEOTAPES FOR COLLEGE MATHEMATICS, by David I. Schneider. An up-to-date listing of films and videotapes available for classroom use. 107 pp. Paperbound.
List: $9.00. MAA Members: $6.00.

A BASIC LIBRARY LIST FOR TWO-YEAR COLLEGES, prepared by the Committee on Basic Library Lists. A recommended library nucleus for two-year colleges. 66. pp. Paperbound.
List: $8.00. MAA Member: $6.00.

A BASIC LIBRARY LIST FOR FOUR-YEAR COLLEGES, prepared by CUPM. Presents listings of books and journals that should be in every college library. 106 pp. Paperbound.
List: $9.00. MAA Member: $6.50

THE CHAUVENET PAPERS. *A Collection of Prize Winning Expository Papers in Mathematics,* edited by James C. Abbott. Two-volumes of the collected prize winning Chauvenet Papers. Vol. 1—312 pp. Hardbound. Vol. 2—282 pp. Hardbound.
List: $21.00 each. MAA Member $16.00 each.
[Two volume sets
List: $36.00. MAA Member: $27.00.]

CRITICAL VARIABLES IN MATHEMATICS EDUCATION: *Findings from a Survey of the Empirical Literature,* by E. G. Begle. A joint publication of the MAA and the National Council of Teachers of Mathematics. List: $8.00. MAA Member: $6.40.

A COMPENDIUM OF CUPM RECOMMENDATIONS. Volumes I and II. A collection of the major recommendations of the Committee on the Undergraduate Program in Mathematics. Two volumes. 756 pp. Hardbound.
List: $16.50. MAA Member: $12.00.

THE WILLIAM LOWELL, PUTNAM MATHEMATICAL COMPETITION: PROBLEMS AND SOLUTIONS—1938-1964. Compiled the R. E. Greenwood, A. M. Gleason, and L. M. Kelly. Contains problems and solutions to the first 25 Putnam Exams. 652 pp. Hardbound.
List: $35.00. MAA Member: $26.00.

THE MATHEMATICAL ASSOCIATION OF AMERICA: *Its First Fifty Years.* An historical perspective of the Association. 170 pp. Hardbound.
List: $10.00. MAA Member: $5.00.

FIFTY YEAR INDEX OF THE MATHEMATICS MAGAZINE, edited by Lynn A. Steen, and J. Arthur Seebach. Cumulative index of volumes 1-50. 163 pp. paperbound.
List: $10.00. MAA Member: $6.50.

INDEX OF THE AMERICAN MATHEMATICAL MONTHLY. Contains the tables of contents for each issue of volumes 1-80 as well as subject and author indices. 269 pp. Hardbound.
List: $19.00. MAA Member: $13.00.

PRIME-80. *Proceedings of a Conference on Prospects in Mathematics Education in the 1980's.* Included are the background of the conference and the recommendations that resulted.
84 pp. Paperbound. $3.50.

PROFESSIONAL OPPORTUNITIES IN THE MATHEMATICAL SCIENCES, Tenth Edition. 1978. Designed for the student interested in a career in mathematics. 35 pp. Paperbound. $1.50 each. 95¢ for orders of five or more.

RECOMMENDATIONS ON A GENERAL MATHEMATICAL SCIENCES PROGRAM. Prepared by the Committee on the Undergraduate Program in Mathematics. (CUPM) 102 pp. Paperbound. $3.50 each.

Order From:

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

Eminent Mathematicians and Mathematical Expositors speak to

STUDENTS and TEACHERS in . . .

The NEW MATHEMATICAL LIBRARY

An internationally acclaimed paperback series providing:

- stimulating excursions for students beyond traditional school mathematics.
- supplementary reading for school and college classrooms.
- valuable background reading for teachers.
- challenging problems for solvers of all ages from high school competitions in the US and abroad.

The New Mathematical Library is published by the MATHEMATICAL ASSOCIATION OF AMERICA. The volumes are paperbound.

For information regarding the price of these publications, please contact The Mathematical Association of America at the address listed below.

NUMBERS: RATIONAL AND IRRATIONAL by Ivan Niven, NML-01

WHAT IS CALCULUS ABOUT? By W. W. Sawyer, NML-02

AN INTRODUCTION TO INEQUALITIES, by E.F. Beckenbach, and R. Bellman, NML-03

GEOMETRIC INEQUALITIES, by N.D. Kazarinoff, NML-04

THE CONTEST PROBLEM BOOK. Problems from the Annual High School Mathematics Contests sponsored by the MAA, NCTM, Mu Alpha Theta, The Society of Actuaries, and the Casualty Actuarial Society. Covers the period 1950-1960. Compiled and with solutions by C.T. Salkind. NML-05

THE LORE OF LARGE NUMBERS, by P.J. Davis, NML-06

USES OF INFINITY, by Leo Zippin, NML-07

GEOMETRIC TRANSFORMATIONS, by I.M. Yaglom, translated by Allen Shields, NML-08

CONTINUED FRACTIONS, by C.D. Olds, NML-09

GRAPHS AND THEIR USES, by Oystein Ore, NML-10

HUNGARIAN PROBLEM BOOKS I and II, based on the Eötvos Competitions 1894-1905 and 1906-1928. Translated by E. Rapaport, NML-11 and NML-12

EPISODES FROM THE EARLY HISTORY OF MATHEMATICS, by A. Aaboe, NML-13

GROUPS AND THEIR GRAPHS, by I. Grossman and W. Magnus, NML-14

THE MATHEMATICS OF CHOICE, by Ivan Niven, NML-15

FROM PYTHAGORAS TO EINSTEIN, by K.O. Friedrichs, NML-16

THE CONTEST PROBLEM BOOK II. A continuation of NML-05 containing problems and solutions from the Annual High School Mathematics Contests for the period 1961-1965. NML-17

FIRST CONCEPTS OF TOPOLOGY, by W.G. Chinn and N.E. Steenrod, NML-18

GEOMETRY REVISITED, by H.S.M. Coxeter, and S.L. Greitzer, NML-19

INVITATION TO NUMBER THEORY, by Oystein Ore, NML-20

GEOMETRIC TRANSFORMATIONS II, by I.M. Yaglom, translated by Allen Shields, NML-21

ELEMENTARY CRYPTANALYSIS — A Mathematical Approach, by Abraham Sinkov, NML-22

INGENUITY BY MATHEMATICS, by Ross Honsberger, NML-23

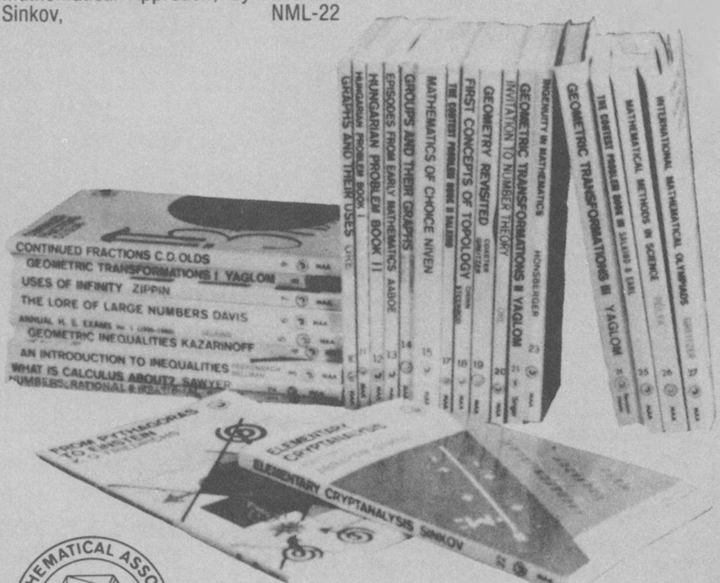
GEOMETRIC TRANSFORMATIONS III, by I.M. Yaglom, translated by Abe Shenitzer, NML-24

THE CONTEST PROBLEM BOOK III. A continuation of NML-05 and NML-17, containing problems and solutions from the Annual High School Mathematics Contests for the period 1966-1972. NML-25

MATHEMATICAL METHODS IN SCIENCE, by George Pólya, NML-26

INTERNATIONAL MATHEMATICAL OLYMPIADS, 1959-1977. Problems, with solutions, from the first nineteen International Mathematical Olympiads. Compiled and with solutions by S.L. Greitzer. NML-27

THE MATHEMATICS OF GAMES AND GAMBLING, by Edward W. Packel, NML-28



Send Orders to: **The Mathematical Association of America**
1529 Eighteenth St., N.W., Washington, D.C. 20036

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, N.W.
Washington, DC 20036
MATHEMATICS MAGAZINE VOL. 55, NO. 5, NOVEMBER 1982