MATHEMATICS MAGAZINE

CONTENTS

VOLUME 45 • SEPTEMBER 1972 • NUMBER 4

MATHEMATICS MAGAZINE

GERHARD N. WOLLAN, *EDITOR*

EDITORIAL CORRESPONDENCE should be sent to the EDITOR, G. N. WOLLAN, Department of Mathematics, Purdue University, Lafayette, Indiana 47907. Articles should be typewritten and triple-spaced on $8\frac{1}{2}$ by 11 paper. The greatest possible care should be taken in preparing the manuscript, and authors should keep a complete copy. Figures should be drawn on separate sheets in India ink and of a suitable size for photographing.

NOTICE OF CHANGE OF ADDRESS and other subscription correspondence should be sent to the Executive Director, A. B. WILLCOX, Mathematical Association of America, Suite 310, 1225 Connecticut Avenue, N. W., Washington, D.C. 20036.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214.

NEW SUBSCRIPTION PRICES. For several years MATHEMATICS MAGAZINE has operated at a substantial deficit. In these days of tight budgets the Association can no longer cover such deficits from general funds, and the MAGAZINE must become essentially self-supporting. For this reason, the U.S. Price Commission has approved an increase in the subscription rates, effective with the September–October 1972 issue. The new rates are: One year subscription (5 issues), \$7.00; one year subscription for members of the MAA and Mu Alpha Theta, \$5.00; single issues, \$1.40.

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

GREATEST COMMON DIVISORS IN ALGEBRAIC NUMBER FIELDS

H. M. EDGAR, San Jose State College

1. Introduction. In a first course in algebraic number theory one encounters a concept of greatest common divisor of two integers in the ring of integers of a given algebraic number field. In an effort to clarify this concept we give three definitions of what one might mean by greatest common divisor. We then compare and contrast these three definitions by means of some examples and theorems. Finally we compare the definitions with the usual definition of the greatest common divisor of two principal (integral) ideals in the ring of integers of a given algebraic number field and with the situation in the ring of all algebraic integers.

2. Notation and definitions. Let K be an arbitrary algebraic number field. Let $[K]$ denote the ring of integers of K and let α and β be arbitrary elements of $[K]$. Let A denote the field of algebraic numbers and let $[A]$ denote the ring of algebraic integers. An element $\alpha \in [K]$ is said to be indecomposable if in every factorization $\alpha = \beta\gamma$ over $[K]$ precisely one factor is a unit. An indecomposable element in a unique factorization domain is called a prime.

In Definition 1 the notation $\alpha | \beta$ will mean

$$\frac{\beta}{\alpha} \in [K]$$

while in Definition 3 the notation $\alpha | \beta$ will mean

$$\frac{\beta}{\alpha} \in [A].$$

The notation $N\alpha | N\beta$ will always mean

$$\frac{N\beta}{N\alpha} \in Z,$$

Z being the ring of rational integers.

DEFINITION 1. *The elements α and β of $[K]$ possess a good* GCD $\gamma \in [K]$ *provided the following conditions hold:*

(1) $\gamma | \alpha$ *and* $\gamma | \beta$,

(2) *if* $\delta \in [K]$ *satisfies* $\delta | \alpha$ *and* $\delta | \beta$ *then* $\delta | \gamma$,

(3) *there exist* $\lambda, \mu \in [K]$ *such that* $\lambda\alpha + \mu\beta = \gamma$.

DEFINITION 2. *The elements α and β of $[K]$ possess a bad* GCD $\gamma \in [K]$ *provided the following conditions hold:*

(4) *α and β fail to possess a good* GCD.

(5) *γ satisfies conditions* (1) *and* (2) *and fails to satisfy condition* (3).

DEFINITION 3. *The elements* α *and* β *of* $[K]$ *possess an ideal* GCD $\gamma \in [A]-[K]$ *provided the following conditions hold*:

(1)′ $\gamma \mid \alpha$ *and* $\gamma \mid \beta$,

(2)′ *if* $\delta \in [A]$ *satisfies* $\delta \mid \alpha$ *and* $\delta \mid \beta$ *then* $\delta \mid \gamma$,

(3)′ *there exist* $\lambda, \mu \in [A]$ *such that* $\lambda\alpha + \mu\beta = \gamma$.

3. Interrelations of definitions. We note that each type of GCD is determined only up to unit factors.

Definition 1 and Definition 2 have been arranged to be mutually exclusive. Example 1 shows that these two definitions are not complementary.

Example 1. Consider $\alpha = 21$ and $\beta = -9 + 3\sqrt{-5}$ in $[Q(\sqrt{-5})]$. Now $\alpha = 3 \cdot 7 = (1 + 2\sqrt{-5}) \cdot (1 - 2\sqrt{-5})$ and $\beta = (1 + 2\sqrt{-5}) \cdot (1 + \sqrt{-5}) = 3(-3 + \sqrt{-5})$ so that 3 and $1 + 2\sqrt{-5}$ are nonunit common factors of α and β. If γ is to be a bad CGD of α and β then we must have $3 \mid \gamma$ and $1 + 2\sqrt{-5} \mid \gamma$. We also require $\gamma \mid \beta$. Taking norms we find that we must have $9 \mid N_\gamma, 21 \mid N\gamma$ and $N\gamma \mid N\beta = 126$. Hence $N\gamma = 63$ or 126 is required. Since $63 = a^2 + 5b^2$ has no rational integral solutions $N_\gamma = 63$ is impossible. It can easily be shown that $N_\gamma = 126$ is impossible. Here, then, is a situation in which α and β fail to possess a bad GCD and also fail to have a good GCD.

In order to compare and contrast Definition 1 and Definition 3 we note the following two results:

THEOREM 1. *The elements* α *and* β *of* $[K]$ *possess a good* GCD γ *if and only if the ideal generated by* α *and* β *is the principal ideal* (γ).

THEOREM 2. *The elements* α *and* β *of* $[K]$ *possess an ideal* GCD *if and only if the ideal generated by* α *and* β *fails to be a principal ideal.*

Proof of Theorem 1. Clear.

Proof of Theorem 2. By Theorem 1 we see that the ideal generated by α and β is nonprincipal if and only if α and β fail to have a good GCD. By Theorem 10.5, p. 113 of [**1**] we see that the ideal generated by α and β is nonprincipal if and only if α and β have an ideal GCD $\gamma \in [A] - [K]$.

Hence Definitions 1 and 3 are complementary.

The next example shows that Definitions 2 and 3 are not mutually exclusive.

Example 2. Consider $\alpha = 3$ and $\beta = 1 + 2\sqrt{-5}$ in $[Q(\sqrt{-5})]$. It can be shown that α and β have $\gamma = 1$ as a bad GCD and have $\sqrt{2 + \sqrt{-5}}$ as an ideal GCD (see, for instance, [**1**], p. 76).

4. Further remarks. As an immediate consequence of Theorem 1 we obtain

THEOREM 3. *The ring* $[K]$ *is a unique factorization domain if and only if every pair of elements* $\alpha, \beta \in [K]$ *possesses a good* GCD.

Proof of Theorem 3. For any algebraic number field K it is well known that $[K]$ is a unique factorization domain if and only if it is a principal ideal domain.

Let α, $\beta \in [K]$ and let (α) and (β) be the corresponding principal ideals. Let $((\alpha),(\beta)) = (\alpha,\beta)$ be the greatest common divisor of the two principal ideals (α) and (β), defined in the usual manner. Then $((\alpha),(\beta))$ is principal if and only if α and β possess a good GCD. In particular (α) and (β) are said to be relatively prime if $(\alpha,\beta) = (1)$.

Example 3. Consider $\alpha = 3$ and $\beta = 4 + \sqrt{-5}$ in $[Q(\sqrt{-5})]$. We show that α and β have $\gamma = 1$ as a bad GCD. Condition (1) is obviously satisfied. Let $\delta \in [Q(\sqrt{-5})]$ have the properties $\delta | 3$ and $\delta | 4 + \sqrt{-5}$. Then $N\delta | 9$ and $N\delta | 21$ so that $N\delta | 3$. If $\delta = a + b\sqrt{-5}$ with $a, b \in Z$ then $N\delta = a^2 + 5b^2$ and so $N\delta = 3$ is impossible. Hence $N\delta = 1$ and so $\delta = \pm 1$ and thus $\delta | 1$. Hence condition (2) holds. Condition (3) can never hold. Now ± 1 are the only elements of $[Q(\sqrt{-5})]$ for which conditions (1) and (2) hold, but upon equating real and imaginary parts in $3(a + b\sqrt{-5}) + (4 + \sqrt{-5})(c + d\sqrt{-5}) = \pm 1$ we obtain $3a + 4c - 5d = \pm 1$ and $3b + c + 4d = 0$ which leads to the impossible equation $3(a - b + c - 3d) = \pm 1$. The numbers α and β of this example are nonassociated indecomposable elements and are sometimes said to be relatively prime (see, for instance, [**2**], p. 57). On the other hand, the corresponding ideals (3) and $(4 + \sqrt{-5})$ cannot be considered as relatively prime ideals.

If the definition of good GCD is recast for $[A]$ instead of $[K]$ by replacing $[K]$ by $[A]$ throughout then it can be shown (for instance, [**2**], p. 85) that any two elements α and $\beta \in [A]$ have a good GCD. On the other hand $[A]$ is not a unique factorization domain and in fact there are no indecomposable (let alone prime!) elements in $[A]$.

References

1. H. Pollard, The Theory of Algebraic Numbers, MAA, Carus Monograph, No. 9, 1960.
2. W. J. LeVeque, Topics in Number Theory, vol. 2, Addison-Wesley, Reading, 1956.

FUNCTIONAL EQUATIONS IN MATHEMATICAL STATISTICS

PETER TAN, Carleton University

1. Introduction. A functional equation is an equation satisfied by an unknown function or functions under certain prescribed conditions or with some predetermined properties. If a functional equation has a unique solution, then the solution can be said to be characterized by the equation under the prescribed conditions.

Functional equations often occur in characterization problems in mathematical statistics and probability theory. In this paper we shall present some of the well-known functional equations that arise in simple statistical problems. In Section 3 a newer method of solution of a functional equation by the elementary theory of linear spaces of functions is also given.

First of all let us introduce the concept of a probability function.

Let α, $\beta \in [K]$ and let (α) and (β) be the corresponding principal ideals. Let $((\alpha),(\beta)) = (\alpha,\beta)$ be the greatest common divisor of the two principal ideals (α) and (β), defined in the usual manner. Then $((\alpha),(\beta))$ is principal if and only if α and β possess a good GCD. In particular (α) and (β) are said to be relatively prime if $(\alpha,\beta) = (1)$.

Example 3. Consider $\alpha = 3$ and $\beta = 4 + \sqrt{-5}$ in $[Q(\sqrt{-5})]$. We show that α and β have $\gamma = 1$ as a bad GCD. Condition (1) is obviously satisfied. Let $\delta \in [Q(\sqrt{-5})]$ have the properties $\delta | 3$ and $\delta | 4 + \sqrt{-5}$. Then $N\delta | 9$ and $N\delta | 21$ so that $N\delta | 3$. If $\delta = a + b\sqrt{-5}$ with $a, b \in Z$ then $N\delta = a^2 + 5b^2$ and so $N\delta = 3$ is impossible. Hence $N\delta = 1$ and so $\delta = \pm 1$ and thus $\delta | 1$. Hence condition (2) holds. Condition (3) can never hold. Now ± 1 are the only elements of $[Q(\sqrt{-5})]$ for which conditions (1) and (2) hold, but upon equating real and imaginary parts in $3(a + b\sqrt{-5}) + (4 + \sqrt{-5})(c + d\sqrt{-5}) = \pm 1$ we obtain $3a + 4c - 5d = \pm 1$ and $3b + c + 4d = 0$ which leads to the impossible equation $3(a - b + c - 3d) = \pm 1$. The numbers α and β of this example are nonassociated indecomposable elements and are sometimes said to be relatively prime (see, for instance, [**2**], p. 57). On the other hand, the corresponding ideals (3) and $(4 + \sqrt{-5})$ cannot be considered as relatively prime ideals.

If the definition of good GCD is recast for $[A]$ instead of $[K]$ by replacing $[K]$ by $[A]$ throughout then it can be shown (for instance, [**2**], p. 85) that any two elements α and $\beta \in [A]$ have a good GCD. On the other hand $[A]$ is not a unique factorization domain and in fact there are no indecomposable (let alone prime!) elements in $[A]$.

References

1. H. Pollard, The Theory of Algebraic Numbers, MAA, Carus Monograph, No. 9, 1960.
2. W. J. LeVeque, Topics in Number Theory, vol. 2, Addison-Wesley, Reading, 1956.

FUNCTIONAL EQUATIONS IN MATHEMATICAL STATISTICS

PETER TAN, Carleton University

1. Introduction. A functional equation is an equation satisfied by an unknown function or functions under certain prescribed conditions or with some predetermined properties. If a functional equation has a unique solution, then the solution can be said to be characterized by the equation under the prescribed conditions.

Functional equations often occur in characterization problems in mathematical statistics and probability theory. In this paper we shall present some of the well-known functional equations that arise in simple statistical problems. In Section 3 a newer method of solution of a functional equation by the elementary theory of linear spaces of functions is also given.

First of all let us introduce the concept of a probability function.

A random phenomenon is described by the probability distribution of a random variable X, which is in turn often specified by a probability function $f(x)$ defined for all real numbers. X is a continuous random variable if its probability function $f(x)$ possesses the following properties:

$$(1) \qquad f(x) \geqq 0 \quad \text{for} \quad -\infty < x < \infty, \int_{-\infty}^{\infty} f(x)dx = 1$$

where the integral is in the sense of Lebesgue, and the probability that X has values that lie in a Borel set E in R^1 is given by

$$P[E] = \int_E f(x)dx.$$

A probability function $f(x)$ often involves some constants called parameters of the probability distribution of X.

2. Characterization of the normal and exponential distributions by functional equations. Let $f(x, \mu) = f(x-\mu)$ be a probability function of a continuous random variable with a parameter μ. We further assume continuous differentiability of the function f and the condition that

$$(2) \qquad g_n(\mu) = f_1(x_1-\mu)f(x_2-\mu)\cdots f(x_n-\mu), \quad -\infty < \mu < \infty,$$

attains its maximum at $\hat{\mu} = (x_1 + x_2 + \cdots + x_n)/n$.

If we disregard the points $(x_1, x_2, \cdots, x_n)$ at which $g_n(\mu) = 0$, then from the conditions that $g_n(\mu)$ has its maximum at $\mu = \hat{\mu}$ and $g_n(\mu)$ has a continuous derivative $g_n'(\mu)$, we obtain

$$(3) \quad g_n'(\hat{\mu}) = -\frac{f'(x_1-\hat{\mu})g_n(\hat{\mu})}{f(x_1-\hat{\mu})} - \frac{f'(x_2-\hat{\mu})g_n(\hat{\mu})}{f(x_2-\hat{\mu})} - \cdots - \frac{f'(x_n-\hat{\mu})g_n(\hat{\mu})}{f(x_n-\hat{\mu})} = 0.$$

Setting $u_i = x_i - \hat{\mu}$ and $g(u_i) = f'(u_i)/f(u_i)$, $i = 1, \cdots, n$, we get Gauss' functional equation [**1**, p. 47]

$$(4) \qquad g(u_1) + g(u_2) + \cdots + g(u_n) = 0, \; u_1 + u_2 + \cdots + u_n = 0,$$

where $g(u)$ is, under our assumptions, a continuous function at all real values u.

The functional equation (4) is easily reduced to Cauchy's (1821) basic functional equation

$$(5) \qquad g(x+y) = g(x) + g(y), \quad -\infty < x, y < \infty,$$

whose continuous solution is of the form

$$(6) \qquad g(u) = cu, \quad -\infty < u < \infty,$$

where c is a real constant. (An excellent elementary exposition on the general solution of (5) is given in [**9**]. An elementary proof of the nature of the discontinuous solutions of (5) is given in [**5**].)

Integration of the differential equation

$$f'(u)/f(u) = cu$$

together with the conditions (1) imposed on the function f gives the unique solution

$$f(x-\mu) = \frac{1}{(\sqrt{2\pi})\sigma} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right], \; -\infty < x < \infty, \tag{7}$$

where σ^2 is a positive constant.

A random variable X having a probability function of the form (7) is said to have a normal distribution with parameters μ and σ. The normal distribution may therefore be characterized by the functional equation (4) which originated from Gauss' (1807) investigation of the functional form $\hat{\mu}$ as a function of $(x_1, x_2, \cdots, x_n)$ that maximizes the function $g_n(\mu)$ in (2).

If X is restricted to be a nonnegative continuous random variable satisfying the functional equation

$$F(x+y) = F(x)F(y) \quad \text{for} \quad x > 0, \quad y > 0, \tag{8}$$

where

$$F(x) = \int_x^\infty f(t)dt > 0, \quad x > 0$$
$$= 1 \qquad\qquad , \quad x \leqq 0, \tag{9}$$

then with the substitution $g(x) = \ln F(x)$, (8) is again reduced to (5) defined on the positive part of the real line. It gives the solution $F(x) = e^{-cx}$ for $x > 0$ where c is a positive constant since F is a bounded function. The corresponding probability function is therefore given by

$$f(x) = ce^{-cx} \quad \text{for} \quad x > 0, \; c > 0,$$
$$= 0 \qquad\qquad \text{otherwise}. \tag{10}$$

A random variable with a probability function of the form (10) is said to have an exponential distribution characterized by the "no-memory" property (8).

3. Characterization of probability distributions by the functional equation

$$h(x+y) = \sum_{k=1}^{n} f_k(x)g_k(y). \tag{11}$$

Let $f(x,\mu) = f(x-\mu)$ be a positive probability function of a continuous random variable with a parameter μ. Suppose that $h(x-\mu)$, the natural logarithm of $f(x-\mu)$, has the form

$$h(x-\mu) = \sum_{k=1}^{n} f_k(x)g_k(\mu), \; -\infty < x, \; \mu < \infty, \tag{12}$$

where the function $h = \ln f$ is differentiable and all real functions h, f_k, g_k are unknown. Probability functions of the form (12) represent a family of distributions often called an "exponential family with a location parameter μ" in statistics. One can easily verify that the probability function (7) of the normal distribution is o this type with $n = 3$, while the probability function (10) of the exponential distribution may also be considered as belonging to this type with $n = 2$ if we replace x by $x - \mu > 0$, $\mu > 0$.

A solution of the general functional equation of the form (11) is given by Aczel in [**1**] for h as unknown under the supposition that h, f_k, g_k $(k = 1,2,\cdots,n)$ are n-times differentiable. Equation (11) is then reduced to a homogeneous linear differential equation of the nth order with constant coefficients. References are also given there as well as in [**2**] for solutions under weaker assumptions. In the following we shall give a solution of (12) by a method slightly different from that presented in [**1**].

For each value of μ the function h considered as a function of x is a linear combination of $f_1(x), f_2(x), \cdots, f_n(x)$. Suppose that the linear space generated by the n functions $f_k(x)$ has dimension n. The linear space generated by the functions $h(x-\mu)$ for all values of μ must have dimension $m \leqq n$. Let $\mu_1, \mu_2, \cdots, \mu_m$ be m distinct values of the parameter μ such that

$$h(x - \mu_1),\ h(x - \mu_2), \cdots, h(x - \mu_m)$$

is a basis of this linear space. Then

$$h(x - \mu) = a_1(\mu)h(x - \mu_1) + \cdots + a_m(\mu)h(x - \mu_m).$$

Replacing μ by $\mu_i + t$ for some $i = 1,2,\cdots,m$, the last equation becomes

$$h(x - \mu_i - t) = \sum_{j=1}^{m} a_j(\mu_i + t)h(x - \mu_j). \tag{13}$$

Since $h(x - \mu_j)$, $j = 1,2,\cdots,m$, are linearly independent functions of x, we can find $x_1, x_2, \cdots, x_m$ such that the determinant $|h(x_k - \mu_j)|$ $(k, j = 1,2,\cdots,m)$ is not zero. (A proof of this last statement is given in [1, p. 201].) Each $a_j(\mu_i + t)$ can then be expressed as a linear combination of the functions $h(x_k - \mu_i - t)$, $(k = 1,2,\cdots,m)$. By the assumption of differentiability of h, each $a_j(\mu_i + t)$ as a function of t is therefore differentiable. Differentiating (13) with respect to t and then putting $t = 0$ we obtain

$$-h'(x - \mu_i) = a_1'(\mu_i)h(x - \mu_1) + \cdots + a_m'(\mu_i)h(x - \mu_m),\ i = 1,2,\cdots,m. \tag{14}$$

The solution of the system (14) of m linear differential equations is

$$\ln f(x - \mu) = h(x - \mu) = \sum_{j=1}^{k} e^{\mu_j(x-\mu)} \left(\sum_{i=1}^{m_j} c_{ij}(x - \mu)^{i-1} \right), \tag{15}$$

where $m_1 + m_2 + \cdots + m_k = m \leqq n$, c_{ij} are constants, the μ_j are k distinct roots of multiplicity m_j of a polynomial equation of degree m, and the m terms in the right-hand member of (15) are linearly independent functions.

4. Remarks. (i) Gauss' result in characterizing the normal distribution (7) by the function $\hat{\mu} = (x_1 + \cdots + x_n)/n$ which maximizes the "likelihood function" $g_n(\mu)$ in (2) remains valid if, instead of assuming f to be continuously differentiable everywhere, we assume merely that $f(x)$ is lower semicontinuous at $x = 0$, [**8**, Theorem 1]. The proof of this generalization of Gauss' result in [**8**] still relies on the solution of the functional equation (5) valid on a restricted domain of definition.

(ii) A generalization of Cauchy's exponential functional equation (8) to the form

$$F(s_1 + t, s_2 + t) = F(s_1,s_2)F(t,t),\ s_1 \geqq 0,\ s_2 \geqq 0,\ t \geqq 0$$

is given in [**6**].

(iii) The statistical significance of equation (12) and its solutions is discussed in [**3**]. In fact, Ferguson [**4**, Theorem 1] shows that without any assumptions on continuity and differentiability of the functions f_k, g_k and h, (15) remains the solution of (12).

(iv) Other examples of applications of functional equations in elementary probability and stochastic processes can be found in [**1**, 106–116] and [**7**, p. 120]. Several articles on more advanced functional equations encountered in mathematical statistics have appeared in recent (1970) issues of Aequationes Mathematicae.

Acknowledgement. The author wishes to thank the referee for helpful suggestions which greatly improved the presentation in this paper. This research was supported by N. R. C. grant No. A7586, and a preliminary version of this paper was prepared at the 1970 Summer Research Institute of the Canadian Mathematical Congress.

References

1. J. Aczel, Lectures on Functional Equations and Their Applications, Academic Press, New York, 1966.

2. ———, On Applications and Theory of Functional Equations, Birkhäuser Verlag, Basel, 1969.

3. E. B. Dynkin, Necessary and sufficient statistics for a family of probability distributions, Selected Transl. Math. Statist. and Prob., 1 (1951) 17–40.

4. T. S. Ferguson, Location and scale parameters in exponential families of distributions, Ann. Math. Statist., 33 (1962) 986–1001.

5. E. Hewitt and H. S. Zuckerman, Remarks on the functional equation $f(x+y)=f(x)+f(y)$, this MAGAZINE, 42 (1969) 121–123.

6. A. W. Marshall and I. Olkin, A multivariate exponential distribution, Amer. Statist. Assoc., 62 (1967) 30–44.

7. E. Parzen, Stochastic Processes, Holden-Day, San Francisco, 1962.

8. H. Teicher, Maximum likelihood characterization of distributions, Ann. Math. Statist., 32 (1961) 1214–1222.

9. A. Wilansky, "Additive functions" in Lectures on Calculus, ed. by K. O. May, Holden-Day, San Francisco, 1969.

MORE ON PALINDROMES BY REVERSAL-ADDITION

CHARLES W. TRIGG, San Diego, California

When the digits of the integer N are written in reverse order, the integer N' is obtained. Let $N + N' = S_1$, $S_1 + S_1' = S_2$, $S_2 + S_2' = S_3, \cdots, S_{k-1} + S_{k-1}' = S_k$. It has been conjectured that for every N there is a k for which S_k is a palindrome.

The validity of this conjecture for $N < 10^4$ has been questioned in [1]. There it is reported that 249 integers $< 10^4$ have reversal-addition sequences that are palindrome-free up to at least $k = 100$. All one-digit and two-digit integers lead to palindromes. No other $N < 10^4$ requires more operations to produce a palindrome than the 24 required to produce $S_{24}(89) = 8813200023188$.

The search for palindrome-free sequences is now extended to $N < 10^5$. To exhaust the field of five-digit integers, only 3420 basic integers need to be dealt with. A *basic integer* is the smallest member and representative of a family all of which have the same S_1. Thus $S_1(10321) = 22622$ indicates that 10321, 11311, 12301, 20320, 21310, and 22300 each produces this palindrome in one operation. These family members have the same middle digits and their symmetrically located digit pairs have the same sums.

Computation can be reduced further by recognizing that

A. Every five-digit N composed of digits < 5 will have a palindromic S_1, as will each N with a middle digit < 5 and digit pairs symmetrical to the middle that have sums < 10.

B. $S_2(a0cd9) = S_2(adc99)$, with $d < 9$ and $c < 5$, $d > 0$.

C. $S_3(9b0d9) = S_3(9b9d9)$ for all digit pairs (b,d) except $(0,0), (0,9)$, and $(9,9)$.

First palindromes with k $\geqq$ 24. In this investigation, if a palindrome appeared in a reversal-addition sequence in fewer than 24 operations, that sequence was not considered further.

There are 1515 five-digit integers representable by 61 basic integers, which require twenty-four or more operations to produce the first palindrome. Many of these sequences merge into common reversal addition sequences, so that only 21 distinct first palindromes are produced by members of the set. They are:

$S_{24}(16991)$, $S_{24}(60649) = S_{24}(64699)$, $S_{24}(50169) = S_{24}(56199)$, $S_{25}(10797) = S_{24}(80499)$, $S_{25}(70779) = S_{25}(77799)$, $S_{26}(90189) = S_{26}(98199)$, $S_{27}(10921) = S_{26}(10853) = S_{25}(12697) = S_{25}(17197) = S_{24}(80339) = S_{24}(83399)$, $S_{27}(70969) = S_{27}(76999)$, $S_{28}(10971) = S_{27}(16893) = S_{26}(12798)$, $S_{29}(13297) = S_{29}(17797) = S_{29}(15097) = S_{29}(19597) = S_{28}(80549) = S_{28}(84599) = S_{28}(80189) = S_{28}(88199)$, $S_{29}(60879) = S_{29}(67899)$, $S_{30}(40849) = S_{30}(44899)$, $S_{30}(60869) = S_{30}(66899)$, $S_{31}(13293) = S_{30}(10548) = S_{29}(90099)$, $S_{31}(89799)$, $S_{33}(20889) = S_{33}(28899) = S_{32}(17793) = S_{31}(14598)$, $S_{38}(13697) =$

$S_{38}(18197) = S_{37}(80359) = S_{37}(85399)$, $S_{39}(10828) = S_{38}(90659) = S_{38}(95699)$, $S_{40}(15891) = S_{39}(10794) = S_{38}(20499)$, $S_{47}(70759) = S_{47}(75799)$, and $S_{55}(10911) = S_{54}(10833) = S_{53}(10677) = S_{52}(70269) = S_{52}(76299) = 466\ 87315\ 96684\ 22486\ 69513\ 78664$.

Palindrome-free sequences. Including those reported in [1], there are 6091 integers $< 10^5$ which lead to no palindromes with fewer than 201 digits. These are representable by 263 basic integers whose reversal-addition sequences merge into 74 distinct sequences. The greatest number of sequences merging into a single sequence is five. No merger occurs beyond any $k = 9$. Both extreme situations occur in $S_9(196) = S_8(689)^* = S_7(1495)^* = S_6(4079)^* = S_6(4799) = S_5(12793)^* = S_4(10538)^* = S_3(90079)^* = S_3(90979) = S_3(97099) = S_3(97999)$. Either a basic integer, marked with an asterisk (*), or one of its family equivalents occurs as a nonstarter in the body of the merged sequences.

A sequence starter for each of the 74 distinct sequences is given below. The five smallest were dealt with in [1]. There are no mergers with the sequences of the five starters preceded by a Δ.

196	10715	10985	19591	50569	80069
879	10735	12393	19792	50579	80269
1997	10783	12797	20459	Δ59299	80279
7059	10785	Δ12898	29499	60289	80289
9999	10787	13097	29899	60489	80439
10553	10843	13197	30389	60659	80489
10563	10883	13393	30399	60689	80669
10577	10933	14698	30929	70379	Δ89099
10583	10963	15297	30959	70559	Δ89899
10585	10965	15393	30979	70949	90379
10638	10969	15597	40069	80049	90579
10663	10977	17190	40429	80059	Δ99999
10668	10983				

Each of these starting integers was fed into an IBM computer by R. S. Cook (who wrote the program) and V. Grannell of Los Angeles City College. Each sequence was extended to the first sum at the 200-digit level. The k's required ranged from 439 for $N = 30979$ to 503 for $N = 29499$. No palindromes or further mergers appeared. Nor did any sequence contain a pattern discernable to us which would assure that no palindrome would appear if the sequence were extended further.

On an IBM 1620, Cook and Grannell extended one sequence to the 4001-digit $S_{9600}(9999)$ without finding any palindromes. On an IBM 360 they developed another sequence to the 1820-digit $S_{4400}(89)$ without encountering any palindrome beyond $S_{24}(89)$. Thus the sequence ending in $S_{4376}(8813200023188)$ is palindrome-free. Don Wall [1] developed several palindrome-free sequences to $S_{4147}(196)$, $S_{3765}(879)$, and $S_{3717}(1997)$ on an IBM 1401.

General observations and conclusions.

1. There are 1515 five-digit integers which need 24 or more reversal-additions to produce a palindrome. The largest first palindrome found in this study and the one requiring the greatest number of operations to produce it is the 28-digit $S_{55}(10911)$.

2. There are two curious consecutive near-palindromic sums consisting principally of pairs of like digits:

$$S_{22}(80289) = 27\ 9988\ 6655\ 8899\ 72,$$

$$S_{23}(80289) = 55\ 9977\ 2222\ 7799\ 44.$$

3. The ratios of 1-digit, 2-digit, 3-digit, 4-digit, and 5-digit integers whch produce palindromes for $k < 439$ to the number of integers in each category, respectively, are 9/9, 90/90, 887/900, 8764/9000, and 84158/90000—a decreasing sequence.

4. There are 74 reversal-addition sequences which are palindrome-free up to at least the 200-digit level, all involving over 438 operations. Four of these sequences and one other have been carried beyond $k = 3710$ without producing a palindrome.

5. The only palindrome in the sequence ending in $S_{4400}(89)$ is $S_{24}(89)$.

6. The conjecture that every N produces a palindrome when subjected to the reiterated reversal-addition operation implies that no sequence thus produced contains a last palindrome. Integers with p digits which do not appear in sequences started with integers of fewer than p digits will start or appear in new sequences. Thus the conjecture requires an infinitude of sequences each containing an infinity of palindromes. (A study of the number of palindromes occurring in various sequences will be reported in a subsequent article.)

7. Items 3, 4, 5 and 6 together tend to negate the conjecture.

8. To replace this unlikely bit of mathematical folklore, we conjecture that in every system of numeration with a positive integral base:

A. For every N, the related reversal-addition sequence S_k is palindrome-free for $k > m$, a constant depending on N;

B. Despite the merging of certain sequences, there is an infinitude of infinitely long palindrome-free disjoint reversal-addition sequences;

C. For every k there is an N whose S_k is a first palindrome. [$S_{55}(10911)$ establishes this for $k < 56$.]

Reference

1. C. W. Trigg, Palindromes by addition, this MAGAZINE, 40 (1967) 26–28.

A NOTE ON PALINDROMES BY REVERSAL-ADDITION

MICHAEL T. REBMANN and FRANK SENTYRZ JR., University of Minnesota

Let N' be the integer obtained by writing the digits of the integer N in reverse order. Define the first *versum* $S_1(N) = N + N'$, and in general the kth versum

General observations and conclusions.

1. There are 1515 five-digit integers which need 24 or more reversal-additions to produce a palindrome. The largest first palindrome found in this study and the one requiring the greatest number of operations to produce it is the 28-digit $S_{55}(10911)$.

2. There are two curious consecutive near-palindromic sums consisting principally of pairs of like digits:

$$S_{22}(80289) = 27\ 9988\ 6655\ 8899\ 72,$$

$$S_{23}(80289) = 55\ 9977\ 2222\ 7799\ 44.$$

3. The ratios of 1-digit, 2-digit, 3-digit, 4-digit, and 5-digit integers whch produce palindromes for $k < 439$ to the number of integers in each category, respectively, are 9/9, 90/90, 887/900, 8764/9000, and 84158/90000—a decreasing sequence.

4. There are 74 reversal-addition sequences which are palindrome-free up to at least the 200-digit level, all involving over 438 operations. Four of these sequences and one other have been carried beyond $k = 3710$ without producing a palindrome.

5. The only palindrome in the sequence ending in $S_{4400}(89)$ is $S_{24}(89)$.

6. The conjecture that every N produces a palindrome when subjected to the reiterated reversal-addition operation implies that no sequence thus produced contains a last palindrome. Integers with p digits which do not appear in sequences started with integers of fewer than p digits will start or appear in new sequences. Thus the conjecture requires an infinitude of sequences each containing an infinity of palindromes. (A study of the number of palindromes occurring in various sequences will be reported in a subsequent article.)

7. Items 3, 4, 5 and 6 together tend to negate the conjecture.

8. To replace this unlikely bit of mathematical folklore, we conjecture that in every system of numeration with a positive integral base:

A. For every N, the related reversal-addition sequence S_k is palindrome-free for $k > m$, a constant depending on N;

B. Despite the merging of certain sequences, there is an infinitude of infinitely long palindrome-free disjoint reversal-addition sequences;

C. For every k there is an N whose S_k is a first palindrome. [$S_{55}(10911)$ establishes this for $k < 56$.]

Reference

1. C. W. Trigg, Palindromes by addition, this MAGAZINE, 40 (1967) 26–28.

A NOTE ON PALINDROMES BY REVERSAL-ADDITION

MICHAEL T. REBMANN and FRANK SENTYRZ JR., University of Minnesota

Let N' be the integer obtained by writing the digits of the integer N in reverse order. Define the first *versum* $S_1(N) = N + N'$, and in general the kth versum

$S_k(N) = S_1(S_{k-1}(N))$ for $k > 1$. A conjecture cited by Trigg [5] states that for every N there is a k such that the kth versum of N is palindromic, that is, that $S_k(N) = (S_k(N))'$. Although this conjecture has been proved false for integers N expressed in bases of the form 2^n [1, 2], it remains unresolved for any other base.

In [4], Trigg displays 6091 decimal integers $N < 10^5$ for which $S_k(N)$ is non-palindromic for each versum of at most 200 digits. It is also shown sufficient to consider only 74 integers to decide for all 6091 whether versum formation will yield a palindrome. To date, the most extensive investigations [3, 4, 5] have reported $S_{4147}(196)$, $S_{3765}(879)$, $S_{3717}(1997)$, and $S_{9600}(9999)$ with no palindrome encountered.

In the spirit that "just one more step" was required, we developed S_{10000} of all 74 integers on the CDC 6600 in use at the University of Minnesota; in each case no palindrome was encountered. On the way to S_{10000}, we displayed the first versum of exactly 4000 digits to determine whether any versum sequences merged. All 4000 digit versums were different, showing 74 distinct sequences, even to this large number of digits. Including input/output routines, the average computation time for S_{10000} (plus checking for palindromes) was 75 seconds, giving us about 4150 digits per number.

References

1. Brother A. Brousseau, Palindromes by addition in base two, this MAGAZINE, 42 (1969) 254–256.
2. D. C. Duncan, Sujet d'étude, No. 74, Sphinx (Bruxelles), 9 (1939) 91–92.
3. M. Gardner, Mathematical games, Scientific American, 223, 2 (1970) 110–112.
4. C. W. Trigg, More on palindromes by reversal-addition, this MAGAZINE, 45 (1972) 184–186.
5. C. W. Trigg, Palindromes by addition, this MAGAZINE, 40 (1967) 26–28.

INEQUALITIES FOR $\sigma(n)$ AND $\phi(n)$

U. ANNAPURNA, Andhra University, Waltair, India

1. Introduction. Let $\sigma(n)$ denote the sum of all the positive divisors of n and $\phi(n)$ denote the Euler's totient function, which is defined to be the number of positive integers $\leqq n$ and relatively prime to n. Recently, A. M. Vaidya [2] proved the following inequality:

(a) $\phi(n) \geqq \sqrt{n}$, except for $n = 2$ or $n = 6$. It is well known that

(b) $\dfrac{6}{\pi^2} < \dfrac{\sigma(n)\phi(n)}{n^2} < 1$ for all $n > 1$.

For a proof of (b), we refer to Hardy and Wright [1, Theorem 329]. In virtue of (a) and (b), it follows that $\sigma(n)/n\sqrt{n} < 1$ for all $n > 1$, except for $n = 2$ or $n = 6$. It is easy to verify that $\sigma(n)/n\sqrt{n} < 1$ for $n = 2$ and $n = 6$ also, so that we have

(c) $\dfrac{\sigma(n)}{n\sqrt{n}} < 1$ for all $n > 1$.

It can be observed that (b) and (c) do not imply (a). The question that arises

$S_k(N) = S_1(S_{k-1}(N))$ for $k > 1$. A conjecture cited by Trigg [5] states that for every N there is a k such that the kth versum of N is palindromic, that is, that $S_k(N) = (S_k(N))'$. Although this conjecture has been proved false for integers N expressed in bases of the form 2^n [1, 2], it remains unresolved for any other base.

In [4], Trigg displays 6091 decimal integers $N < 10^5$ for which $S_k(N)$ is non-palindromic for each versum of at most 200 digits. It is also shown sufficient to consider only 74 integers to decide for all 6091 whether versum formation will yield a palindrome. To date, the most extensive investigations [3, 4, 5] have reported $S_{4147}(196)$, $S_{3765}(879)$, $S_{3717}(1997)$, and $S_{9600}(9999)$ with no palindrome encountered.

In the spirit that "just one more step" was required, we developed S_{10000} of all 74 integers on the CDC 6600 in use at the University of Minnesota; in each case no palindrome was encountered. On the way to S_{10000}, we displayed the first versum of exactly 4000 digits to determine whether any versum sequences merged. All 4000 digit versums were different, showing 74 distinct sequences, even to this large number of digits. Including input/output routines, the average computation time for S_{10000} (plus checking for palindromes) was 75 seconds, giving us about 4150 digits per number.

References

1. Brother A. Brousseau, Palindromes by addition in base two, this MAGAZINE, 42 (1969) 254–256.
2. D. C. Duncan, Sujet d'étude, No. 74, Sphinx (Bruxelles), 9 (1939) 91–92.
3. M. Gardner, Mathematical games, Scientific American, 223, 2 (1970) 110–112.
4. C. W. Trigg, More on palindromes by reversal-addition, this MAGAZINE, 45 (1972) 184–186.
5. C. W. Trigg, Palindromes by addition, this MAGAZINE, 40 (1967) 26–28.

INEQUALITIES FOR $\sigma(n)$ AND $\phi(n)$

U. ANNAPURNA, Andhra University, Waltair, India

1. Introduction. Let $\sigma(n)$ denote the sum of all the positive divisors of n and $\phi(n)$ denote the Euler's totient function, which is defined to be the number of positive integers $\leqq n$ and relatively prime to n. Recently, A. M. Vaidya [2] proved the following inequality:

(a) $\phi(n) \geqq \sqrt{n}$, except for $n = 2$ or $n = 6$. It is well known that

(b) $\dfrac{6}{\pi^2} < \dfrac{\sigma(n)\phi(n)}{n^2} < 1$ for all $n > 1$.

For a proof of (b), we refer to Hardy and Wright [1, Theorem 329]. In virtue of (a) and (b), it follows that $\sigma(n)/n\sqrt{n} < 1$ for all $n > 1$, except for $n = 2$ or $n = 6$. It is easy to verify that $\sigma(n)/n\sqrt{n} < 1$ for $n = 2$ and $n = 6$ also, so that we have

(c) $\dfrac{\sigma(n)}{n\sqrt{n}} < 1$ for all $n > 1$.

It can be observed that (b) and (c) do not imply (a). The question that arises naturally in this connection is the following: What improvement in (c) when combined with (b) will yield (a)? It is clear that the answer to this question is that the inequality $\sigma(n)/n\sqrt{n} < 6/\pi^2$ when combined with (b) will yield (a).

The main object of this note is to prove the following:

(d) $\dfrac{\sigma(n)}{n\sqrt{n}} < \dfrac{6}{\pi^2}$ for all $n > 1$, except for $n = 2,3,4,6,8$ and 12.

It may be noted that (a) and (b) do not imply (d) and in this sense (d) is stronger than (a). Another question that arises in this connection is the following: Can Vaidya's inequality, namely (a), be sharpened so that this sharpened version when used with (b) will give (d)? We cannot answer this question now.

In this note we also prove the following which is stronger than (d) for certain values of n: Let p denote a prime number and $H(p,n)$ denote the nonnegative integer α such that $p^{\alpha} | n$ and $p^{\alpha+1} \nmid n$. Then for all even n with $H(2,n) \geqq 4$ and for all odd $n > 1$, with $H(3,n) \neq 1$,

(e) $\dfrac{\sigma(n)}{n\sqrt{n}} < \left(\dfrac{6}{\pi^2}\right)^{\omega(n)}$

where $\omega(n)$ is the number of distinct prime factors of n.

It can be easily seen that (e) when combined with (b) will give

(f) $\dfrac{\phi(n)}{\sqrt{n}} > \left(\dfrac{\pi^2}{6}\right)^{\omega(n)-1}$

which is stronger than (a) for all even n, with $H(2,n) \geqq 4$ and for all odd $n > 1$, with $H(3,n) \neq 1$.

2. A lemma. Before we give the proofs of (d) and (e), we prove the following lemma, which is essential in our discussion. It should be mentioned that the techniques adopted in this note are similar to those of Vaidya [2] in establishing (a).

LEMMA. *If* $f(x) = (x^{m+1} - 1)/(x - 1)x^{3m/2}$, *then* $f(x)$ *is monotonically decreasing for* $x \geqq 2$ *and* $m \geqq 1$.

Proof. To prove this lemma, we prove that $f'(x) < 0$ for $x \geqq 2$ and $m \geqq 1$, where $f'(x)$ is the derivative of $f(x)$. It is clear that

$$(x-1)^2 x^{3m} f'(x) = (x-1)x^{3m/2}(m+1)x^m - (x^{m+1}-1)\left\{x^{3m/2} + (x-1)\frac{3m}{2}x^{(3m/2)-1}\right\},$$

so that

$$(x-1)^2 x^{(3m/2)+1} f'(x) = (x-1)(m+1)x^{m+1} - (x^{m+1}-1)\left\{x + \frac{3m}{2}(x-1)\right\} < 0,$$

provided

$$(x-1)(m+1)(x^{m+1}-1+1) < (x^{m+1}-1)\left\{\left(1+m+\frac{m}{2}\right)(x-1)+1\right\},$$

that is,

$$(x-1)(m+1)(x^{m+1}-1)+(x-1)(m+1)$$
$$<(x^{m+1}-1)(m+1)(x-1)+\frac{m}{2}(x^{m+1}-1)(x-1)+(x^{m+1}-1),$$

that is,

$$m+1<\frac{m}{2}(x^{m+1}-1)+\frac{x^{m+1}-1}{x-1},$$

that is,

$$\frac{3m}{2}+1<\frac{m}{2}x^{m+1}+\frac{x^{m+1}-1}{x-1},$$

which is true, since

$$\frac{m}{2}x^{m+1}\geqq\frac{m}{2}2^{m+1}\geqq\frac{3m}{2},\quad\text{and}\quad\frac{x^{m+1}-1}{x-1}\geqq x+1>1.$$

Hence $f'(x)<0$ for $x\geqq 2$ and $m\geqq 1$, so that $f(x)$ is monotonically decreasing.

3. Proof of (d). Let $n=2^\alpha 3^\beta\prod_{p|n}p^\gamma$, where $\alpha\geqq 0$, $\beta\geqq 0$, $\gamma\geqq 0$ (not all simultaneously zero), with the convention that the empty product is taken to be 1.

(i) If $\alpha\geqq 4$, then by the lemma,

$$\frac{\sigma(2^\alpha)}{2^{3\alpha/2}}=\frac{2^{\alpha+1}-1}{2^{3\alpha/2}}\leqq\frac{31}{64}<\frac{6}{\pi^2}.$$

Also, if $\beta\geqq 1$, then by the lemma,

$$\frac{\sigma(2^\alpha\cdot 3^\beta)}{(2^\alpha\cdot 3^\beta)^{\frac{3}{2}}}=\frac{\sigma(2^\alpha)}{2^{3\alpha/2}}\cdot\frac{3^{\beta+1}-1}{2\cdot 3^{3\beta/2}}\leqq\frac{31}{64}\cdot\frac{4}{3\sqrt{3}}<\frac{6}{\pi^2}.$$

(ii) Suppose $\beta\geqq 2$, then by the lemma,

$$\frac{\sigma(3^\beta)}{3^{3\beta/2}}=\frac{3^{\beta+1}-1}{2\cdot 3^{3\beta/2}}\leqq\frac{13}{27}<\frac{6}{\pi^2}.$$

Also, it is easy to verify, in a similar fashion as above, that

$$\frac{\sigma(2^\alpha\cdot 3^\beta)}{(2^\alpha\cdot 3^\beta)^{\frac{3}{2}}}<\frac{6}{\pi^2}\quad\text{for }\alpha=1,2,3;$$

and

$$\frac{\sigma(2^3\cdot 3)}{(2^3\cdot 3)^{\frac{3}{2}}}<\frac{6}{\pi^2}.$$

(iii) Suppose $p \geqq 5$ and $\beta \geqq 1$, $\gamma \geqq 1$. Then by the lemma,

$$\frac{\sigma(p^{\gamma})}{p^{3\gamma/2}} = \frac{p^{\gamma+1}-1}{(p-1)p^{3\gamma/2}} \leqq \frac{5^2-1}{4\cdot 5\sqrt{5}} = \frac{6}{5\sqrt{5}} < \frac{6}{\pi^2}.$$

Also, it is easy to verify that

$$\frac{\sigma(2^{\alpha}\cdot p^{\gamma})}{(2^{\alpha}\cdot p^{\gamma})^{\frac{3}{2}}} < \frac{6}{\pi^2} \quad \text{for } \alpha = 1,2,3;$$

and

$$\frac{\sigma(2^{\alpha}\cdot 3^{\beta}\cdot p^{\gamma})}{(2^{\alpha}\cdot 3^{\beta}\cdot p^{\gamma})^{\frac{3}{2}}} < \frac{6}{\pi^2} \quad \text{for } \alpha = 0,1,2.$$

Since $\sigma(n)/n\sqrt{n}$ is multiplicative, (d) follows by making use of the above discussion in (i), (ii) and (iii) and using the fact that $6/\pi^2 < 1$ whenever necessary.

4. Proof of (e). Let $n = \prod_{i=1}^{r} p_i^{\alpha_i}$ be the canonical representation of $n > 1$, then

$$\frac{\sigma(n)}{n\cdot\sqrt{n}} = \prod_{i=1}^{r}\left[\frac{\sigma(p_i^{\alpha_i})}{p_i^{3\alpha_i/2}}\right] < \left(\frac{6}{\pi^2}\right)^r,$$

by the discussion in the proof of (d) and the hypothesis of (e).

5. Acknowledgement. The author is grateful to Dr. D. Suryanarayana for his help and guidance during the preparation of this note.

References

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 4th ed., Oxford, New York, 1965.
2. A. M. Vaidya, An inequality for Euler's totient function, Math. Student, 35 (1967) 79–80.

SERIES REPRESENTATION OF ELEMENTS IN SEPARABLE BANACH SPACES

ROBERT C. HANSEN, University of Missouri

If b is any real number and $\{a_k\}_{k=1}^{\infty}$ is a real sequence satisfying

(1) $$\sum_{k=1}^{\infty} |a_k| = +\infty \text{ and } \lim_{k\to\infty} a_k = 0$$

(iii) Suppose $p \geqq 5$ and $\beta \geqq 1$, $\gamma \geqq 1$. Then by the lemma,

$$\frac{\sigma(p^\gamma)}{p^{3\gamma/2}} = \frac{p^{\gamma+1}-1}{(p-1)p^{3\gamma/2}} \leqq \frac{5^2-1}{4 \cdot 5\sqrt{5}} = \frac{6}{5\sqrt{5}} < \frac{6}{\pi^2}.$$

Also, it is easy to verify that

$$\frac{\sigma(2^\alpha \cdot p^\gamma)}{(2^\alpha \cdot p^\gamma)^{\frac{3}{2}}} < \frac{6}{\pi^2} \quad \text{for } \alpha = 1,2,3;$$

and

$$\frac{\sigma(2^\alpha \cdot 3^\beta \cdot p^\gamma)}{(2^\alpha \cdot 3^\beta \cdot p^\gamma)^{\frac{3}{2}}} < \frac{6}{\pi^2} \quad \text{for } \alpha = 0,1,2.$$

Since $\sigma(n)/n\sqrt{n}$ is multiplicative, (d) follows by making use of the above discussion in (i), (ii) and (iii) and using the fact that $6/\pi^2 < 1$ whenever necessary.

4. Proof of (e). Let $n = \prod_{i=1}^r p_i^{\alpha_i}$ be the canonical representation of $n > 1$, then

$$\frac{\sigma(n)}{n \cdot \sqrt{n}} = \prod_{i=1}^r \left[\frac{\sigma(p_i^{\alpha_i})}{p_i^{3\alpha_i/2}}\right] < \left(\frac{6}{\pi^2}\right)^r,$$

by the discussion in the proof of (d) and the hypothesis of (e).

5. Acknowledgement. The author is grateful to Dr. D. Suryanarayana for his help and guidance during the preparation of this note.

References

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 4th ed., Oxford, New York, 1965.
2. A. M. Vaidya, An inequality for Euler's totient function, Math. Student, 35 (1967) 79–80.

SERIES REPRESENTATION OF ELEMENTS IN SEPARABLE BANACH SPACES

ROBERT C. HANSEN, University of Missouri

If b is any real number and $\{a_k\}_{k=1}^\infty$ is a real sequence satisfying

$$(1) \qquad \sum_{k=1}^\infty |a_k| = +\infty \text{ and } \lim_{k\to\infty} a_k = 0$$

then

(A) there is a sequence $\{t_k\}_{k=1}^{\infty}$ with each $t_k = +1$ or -1 such that

$$\sum_{k=1}^{\infty} a_k t_k = b,$$

(B) if $\{t_k\}_{k=1}^{\infty}$ is a sequence of $+1$'s and -1's with infinitely many $+1$'s and infinitely many -1's, then there is a rearrangement $\{a_{n(k)}\}_{k=1}^{\infty}$ of $\{a_k\}_{k=1}^{\infty}$ such that $\sum_{k=1}^{\infty} a_{n(k)} t_k = b$.

These are familiar facts. Regarding the real numbers as a separable Banach space of dimension one we obtain generalizations of these facts to a separable Banach space X of dimension greater than one. In this more general setting a sequence $\{x_k\}_{k=1}^{\infty}$ of elements of X satisfying

$$(2) \qquad \|x_k\| = 1, \ k = 1,2,\cdots \text{ and}$$
$$\{x_k\}_{k=1}^{\infty} \text{ is dense in the subset } S(0,1) = \{x: \|x\| = 1\} \text{ of } X$$

plays the role of the sequence $\{t_k\}_{k=1}^{\infty}$ of $+1$'s and -1's.

THEOREM 1. *Let $\{a_k\}_{k=1}^{\infty}$ be a sequence satisfying* (1) *and $\{x_n\}_{n=1}^{\infty}$ a sequence satisfying* (2). *Then given $x \in X$ there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ such that $\sum_{k=1}^{\infty} a_k x_{n_k} = x$ (i.e., $\lim_{m\to\infty} \sum_{k=1}^{m} a_k x_{n_k} = x$).*

THEOREM 2. *Suppose $\{a_k\}_{k=1}^{\infty}$ satisfies* (1) *and $\{x_n\}_{n=1}^{\infty}$ satisfies* (2). *Then given $x \in X$ there exists a rearrangement of $\{a_k\}_{k=1}^{\infty}$, denoted by $\{a_{p(i)}\}_{i=1}^{\infty}$, such that $\sum_{i=1}^{\infty} a_{p(i)} x_i = x$.*

Theorems 1 and 2 generalize A and B respectively. To prove the theorems we use the following lemma:

LEMMA. *Suppose $\{x_n\}$ satisfies* (2) *and a is any real number such that $|a| \leqq r$ where $r > 0$. If $y \in D(0,r) = \{x: \|x - 0\| \leqq r\}$ then there is an x_i in $\{x_n\}$ such that $y + ax_i \in D(0,r)$.*

If $a = 0$ or $y = 0$ then any x_i will suffice. Now if $a \neq 0$ and $y \neq 0$ then since $\{x_i\}$ is dense in $S(0,1)$ we have $\{y + ax_i\}$ is dense in $S(y,|a|)$. Note that $y - |a| \, \|y\|^{-1} y \in S(y,|a|)$ since $\|(y - |a| \, \|y\|^{-1} y) - y\| = |a|$, and so to complete the proof we show $y - |a| \, \|y\|^{-1} y$ is in the interior of $D(0,r)$. This follows since

$$\|y - |a| \, \|y\|^{-1} y\| = \| \, \|y\| \, \|y\|^{-1} y - |a| \, \|y\|^{-1} y\| = \Big| \|y\| - |a| \Big| < r$$

(the inequality obtained by adding the known inequalities $0 < \|y\| \leqq r$ and $-r \leqq -|a| < 0$).

Proof of Theorem 1. We assume without loss of generality $0 < |a_k| \leqq 1$, for all k. Inductively we choose a strictly increasing sequence of positive integers $\{l_i\}_{i=1}^{\infty}$ such that if $k > l_i$ then $|a_k| < 1/2^i$. Let $x \in X$, $x \neq 0$. (The case when $x = 0$ will be dealt with later.) Since $\sum_{k=1}^{\infty} |a_k| = +\infty$, there exists a least positive integer L_1 such that $\|x\| < \sum_{k=1}^{L_1}|a_k| \leqq \|x\| + 1$. Note that $1 \geqq \delta_1 = \sum_{k=1}^{L_1}|a_k| - \|x\| > 0$. Since $\{x_n\}$ is dense in S, we may choose $x_{n_1}, x_{n_2}, \cdots, x_{n_{L_1}}$ with $n_1 < n_2 < \cdots < n_{L_1}$ such that if $a_k > 0$, $\|x_{n_k} - \|x\|^{-1}x\| < \eta_1$ and if $a_k < 0$, $\|x_{n_k} + \|x\|^{-1}x\| < \eta_1$ where $\eta_1 = (\sum_{k=1}^{L_1}|a_k|)^{-1}$. Note that if $a_k > 0$ then $\|a_kx_{n_k} - \|x\|^{-1}|a_k|x\| = |a_k|\,\|x_{n_k} - \|x\|^{-1}x\| < |a_k|\eta_1$, and if $a_k < 0$ then $\|a_kx_{n_k} - \|x\|^{-1}|a_k|x\| = |a_k|\,\|x_{n_k} + \|x\|^{-1}x\| < |a_k|\eta_1$ for $1 \leqq k \leqq L_1$.

Now

$$\begin{aligned}
\left\|\sum_{k=1}^{L_1} a_kx_{n_k} - x\right\| &= \left\|\sum_{k=1}^{L_1} a_kx_{n_k} - \|x\|^{-1}\|x\|x\right\| \\
&= \left\|\sum_{k=1}^{L_1} a_kx_{n_k} - \left[\sum_{k=1}^{L_1}|a_k|\right]\|x\|^{-1}x + \left[\sum_{k=1}^{L_1}|a_k| - \|x\|\right]\|x\|^{-1}x\right\| \\
&\leqq \sum_{k=1}^{L_1}\left\|a_kx_{n_k} - \|x\|^{-1}|a_k|x\right\| + \|\delta_1\|x\|^{-1}x\| \\
&= \sum_{k=1}^{L_1}|a_k|\,\left\|x_{n_k} \pm \|x\|^{-1}x\right\| + \delta_1 \\
&\leqq \left[\sum_{k=1}^{L_1}|a_k|\right]\eta_1 + \delta_1,
\end{aligned}$$

where the choice of signs in the next to last inequality depends on whether a_k is positive or negative. We see from the above that $\|\sum_{k=1}^{L_1}a_kx_{n_k} - x\| \leqq 1 + \delta_1 \leqq 2$.

Now either $L_1 < l_1$ or $L_1 \geqq l_1$. If $L_1 < l_1$, we choose $x_{n_{L_1}+1}$ such that $a_{L_1+1}x_{n_{L_1}+1} \in D(0,1)$. By using the lemma we choose $x_{n_{L_1}+2}$ such that $\sum_{k=L_1+1}^{L_1+2} a_kx_{n_k} \in D(0,1)$. In a similar fashion we can choose x_{n_k}, $L_1 + 1 \leqq k \leqq l_1$ such that $\sum_{k=L_1+1}^{r}a_kx_{n_k} \in D(0,1)$ for $L_1 + 1 \leqq r \leqq l_1$. We thus obtain $\|\sum_{k=1}^{l_1}a_kx_{n_k} - x\| \leqq \|\sum_{k=1}^{L_1}a_kx_{n_k} - x\| + \|\sum_{k=L_1+1}^{l_1}a_kx_{n_k}\| \leqq 2 + 1 = 3$. Let $M_1 = \max\{L_1, l_1\}$ then we have $\|\sum_{k=1}^{M_1}a_kx_{n_k} - x\| \leqq 3$. Proceeding by induction we obtain $\{L_k\}_{k=1}^{t}$, $\{M_k\}_{k=1}^{t}$, $\{x_{n_k}\}_{k=1}^{M_t}$ and $\{z_k\}_{k=1}^{t}$ where $z_k = x - \sum_{i=1}^{M_k}a_ix_{n_i}$,

$$\left\|\sum_{i=1}^{L_k} a_ix_{n_i} - x\right\| \leqq 1/2^{k-2},$$

$\|z_k\| = \|\sum_{i=1}^{M_k}a_ix_{n_i} - x\| \leqq 3/2^{k-1}$ for $1 \leqq k \leqq t$, and $\sum_{L_k+1}^{r}a_kx_{n_k} \in D(0,1/2^{k-1})$ for $L_k + 1 \leqq r \leqq M_k$. Now since $\{x_n\}_{n=M_t+1}^{\infty}$ satisfies (2) and $\{a_k\}_{M_t+1}^{\infty}$ satisfies (1) and $|a_k| \leqq 1/2^t$ for $k \geqq M_t + 1$, we can choose the least positive integer L_{t+1} such that $\|z_t\| < \sum_{k=M_t+1}^{L_{t+1}}|a_k| \leqq \|z_t\| + 1/2^t$. As in the previous calculations, assuming $z_t \neq 0$, with $\eta_t = (2^{t-1}\sum_{k=M_t+1}^{L_{t+1}}|a_k|)^{-1}$ we can choose x_{n_k}, $M_t + 1 \leqq k \leqq L_{t+1}$, in increasing fashion such that

$$\left\|\sum_{k=1}^{L_{t+1}} a_kx_{n_k} - x\right\| = \left\|\sum_{k=M_t+1}^{L_{t+1}} a_kx_{n_k} - z_t\right\| < 1/2^{t-1}.$$

By using the lemma we can choose x_{n_k} for $L_{t+1}+1 \leqq k \leqq M_{t+1} = \min\{L_{t+1}, l_{t+1}\}$ such that $\sum_{k=L_{t+1}+1}^{r} a_k x_{n_k} \in D(0, 1/2^t)$ for $L_{t+1}+1 \leqq r \leqq M_{t+1}$. Thus we have

$$\left\| \sum_{k=1}^{M_{t+1}} a_k x_{n_k} - x \right\| \leqq \left\| \sum_{k=1}^{L_{t+1}} a_k x_{n_k} - x \right\| + \left\| \sum_{k=L_{t+1}+1}^{M_{t+1}} a_k x_{n_k} \right\|$$

$$\leqq 1/2^{t-1} + 1/2^t = 3/2^t.$$

If at any step in the induction we obtain $z_n = 0$ (or $x = 0$ in the beginning) let $L_{n+1} = M_n + 1$ ($L_1 = 1$) and proceed as usual.

Now let $S_m = \sum_{k=1}^{m} a_k x_{n_k}$ in the series we have constructed by induction. We have shown that $\lim_{k\to\infty} S_{M_k} = x$, (also $\lim_{k\to\infty} S_{L_k} = x$) in the norm topology and it remains to show $\lim_{k\to\infty} S_k = x$. Let t be a positive integer with $L_n < t < M_n$. Then $\| S_t - x \| \leqq \| S_{L_n} - x \| + \| S_t - S_{L_n} \| \leqq 1/2^{n-2} + 1/2^{n-1} = 3/2^{n-1}$. Now suppose $M_n < t < L_{n+1}$. Then

$$\begin{aligned} \| S_t - x \| &= \| S_{M_n} - x \| + \| S_t - S_{M_n} \| \\ &\leqq 3/2^{n-1} + \sum_{k=M_n+1}^{t} \| a_k x_{n_k} \| \\ &\leqq 3/2^{n-1} + \sum_{k=M_n+1}^{t} | a_k | \\ &\leqq 3/2^{n-1} + \| z_n \| + 1/2^{n-1} \\ &\leqq 3/2^{n-1} + 3/2^{n-1} + 1/2^{n-1}. \end{aligned}$$

It now follows that $\lim_{k\to\infty} S_k = x$.

Proof of Theorem II. Given $\{a_k\}$, extract a subsequence $\{a_{n_k}\}$ such that $| a_{n_k} | \leqq 1/2^k$ for each k. Let $\{a_{m_k}\}$ be the remaining subsequence after $\{a_{n_k}\}$ has been extracted. Rename $\{a_{m_k}\}$ as $\{b_k\}$ and $\{a_{n_k}\}$ as $\{d_k\}$ and note that $\{b_k\}$ satisfies (1). Let $x \in X$ be given. As in the proof of the previous theorem we can obtain x_{n_k}, $1 \leqq k \leqq M_1$, so $\| \sum_{k=1}^{M_1} b_k x_{n_k} - x \| < 3$. Let $x_{m_1}, x_{m_2}, \cdots, x_{m_{R_1}}$ be the elements of $\{x_k\}$ for $1 \leqq k \leqq n_{M_1} + 1$ not already used and note $\| \sum_{k=1}^{R_1} d_k x_{m_k} \| < 2$. Thus if we define $a_{p(i)}$ for $1 \leqq i \leqq n_{M_1} + 1$ to be the appropriate b_k or d_k as the above we obtain

$$\left\| \sum_{i=1}^{n_{M_1}+1} a_{p(i)} x_i - x \right\| \leqq \left\| \sum_{k=1}^{M_1} b_k x_n - x \right\| + \left\| \sum_{j=1}^{R_1} d_j x_{n_j} \right\|$$

$$< 3 + 2 = 5.$$

Proceeding by induction we obtain increasing sequences of real numbers $\{M_k\}_{k=1}^{\infty}$ and $\{R_k\}_{k=1}^{\infty}$, a rearrangement of $\{a_k\}$ denoted by $\{a_{p(i)}\}_{i=1}^{\infty}$ and a sequence of elements $\{z_k\}_{k=1}^{\infty}$ with $z_k = x - \sum_{i=1}^{n_{M_k}} a_{p(i)} x_i$ such that

$$\left\| \sum_{i=1}^{n_{M_k}} a_{p(i)} x_i - x \right\| = \left\| \sum_{i=n_{M_{k-1}}+1}^{n_{M_k}} a_{p(i)} x_i - z_{k-1} \right\|$$

$$\leq \left\| \sum_{i=M_{k-1}+1}^{M_k} b_k x_{n_k} - z_{k-1} \right\| + \left\| \sum_{j=R_{k-1}+1}^{R_k} d_k x_{m_k} \right\|$$

$$< 3/2^{k-1} + 1/2^{k-2} = 5/2^{k-1}$$

for every k. The above induction shows that $\lim_{k\to\infty} \sum_{i=1}^{n_{M_k}+1} a_{p(i)} x_i = x$. To show that $\lim_{m\to\infty} \sum_{i=1}^{m} a_{p(i)} x_i = x$ we can use arguments similar to those in the proof of Theorem I and the fact that $\lim_{k\to\infty} \sum_{i=k}^{\infty} |d_i| = 0$.

The author wishes to express his gratitude to the referee for his suggestions and comments.

THE DISTRIBUTION OF QUADRATIC RESIDUES IN FIELDS OF ORDER p^2

G. E. BERGUM, South Dakota State University and
J. H. JORDAN, Washington State University

1. Introduction. In [1], N. R. Hardman and J. H. Jordan discuss the distribution of quadratic residues in fields of order p^2 where p is a prime of the form $4k + 3$. In this paper, we shall examine the distribution in fields of order p^2 where p is a prime of the form $8k + 5$.

Throughout this paper, small Greek letters will represent integers in $Z(\sqrt{-2})$ and Latin letters will represent rational integers with the exception of i which is the imaginary unit in the field of complex numbers. We will illustrate the integers in $Z(\sqrt{-2})$ by lattice points in a Cartesian coordinate system where the horizontal grid lines are $\sqrt{2}$ units apart and the vertical grid lines are one unit apart. We shall say that $\rho = a + b\sqrt{2}i$ is a prime iff $\rho = \alpha\beta$ implies that α or β (but not both) is a unit. For $\alpha = a + b\sqrt{2}i$, we define the conjugate of α by $\bar{\alpha} = a - b\sqrt{2}i$ and the norm of α as $N(\alpha) = \alpha\bar{\alpha}$. Therefore, $N(\alpha) = a^2 + 2b^2$. We say that α is a unit iff α divides β for all β in $Z(\sqrt{-2})$. Since the norm is multiplicative, it is a trivial matter to show that α is a unit iff $N(\alpha) = 1$. Hence, the only units in $Z(\sqrt{-2})$ are ± 1. Obviously, if ρ is a prime so is $\bar{\rho}$. If ρ is a prime in $Z(\sqrt{-2})$ and in Z we say that it is a real prime in $Z(\sqrt{-2})$. Primes p in Z will, in contrast, be referred to as rational primes.

In [2], the following theorems are established.

THEOREM 1.1. (a) *$\rho = \pm\sqrt{2}i$ is a prime in $Z(\sqrt{-2})$.*

(b) *ρ is a real prime iff ρ is a rational prime congruent to* 5 *or* 7 modulo 8.

(c) *$\rho \neq \pm\sqrt{2}i$ is a nonreal prime iff $N(\rho)$ is a rational prime congruent to* 1 *or* 3 modulo 8.

THEOREM 1.2. *Let $\hat{\gamma}$ be the collection of lattice points within the rectangle whose vertices are $(\pm 1 \pm \sqrt{2}i)\gamma/2$ and on the half open line segments*

$$(\pm(-1 + \sqrt{2}i)\gamma/2,\ (-1 - \sqrt{2}i)\gamma/2]\,;$$

$\hat{\gamma}$ is a complete residue system modulo *γ and the cardinality of $\hat{\gamma}$ is $N(\gamma)$.*

$$\leq \left\| \sum_{i=M_{k-1}+1}^{M_k} b_k x_{n_k} - z_{k-1} \right\| + \left\| \sum_{j=R_{k-1}+1}^{R_k} d_k x_{m_k} \right\|$$

$$< 3/2^{k-1} + 1/2^{k-2} = 5/2^{k-1}$$

for every k. The above induction shows that $\lim_{k\to\infty} \sum_{i=1}^{n_{M_k}+1} a_{p(i)} x_i = x$. To show that $\lim_{m\to\infty} \sum_{i=1}^{m} a_{p(i)} x_i = x$ we can use arguments similar to those in the proof of Theorem I and the fact that $\lim_{k\to\infty} \sum_{i=k}^{\infty} |d_i| = 0$.

The author wishes to express his gratitude to the referee for his suggestions and comments.

THE DISTRIBUTION OF QUADRATIC RESIDUES IN FIELDS OF ORDER p^2

G. E. BERGUM, South Dakota State University and
J. H. JORDAN, Washington State University

1. Introduction. In [**1**], N. R. Hardman and J. H. Jordan discuss the distribution of quadratic residues in fields of order p^2 where p is a prime of the form $4k+3$. In this paper, we shall examine the distribution in fields of order p^2 where p is a prime of the form $8k+5$.

Throughout this paper, small Greek letters will represent integers in $Z(\sqrt{-2})$ and Latin letters will represent rational integers with the exception of i which is the imaginary unit in the field of complex numbers. We will illustrate the integers in $Z(\sqrt{-2})$ by lattice points in a Cartesian coordinate system where the horizontal grid lines are $\sqrt{2}$ units apart and the vertical grid lines are one unit apart. We shall say that $\rho = a + b\sqrt{2}i$ is a prime iff $\rho = \alpha\beta$ implies that α or β (but not both) is a unit. For $\alpha = a + b\sqrt{2}i$, we define the conjugate of α by $\bar{\alpha} = a - b\sqrt{2}i$ and the norm of α as $N(\alpha) = \alpha\bar{\alpha}$. Therefore, $N(\alpha) = a^2 + 2b^2$. We say that α is a unit iff α divides β for all β in $Z(\sqrt{-2})$. Since the norm is multiplicative, it is a trivial matter to show that α is a unit iff $N(\alpha) = 1$. Hence, the only units in $Z(\sqrt{-2})$ are ± 1. Obviously, if ρ is a prime so is $\bar{\rho}$. If ρ is a prime in $Z(\sqrt{-2})$ and in Z we say that it is a real prime in $Z(\sqrt{-2})$. Primes p in Z will, in contrast, be referred to as rational primes.

In [**2**], the following theorems are established.

THEOREM 1.1. (a) *$\rho = \pm\sqrt{2}i$ is a prime in $Z(\sqrt{-2})$.*

(b) *ρ is a real prime iff ρ is a rational prime congruent to* 5 *or* 7 modulo 8.

(c) *$\rho \neq \pm\sqrt{2}i$ is a nonreal prime iff $N(\rho)$ is a rational prime congruent to* 1 *or* 3 modulo 8.

THEOREM 1.2. *Let $\hat{\gamma}$ be the collection of lattice points within the rectangle whose vertices are $(\pm 1 \pm \sqrt{2}i)\gamma/2$ and on the half open line segments*

$$(\pm(-1+\sqrt{2}i)\gamma/2,\ (-1-\sqrt{2}i)\gamma/2]\,;$$

$\hat{\gamma}$ is a complete residue system modulo *γ and the cardinality of $\hat{\gamma}$ is $N(\gamma)$.*

It is easy to show that if ρ is a prime then $\hat{\rho}$ is a field with $N(\rho)$ elements. In particular, if ρ is a real prime then $\hat{\rho}$ is a field with p^2 elements.

Let $\alpha \neq 0$ be an element in $\hat{\rho}$. If $\eta^2 \equiv \alpha \pmod{\rho}$ is solvable, then α is said to be a quadratic residue modulo ρ, otherwise it is called a quadratic nonresidue modulo ρ. The quadratic residues for 5, 13, and 29 are illustrated respectively in Figures 1 through 3.

Quadratic Residues Modulo 5

·	x	·	x	·
x	·	·	·	x
x	x	0	x	x
x	·	·	·	x
·	x	·	x	·

FIG 1.

Quadratic Residues Modulo 13

x	·	x	x	·	·	·	·	·	x	x	·	x
·	x	x	·	·	x	·	x	·	·	x	x	·
x	·	x	·	x	·	·	·	x	·	x	·	x
·	x	·	x	x	·	·	·	x	x	·	x	·
·	·	·	x	x	x	·	x	x	x	·	·	·
x	x	·	·	·	x	·	x	·	·	·	x	x
x	x	x	x	x	x	0	x	x	x	x	x	x
x	x	·	·	·	x	·	x	·	·	·	x	x
·	·	·	x	x	x	·	x	x	x	·	·	·
·	x	·	x	x	·	·	·	x	x	·	x	·
x	·	x	·	x	·	·	·	x	·	x	·	x
·	x	x	·	·	x	·	x	·	·	x	x	·
x	·	x	x	·	·	·	·	·	x	x	·	x

FIG. 2.

If $\rho \neq \pm\sqrt{2}i$ and $\alpha \not\equiv 0 \pmod{\rho}$, we define $\langle\alpha/\rho\rangle = 1$ if α is a quadratic residue modulo ρ; $\langle\alpha/\rho\rangle = -1$ if α is a quadratic nonresidue modulo ρ. The symbol has the following properties of the Legendre symbol (a/p); namely, if $\alpha \equiv \beta \pmod{\rho}$ then $\langle\alpha/\rho\rangle = \langle\beta/\rho\rangle$ and $\langle\alpha/\rho\rangle\,\langle\beta/\rho\rangle = \langle\alpha\beta/\rho\rangle$.

Examining Figures 1–3, where p is a prime of the form $8k+5$, we make the following conjectures:

1. Every point on the x-axis is a quadratic residue.
2. No point on the y-axis is a quadratic residue.
3. $\langle\alpha/\rho\rangle = \langle-\alpha/\rho\rangle = \langle\bar{\alpha}/\rho\rangle = \langle-\bar{\alpha}/\rho\rangle$.
4. If $1+\sqrt{2}i$ is a quadratic residue (nonresidue) then $\pm a \pm a\sqrt{2}i$ is a quadratic residue (nonresidue).
5. The number of quadratic residues in $\hat{\rho}$ on every vertical line, with the exception of the y-axis, is one more than the number of quadratic nonresidues.
6. The number of quadratic residues in $\hat{\rho}$ on every horizontal line, with the exception of the x-axis, is one less than the number of quadratic residues.

The remainder of this paper is devoted to the proofs of these conjectures as well as to the establishment of a relationship between $\langle\,/\,\rangle$ and $(\,/\,)$. Some results which are not obvious from an examination of Figures 1–3 will also be given.

Quadratic Residues Modulo 29

·	·	·	x	x	x	·	x	x	·	·	x	·	x	·	x	·	x	·	·	x	x	·	x	x	x	·	·	·
·	·	·	x	·	x	x	·	·	·	x	x	x	x	·	x	x	x	x	·	·	·	x	x	·	x	·	·	·
x	x	·	·	·	·	x	·	x	x	·	x	·	x	·	x	·	x	·	x	x	·	x	·	·	·	·	x	x
·	x	x	·	x	x	x	x	·	x	·	·	·	·	·	·	·	·	·	x	·	x	x	x	x	·	x	x	·
·	·	x	·	·	x	·	·	x	x	x	x	x	·	·	·	x	x	x	x	x	·	·	x	·	·	x	·	·
·	·	x	x	x	·	x	·	x	x	x	·	·	·	·	·	·	·	x	x	x	·	x	·	x	x	x	·	·
x	x	·	·	x	x	·	·	·	·	x	·	x	x	·	x	x	·	x	·	·	·	·	x	x	·	·	x	x
x	x	·	x	x	x	·	·	·	x	·	x	·	·	·	·	·	x	·	x	·	·	·	x	x	x	·	x	x
x	x	x	·	·	·	x	x	·	·	x	x	·	·	·	·	·	x	x	·	·	x	x	·	·	·	x	x	x
·	x	x	·	x	·	·	·	x	·	·	x	x	x	·	x	x	x	·	·	x	·	·	·	x	·	x	x	·
x	·	·	·	x	·	x	x	·	x	·	·	x	x	·	x	x	·	·	x	·	x	x	·	x	·	·	·	x
·	x	·	x	·	·	x	x	x	·	x	·	x	·	·	·	x	·	x	·	x	x	x	·	·	x	·	x	·
x	·	x	x	·	·	·	x	·	x	x	·	·	x	·	x	·	·	x	x	·	x	·	·	·	x	x	·	x
x	·	x	x	·	x	·	x	x	·	·	·	x	·	·	·	x	·	·	·	x	x	·	x	·	x	x	·	x
x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x
x	·	x	x	·	x	·	x	x	·	·	·	x	·	·	·	x	·	·	·	x	x	·	x	·	x	x	·	x
x	·	x	x	·	·	·	x	·	x	x	·	·	x	·	x	·	·	x	x	·	x	·	·	·	x	x	·	x
·	x	·	x	·	·	x	x	x	·	x	·	x	·	·	·	x	·	x	·	x	x	x	·	·	x	·	x	·
x	·	·	·	x	·	x	x	·	x	·	·	x	x	·	x	x	·	·	x	·	x	x	·	x	·	·	·	x
·	x	x	·	x	·	·	·	x	·	·	x	x	x	·	x	x	x	·	·	x	·	·	·	x	·	x	x	·
x	x	x	·	·	·	x	x	·	·	x	x	·	·	·	·	·	x	x	·	·	x	x	·	·	·	x	x	x
x	x	·	x	x	x	·	·	·	x	·	x	·	·	·	·	·	x	·	x	·	·	·	x	x	x	·	x	x
x	x	·	·	x	x	·	·	·	·	x	·	x	x	·	x	x	·	x	·	·	·	·	x	x	·	·	x	x
·	·	x	x	x	·	x	·	x	x	x	·	·	·	·	·	·	·	x	x	x	·	x	·	x	x	x	·	·
·	·	x	·	·	x	·	·	x	x	x	x	x	·	·	·	x	x	x	x	x	·	·	x	·	·	x	·	·
·	x	x	·	x	x	x	x	·	x	·	·	·	·	·	·	·	·	·	x	·	x	x	x	x	·	x	x	·
x	x	·	·	·	·	x	·	x	x	·	x	·	x	·	x	·	x	·	x	x	·	x	·	·	·	·	x	x
·	·	·	x	·	x	x	·	·	·	x	x	x	x	·	x	x	x	x	·	·	·	x	x	·	x	·	·	·
·	·	·	x	x	x	·	x	x	·	·	x	·	x	·	x	·	x	·	·	x	x	·	x	x	x	·	·	·

FIG. 3.

2. Euler's criterion. The generalization of Euler's criterion would be

THEOREM 2.1. *If* $\alpha \not\equiv 0 \pmod{\rho}$ *then* $\langle\alpha/\rho\rangle \equiv \alpha^{(N(\rho)-1)/2} \pmod{\rho}$.

Proof. Since $\hat{\rho}$ is a finite field the nonzero elements form a cyclic group under multiplication. Let γ be a generator. If α is a quadratic residue modulo ρ then there exists a β such that $\beta^2 \equiv \alpha \pmod{\rho}$. Furthermore, there exists an integer n such that $\gamma^n \equiv \beta \pmod{\rho}$. Hence, $\gamma^{2n} \equiv \alpha \pmod{\rho}$ or

$$\gamma^{2n(N(\rho)-1)/2} \equiv \alpha^{(N(\rho)-1)/2} \pmod{\rho}.$$

Since $\gamma^{N(\rho)-1} \equiv 1 \pmod{\rho}$, we have $\alpha^{(N(\rho)-1)/2} \equiv \langle\alpha/\rho\rangle \pmod{\rho}$.

Suppose $\langle\alpha/\rho\rangle = -1$, then for some integer m we have $\gamma^{2m+1} \equiv \alpha \pmod{\rho}$ or

$$\gamma^{(2m+1)(N(\rho)-1)/2} \equiv \alpha^{(N(\rho)-1)/2} \pmod{\rho}.$$

But, γ is a generator so $\gamma^{(N(\rho)-1)/2} \equiv -1 \pmod{\rho}$. Therefore

$$\gamma^{(2m+1)(N(\rho)-1)/2} \equiv (-1)^{2m+1} \equiv -1 \pmod{\rho}$$

or $\alpha^{(N(\rho)-1)/2} \equiv \langle \alpha/\rho \rangle \pmod{\rho}$ and the theorem is proved.

For the remainder of this paper, unless stated otherwise, we will assume that ρ is a prime of the form $8k+5$. To establish conjecture one, we have

COROLLARY 2.1. $\langle a/p \rangle = 1$.

Proof. By Fermat's little theorem $a^{p-1} \equiv 1 \pmod{p}$. Hence,

$$a^{(p-1)(p+1)/2} \equiv a^{(N(p)-1)/2} \equiv 1 \pmod{p}.$$

For conjecture two, we show

COROLLARY 2.2. $\langle a\sqrt{2}i/p \rangle = -1$.

Proof.
$$\begin{aligned}\langle a\sqrt{2}i/p \rangle &\equiv (a\sqrt{2}i)^{(N(p)-1)/2} \pmod{p}\\ &\equiv a^{(N(p)-1)/2}2^{(N(p)-1)/4}i^{(N(p)-1)/2} \pmod{p}\\ &\equiv 2^{(N(p)-1)/4} \pmod{p}.\end{aligned}$$

Since $p \equiv 5 \pmod{8}$, $(2/p) = -1$ or by Euler's criterion $2^{(p-1)/2} \equiv -1 \pmod{p}$. Therefore, $\langle a\sqrt{2}i/p \rangle \equiv 2^{(N(p)-1)/4} \equiv (-1)^{(p+1)/2} \equiv -1 \pmod{p}$ and the second conjecture is proved.

For conjecture three, we have

COROLLARY 2.3. $\langle \alpha/p \rangle = \langle -\alpha/p \rangle = \langle \bar{\alpha}/p \rangle = \langle -\bar{\alpha}/p \rangle$.

Proof.
$$\begin{aligned}\langle -\alpha/p \rangle &\equiv (-\alpha)^{(N(p)-1)/2} \pmod{p}\\ &\equiv (-1)^{(N(p)-1)/2}\alpha^{(N(p)-1)/2} \pmod{p}\\ &\equiv \langle \alpha/p \rangle \pmod{p}. \text{ Hence, } \langle \alpha/p \rangle = \langle -\alpha/p \rangle.\end{aligned}$$

Similarly, $\langle -\bar{\alpha}/p \rangle = \langle \alpha/p \rangle$. Suppose $\langle \alpha/p \rangle = 1$, then there exists a β such that $\beta^2 \equiv \alpha \pmod{p}$. Therefore, $\bar{\beta}^2 \equiv \bar{\alpha} \pmod{p}$ or $\langle \bar{\alpha}/p \rangle = 1$ and conjecture three is established.

Conjecture four is equivalent to

COROLLARY 2.4. $\langle 1+\sqrt{2}i/p \rangle = \langle a+a\sqrt{2}i/p \rangle$.

Proof.
$$\begin{aligned}\langle a+a\sqrt{2}i/p \rangle &\equiv (a+a\sqrt{2}i)^{(N(p)-1)/2} \pmod{p}\\ &\equiv a^{(N(p)-1)/2}(1+\sqrt{2}i)^{(N(p)-1)/2} \pmod{p}\\ &\equiv (1+\sqrt{2}i)^{(N(p)-1)/2} \pmod{p}\\ &\equiv \langle 1+\sqrt{2}i/p \rangle \pmod{p}.\end{aligned}$$

Hence, $\langle 1+\sqrt{2}i/p \rangle = \langle a+a\sqrt{2}i/p \rangle$.

From Figures 1–3, we see that there is no symmetry about the diagonals. After we establish a relationship between $\langle\ /\ \rangle$ and $(\ /\ )$, we will be able to determine under what conditions $1+\sqrt{2}i$ and hence the elements on the diagonals are quadratic residues or nonresidues. Furthermore, we will investigate under what conditions $a+b\sqrt{2}i$ and $b+a\sqrt{2}i$ are both quadratic residues or nonresidues.

3. The number of residues per line. In this section, we shall prove that conjectures five and six are valid. Because of symmetry about the x and y-axis (Corollary 2.3) and since every value on the x-axis is a quadratic residue while every value on the y-axis is a quadratic nonresidue, conjectures five and six are equivalent to

THEOREM 3.1. *If* $c \neq 0$ *and* $b \neq 0$ *then*

$$\sum_{d=1}^{(p-1)/2} \langle c+d\sqrt{2}i/p\rangle = \sum_{a=1}^{(p-1)/2} \langle a+b\sqrt{2}i/p\rangle = 0.$$

Proof. There are $(p^2-1)/2$ quadratic residues and $p-1$ of them are on the x-axis. There are therefore $(p^2-2p+1)/2$ residues equally distributed in the four quadrants, a total of $(p^2-2p+1)/8$ residues in each quadrant. Similarly there are $(p^2-2p+1)/8$ nonresidues in each quadrant. Hence,

$$\sum_{c=1}^{(p-1)/2}\sum_{d=1}^{(p-1)/2} \langle c+d\sqrt{2}i/p\rangle = (p^2-2p+1)/8-(p^2-2p+1)/8 = 0.$$

Since $d \neq 0$ there exists a rational integer x_d such that $dx_d \equiv 1 \pmod{p}$ and $\langle x_d/p\rangle = 1$. Therefore $cx_d+dx_d\sqrt{2}i \equiv cx_d+\sqrt{2}i \pmod{p}$ or

$$\sum_{c=1}^{(p-1)/2} \langle cx_d+dx_d\sqrt{2}i/p\rangle = \sum_{c=1}^{(p-1)/2} \langle c+d\sqrt{2}i/p\rangle = \sum_{c=1}^{(p-1)/2} \langle cx_d+\sqrt{2}i/p\rangle.$$

Now $\hat{p}$ is a complete residue system so cx_d for each value of c is congruent modulo p to some element of $\hat{p}$. Since p is a real prime, cx_d must be congruent to some value on the x-axis. Therefore, the elements of the set

$$\{x_d, 2x_d, \cdots, (p-1)x_d/2\}$$

are in some way congruent to

$$\{\zeta_1, 2\zeta_2, \cdots, \zeta_{p-1}(p-1)/2\} \text{ modulo } p \text{ where } \zeta_j = \pm 1.$$

Because of symmetry with respect to the y-axis (Corollary 2.3), the following equality holds:

$$\sum_{c=1}^{(p-1)/2} \langle c+d\sqrt{2}i/p\rangle = \sum_{c=1}^{(p-1)/2} \langle cx_d+\sqrt{2}i/p\rangle$$
$$= \sum_{j=1}^{(p-1)/2} \langle \zeta_j j+\sqrt{2}i/p\rangle = \sum_{c=1}^{(p-1)/2} \langle c+\sqrt{2}i/p\rangle.$$

Summing the expression over d as d ranges between 1 and $(p-1)/2$ and interchanging the summation signs, we have

$$\sum_{c=1}^{(p-1)/2} \sum_{d=1}^{(p-1)/2} \langle c + d\sqrt{2}i/p \rangle = (p-1)/2 \left(\sum_{c=1}^{(p-1)/2} \langle c + \sqrt{2}i/p \rangle \right) = 0$$

or

$$\sum_{c=1}^{(p-1)/2} \langle c + d\sqrt{2}i/p \rangle = 0.$$

A similar argument will show that

$$\sum_{a=1}^{(p-1)/2} \langle a + b\sqrt{2}i/p \rangle = 0.$$

4. A connection between $\langle / \rangle$ and $(/)$. The relationship which we wish to establish is

THEOREM 4.1. $\langle a + b\sqrt{2}i/p \rangle = (a^2 + 2b^2/p)$.

Proof. Suppose $\langle a + b\sqrt{2}i/p \rangle = 1$ then there is a β such that

$$\beta^2 \equiv (a + b\sqrt{2}i) \pmod{p}.$$

Consequently, $\bar{\beta}^2 \equiv (a - b\sqrt{2}i) \pmod{p}$ or $(\beta\bar{\beta})^2 \equiv a^2 + 2b^2 \pmod{p}$ so that $(a^2 + 2b^2/p) = 1$.

Now assume that $\langle a + b\sqrt{2}i/p \rangle = -1$. Let q be any prime of the form $8np + (3p + 2)$. Since $p = 8k + 5$, $q = 8(np + 3k + 2) + 1$ or $q \equiv 1 \pmod{8}$. Therefore, there exist integers c and d such that $q = c^2 + 2d^2$. But $q \equiv 2 \pmod{p}$ so $(q/p) = (c^2 + 2d^2/p) = (2/p) = -1$. By the first part of this proof this implies that $\langle c + d\sqrt{2}i/p \rangle = -1$. Therefore,

$$\langle a + b\sqrt{2}i/p \rangle \langle c + d\sqrt{2}i/p \rangle = 1$$

or $\langle (ac - 2bd) + (ad + bc)\sqrt{2}i/p \rangle = 1$.

Using the first part of this proof again, we have

$$(ac - 2bd)^2 + 2(ad + bc)^2/p = (a^2 + 2b^2/p)(c^2 + 2d^2/p) = 1.$$

Hence, $(a^2 + 2b^2/p) = -1$ or $\langle a + b\sqrt{2}i/p \rangle = (a^2 + 2b^2/p)$.

An immediate consequence of Theorems 3.1 and 4.1 is

COROLLARY 4.1. *If $b \neq 0$ then $\sum_{a=1}^{(p-1)/2}(a^2 + 2b^2/p) = 0$.*

We also have

COROLLARY 4.2. *If $a \neq 0$ then $\sum_{b=1}^{(p-1)/2}(a^2 + 2b^2/p) = 0$.*

Applying Theorem 4.1, we are able to determine the quadratic nature of the elements on the diagonals. The result, with $a = b = 1$, is

COROLLARY 4.3. $\langle 1 + \sqrt{2}i/p \rangle = 1$ *iff* $(3/p) = 1$.

Since $(3/p) = 1$ iff $p \equiv \pm 1 \pmod{12}$ and $(3/p) = -1$ iff $p \equiv \pm 5 \pmod{12}$, we have

COROLLARY 4.4. *$\langle 1 + \sqrt{2}i/p \rangle = 1$ if $p \equiv 1 \pmod{12}$ and $\langle 1 + \sqrt{2}i/p \rangle = -1$ if $p \equiv 5 \pmod{12}$.*

With regard to symmetry about the diagonals, we have

COROLLARY 4.5. *$\langle a + b\sqrt{2}i/p \rangle = \langle b + a\sqrt{2}i/p \rangle$ iff $(2a^4 + 5a^2b^2 + 2b^4/p) = 1$.*

Proof. $\langle a + b\sqrt{2}i/p \rangle = \langle b + a\sqrt{2}i/p \rangle$ iff $\langle a + b\sqrt{2}i/p \rangle \langle b + a\sqrt{2}i/p \rangle = 1$ iff $\langle -ab + (a^2 + b^2)\sqrt{2}i/p \rangle = 1$ iff $(2a^4 + 5a^2b^2 + 2b^4/p) = 1$.

References

1. N. R. Hardman and J. H. Jordan, The distribution of quadratic residues in fields of order p^2, this MAGAZINE, 42 (1969) 12–17.

2. C. J. Potratz, Character sums in $Z(\sqrt{-2})$ / (p), doctoral dissertation, Washington State University, 1966.

A CHARACTERIZATION OF CONTINUOUS CLOSED REAL FUNCTIONS

M. SOLVEIG ESPELIE, Howard University and
JAMES E. JOSEPH, Federal City College

It is a well known simple exercise to prove that a continuous function from the reals, R, into R is open if and only if it is monotone. In this note we will characterize the set of continuous closed functions from R into R. (We say that a function is closed if the image of any closed set is closed.) In the proof we use only ideas from advanced calculus so that the student should be able to read the paper and prove the corollaries.

DEFINITION. *Let f be a continuous function from R into R. The sequence $\{f(x_n)\}$ is an asymptotic sequence for f if and only if $\{x_n\}$ has no Cauchy subsequence, $\lim_{n\to\infty} f(x_n) = y$ is finite and $y \neq f(x_n)$ for each n.*

THEOREM. *Suppose f is a continuous function from R into R, then f is closed if and only if f does not admit an asymptotic sequence.*

Proof. Clearly, if f admits an asymptotic sequence, $\{f(x_n)\}$, f is not closed. For then the sequence $\{x_n\}$ has no convergent subsequence so that $\{x_n\}$ considered as a set has no limit point. Thus the set $\{x_n\}$ is closed but its image under f is not closed.

Conversely, suppose that f is not closed. We will show that f admits an asymptotic sequence. Let S be a closed set in R such that $f(S)$ is not closed. Then there exists a limit point y of $f(S)$ such that $y \notin f(S)$. Hence there is a sequence of distinct elements of $f(S)$, $\{f(x_n)\}$, such that $\lim_{n\to\infty} f(x_n) = y$ and $f(x_n) \neq y$ for any n. Now

Since $(3/p) = 1$ iff $p \equiv \pm 1 \pmod{12}$ and $(3/p) = -1$ iff $p \equiv \pm 5 \pmod{12}$, we have

COROLLARY 4.4. *$\langle 1 + \sqrt{2}i/p \rangle = 1$ if $p \equiv 1 \pmod{12}$ and $\langle 1 + \sqrt{2}i/p \rangle = -1$ if $p \equiv 5 \pmod{12}$.*

With regard to symmetry about the diagonals, we have

COROLLARY 4.5. *$\langle a + b\sqrt{2}i/p \rangle = \langle b + a\sqrt{2}i/p \rangle$ iff $(2a^4 + 5a^2b^2 + 2b^4/p) = 1$.*

Proof. $\langle a + b\sqrt{2}i/p \rangle = \langle b + a\sqrt{2}i/p \rangle$ iff $\langle a + b\sqrt{2}i/p \rangle \langle b + a\sqrt{2}i/p \rangle = 1$ iff $\langle -ab + (a^2 + b^2)\sqrt{2}i/p \rangle = 1$ iff $(2a^4 + 5a^2b^2 + 2b^4/p) = 1$.

References

1. N. R. Hardman and J. H. Jordan, The distribution of quadratic residues in fields of order p^2, this MAGAZINE, 42 (1969) 12–17.
2. C. J. Potratz, Character sums in $Z(\sqrt{-2})$ / (p), doctoral dissertation, Washington State University, 1966.

A CHARACTERIZATION OF CONTINUOUS CLOSED REAL FUNCTIONS

M. SOLVEIG ESPELIE, Howard University and
JAMES E. JOSEPH, Federal City College

It is a well known simple exercise to prove that a continuous function from the reals, R, into R is open if and only if it is monotone. In this note we will characterize the set of continuous closed functions from R into R. (We say that a function is closed if the image of any closed set is closed.) In the proof we use only ideas from advanced calculus so that the student should be able to read the paper and prove the corollaries.

DEFINITION. *Let f be a continuous function from R into R. The sequence $\{f(x_n)\}$ is an asymptotic sequence for f if and only if $\{x_n\}$ has no Cauchy subsequence, $\lim_{n\to\infty} f(x_n) = y$ is finite and $y \neq f(x_n)$ for each n.*

THEOREM. *Suppose f is a continuous function from R into R, then f is closed if and only if f does not admit an asymptotic sequence.*

Proof. Clearly, if f admits an asymptotic sequence, $\{f(x_n)\}$, f is not closed. For then the sequence $\{x_n\}$ has no convergent subsequence so that $\{x_n\}$ considered as a set has no limit point. Thus the set $\{x_n\}$ is closed but its image under f is not closed.

Conversely, suppose that f is not closed. We will show that f admits an asymptotic sequence. Let S be a closed set in R such that $f(S)$ is not closed. Then there exists a limit point y of $f(S)$ such that $y \notin f(S)$. Hence there is a sequence of distinct elements of $f(S)$, $\{f(x_n)\}$, such that $\lim_{n\to\infty} f(x_n) = y$ and $f(x_n) \neq y$ for any n. Now

$\{x_n\}$ contains no Cauchy subsequence, for if $\{x_{n_m}\}$ were a Cauchy subsequence it would converge to a point x of the closed set S. Then by the continuity of f, $\lim_{m\to\infty} f(x_{n_m}) = f(x) \in f(S)$. Since $f(x_n)$ converges to y, the subsequence $\{f(x_{n_m})\}$ must converge to y so that $f(x) = y$. This is impossible and $\{f(x_n)\}$ is asymptotic.

We note that our result is valid for more general spaces. The domain may be a complete metric space and the range may be a first countable topological space such that converging sequences have unique limits.

It is clear that any continuous function from R into R with a horizontal asymptote is not closed. However, the concept of horizontal asymptotes is not sufficient to characterize continuous closed real functions, for $f(x) = \sin x$ is not closed and does not have a horizontal asymptote. Notice that $\{f(2n\pi + 1/n)\}$ is an asymptotic sequence for this function.

COROLLARY 1. *If f is continuous from R into R, $\lim_{x\to\infty} f(x) = \pm\infty$ and $\lim_{x\to-\infty} f(x) = \pm\infty$, then f is closed.*

COROLLARY 2. *The polynomial and logarithmic functions are closed. The trigonometric and exponential functions are not closed. The nonconstant rational function $f(x) = (a_m x^m + \cdots + a_1 x + a_0)/(b_n x^n + \cdots + b_1 x + b_0)$ is closed if and only if $m > n$.*

ON CONDITIONS IMPLYING CONTINUITY OF REAL-VALUED FUNCTIONS

R. F. DICKMAN, JR., Virginia Polytechnic Institute and State University

There are several known conditions on real-valued functions on metric spaces which imply continuity. For example every connectedness-preserving real-valued function with closed point inverses which is defined on a locally connected metric space is continuous [**2**]. In this paper we introduce a property which is weaker than closedness of a function and which is enjoyed by every continuous function and show that the assumption that certain products have this property yields continuity. We also extend several known results.

DEFINITION. *Let (X,d) be a metric space. We say that a real-valued function f on X is weakly-closed at $x_0 \in X$ provided whenever $\{x_i\}_{i=1}^{\infty}$ is a sequence in X which converges to x_0, $f(\bigcup_{i=0}^{\infty} x_i)$ is a closed set. If f is weakly-closed at each point of X, we say that f is weakly-closed.*

LEMMA 1. *Let (X,d) be a metric space and let $f: X \to$ Reals be a bounded real-valued function with closed point inverses. If f is weakly-closed at x_0, then f is continuous at x_0.*

Proof. Suppose that f is not continuous at x_0. Then since f is bounded, there exists a sequence $\{x_i\}_{i=1}^{\infty}$ in X converging to x_0 such that $\{f(x_i)\}_{i=1}^{\infty}$ converges to a point

$\{x_n\}$ contains no Cauchy subsequence, for if $\{x_{n_m}\}$ were a Cauchy subsequence it would converge to a point x of the closed set S. Then by the continuity of f, $\lim_{m\to\infty} f(x_{n_m}) = f(x) \in f(S)$. Since $f(x_n)$ converges to y, the subsequence $\{f(x_{n_m})\}$ must converge to y so that $f(x) = y$. This is impossible and $\{f(x_n)\}$ is asymptotic.

We note that our result is valid for more general spaces. The domain may be a complete metric space and the range may be a first countable topological space such that converging sequences have unique limits.

It is clear that any continuous function from R into R with a horizontal asymptote is not closed. However, the concept of horizontal asymptotes is not sufficient to characterize continuous closed real functions, for $f(x) = \sin x$ is not closed and does not have a horizontal asymptote. Notice that $\{f(2n\pi + 1/n)\}$ is an asymptotic sequence for this function.

COROLLARY 1. *If f is continuous from R into R, $\lim_{x\to\infty} f(x) = \pm\infty$ and $\lim_{x\to-\infty} f(x) = \pm\infty$, then f is closed.*

COROLLARY 2. *The polynomial and logarithmic functions are closed. The trigonometric and exponential functions are not closed. The nonconstant rational function $f(x) = (a_m x^m + \cdots + a_1 x + a_0)/(b_n x^n + \cdots + b_1 x + b_0)$ is closed if and only if $m > n$.*

ON CONDITIONS IMPLYING CONTINUITY OF REAL-VALUED FUNCTIONS

R. F. DICKMAN, JR., Virginia Polytechnic Institute and State University

There are several known conditions on real-valued functions on metric spaces which imply continuity. For example every connectedness-preserving real-valued function with closed point inverses which is defined on a locally connected metric space is continuous [**2**]. In this paper we introduce a property which is weaker than closedness of a function and which is enjoyed by every continuous function and show that the assumption that certain products have this property yields continuity. We also extend several known results.

DEFINITION. *Let (X,d) be a metric space. We say that a real-valued function f on X is weakly-closed at $x_0 \in X$ provided whenever $\{x_i\}_{i=1}^{\infty}$ is a sequence in X which converges to x_0, $f(\bigcup_{i=0}^{\infty} x_i)$ is a closed set. If f is weakly-closed at each point of X, we say that f is weakly-closed.*

LEMMA 1. *Let (X,d) be a metric space and let $f: X \to$ Reals be a bounded real-valued function with closed point inverses. If f is weakly-closed at x_0, then f is continuous at x_0.*

Proof. Suppose that f is not continuous at x_0. Then since f is bounded, there exists a sequence $\{x_i\}_{i=1}^{\infty}$ in X converging to x_0 such that $\{f(x_i)\}_{i=1}^{\infty}$ converges to a point

y and $y \neq f(x_0)$. For each positive integer n, let $C_n = f(x_0) \cup f(\bigcup_{i=n}^{\infty} x_i)$. Then since f is weakly-closed at x_0, each of the sets C_n is closed and this implies that for infinitely many integers j, $y = f(x_j)$. But this means that x_0 is a limit point of $f^{-1}(y)$ and since $f^{-1}(y)$ is a closed set, $x_0 \in f^{-1}(y)$. Then $f(x_0) = x$ and this contradicts our selection of the x_i's. Hence f is continuous at x_0.

THEOREM 1. *Let (X,d) be a metric space, let $x_0 \in X$ and let $f: X \to$ Reals be a bounded real-valued function such that $f(x) \neq 0$ for all $x \in X$. Then if f and $g(x) = (1 + d(x,x_0))f(x)$ are weakly-closed at x_0, f is continuous at x_0.*

Proof. Suppose that f is not continuous at x_0. Then there exists a sequence $\{x_i\}_{i=1}^{\infty}$ which converges to x_0 such that for all i, $x_i \neq x_0$ and $\{f(x_i)\}_{i=1}^{\infty}$ converges to a number $y \neq f(x_0)$. As in the proof of Lemma 1, we must have that $f(x_i) = y$ for infinitely many positive integers i, so we may and do assume that for every i, $f(x_i) = y$. We now argue that for every nonnegative integer i, $g(x_i) \neq y$. First of all $g(x_0) = (1+0)f(x_0) \neq y$ by our selection of y. Now for $i > 0$, $g(x_i) = (1 + d(x_i, x_0))f(x_i) = (1 + d(x_i,x_0))y \neq y$ since $d(x_i,x_0) > 0$ and $y \neq 0$. Thus y is an element of the closure of $C = g(\bigcup_{i=0}^{\infty} x_i)$ which does not belong to C. This contradicts our assumption that g is weakly-closed at x_0 and hence f is continuous at x_0.

COROLLARY (1.1). *Let (X,d) be a metric space, let $x_0 \in X$ and let $f: X \to$ reals be a bounded real-valued function. Then if f is weakly-closed at x_0 and if for every real number k, $g_k(x) = (1 + d(x,x_0))(f(x) + k)$ is weakly-closed at x_0, f is continuous at x_0.*

Proof. Since f is bounded there exists a real number k such that for all $x \in X$, $f(x) + k \neq 0$. Define $h(x) = f(x) + k$. Then h is f composed with a homeomorphism on the reals and as such, h is weakly-closed at x_0 and by our hypothesis $(1 + d(x,x_0))\, h(x)$ is weakly-closed at x_0. By Theorem 1, h is continuous at x_0 from whence it follows that f is continuous at x_0.

LEMMA 2. *Let X be a subset of the reals, let $x_0 \in X$ and let f be a bounded real-valued function on X such that $f(x) \neq 0$ for all $x \in X$. If $g(x) = (1 + |x - x_0|)f(x)$ is weakly-closed, then every point inverse of f is a closed set.*

Proof. Let $y \in f(x)$ and suppose that $f^{-1}(y)$ is not closed. Then there exists a sequence $\{z_i\}_{i=1}^{\infty}$ converging to a point $z_0 \in X$ such that for all $i > 0$, $f(z_i) = y$ and $f(z_0) \neq y$. Since X is a subset of the reals we may and do suppose that the z_i's have been chosen so that for $i > 0$, $|x_0 - z_i| \neq |x_0 - z_0|$. Now the sequence $\{g(z_i) = (1 + |x_0 - z_i|)y\}_{i=1}^{\infty}$ converges to $a = (1 + |x_0 - z_0|)y$ and since g is weakly-closed we must have that for some z_i, $g(z_i) = a$. Clearly $g(z_0) \neq a$, since $g(z_0) = (1 + |x_0 - z_0|)f(z_0)$ and $f(z_0) \neq y$. Hence for some $i > 0$, $(1 + |x_0 - z_1|)y = (1 + |x_0 - z_0|)y$. But since $y \neq 0$ this implies that $|x_0 - z_0| = |x_0 - z_i|$ for some $i > 0$ and this contradicts our selection of the z_i's. Hence $f^{-1}(y)$ is closed.

THEOREM 2. *Let X be a subset of the reals, let $x_0 \in X$ and let f be a bounded real-valued function on X such that for all $x \in X$, $f(x) \neq 0$ and $g(x) = (1 + |x_0 - x|)f(x)$ is weakly-closed. Then if f is weakly-closed at $x_1 \in X$, f is continuous at x_1.*

Proof. This result follows immediately from Lemmas 1 and 2.

COROLLARY (2.1). *Let X be a subset of the reals, let $x_0 \in X$ and let f be a bounded real-valued function on X such that for every real number k, $g_k(x) = (1 + |x - x_0|)(f(x) + k)$ is weakly-closed. Then if f is weakly-closed at $x_1 \in X$, f is continuous at x_1.*

Example 1. This is an example of a closed function on a metric space where for some x_0, $g_k(x) = (1 + d(x, x_0))(f(x) + k)$ is weakly-closed for any real number k, however f is not continuous at some point $x_1 \in X$. Thus the hypothesis X is a subset of the reals is necessary in Theorem 2. Let $Y = \{(x, y) : x^2 + y^2 = 1\}$ be the unit circle in the plane, let $Z = \{(x, y) : x = 0 \text{ and } 0 \leqq y \leqq 1\}$ and let $X = Y \cup Z$. Define f on X by $f(x, y) = 2$ for $(x, y) \neq (0, 1)$ and let $f(0, 1) = 1$. Choose $x_0 = (0, 0)$. Then f is closed and every $g_k = (1 + d(x, x_0))_{(f(x)+k)}$ is weakly-closed. Clearly f is not continuous at $(0, 1)$.

Example 2. This is an example of a bounded real-valued function defined on $[0, 1]$ where $f(x) \neq 0$ for all $x \in [0, 1]$, point inverses of f are closed, but $g(x) = (1 + x) f(x)$ is not weakly-closed at $x_0 = 0$. Hence the condition $g(x)$ is weakly-closed is not equivalent to f having closed point inverses. Define G on $[0, 1]$ by $G(x) = -x$ for $x \neq 0$ and $G(0) = -1$ and let $f(x) = G(x)/(1 + x)$. Then $g(x) = G(x)$ is not weakly closed at $x_0 = 0$, but point inverses of f are finite sets.

Example 3. This is an example of a function f on $[0, 1]$ where $f(x) \neq 0$ for all x and $g(x) = (1 + x) f(x)$ is a closed function and f is not weakly closed at $x_0 = 0$. Let G be defined by $G(x) = 2$ if $x \neq 0$ and $G(0) = 1$ and let $f(x) = G(x)/(1 + x)$. Then f is not weakly-closed at 0 but $g(x) = (1 + x) f(x) = G(x)$ is a closed function.

DEFINITION. *Let X be a connected subset of the reals. A real-valued function f on X is called a Darboux function provided for every connected set C in X, $f(C)$ is a connected set. It is easy to show that every Darboux function with closed point inverses is continuous* [**1**].

THEOREM 3. *Let f be a Darboux function defined on a connected subset of the reals. If f is weakly closed, f is continuous.*

Proof. It will be sufficient to prove that point inverses of f are closed. To this end let $y \in f(X)$ and suppose that $f^{-1}(y)$ is not closed. Then there exists a sequence $\{x_i\}_{i=1}^{\infty}$ in X converging to a point x_0 such that $f(x_0) \neq y$ for $i > 0, f(x_i) = y$. Now let $\{y_i\}$ be any sequence in the open interval between $f(x_0)$ and y which converges to y. For each $i > 0$ let A_i denote the closed interval between x_0 and x_i. Then for $i > 0$, $f(A_i)$ is a connected set which contains y and $f(x_0)$ and so there must be a point z_i in A_i such that $f(z_i) = y_i$. Then the sequence $\{z_i\}_{i=1}^{\infty}$ converges to x_0 and by the weak-closedness of f at x_0, $C = f(\bigcup_{i=1}^{\infty} z_i) \cup f(x_0)$ is a closed set. But this is impossible since y is a limit point of C which does not belong to C. Hence point inverses of f are closed and by [**1**], f is continuous.

References

1. A. M. Bruckner and J. B. Bruckner, Darboux transformations, Trans. Amer. Math. Soc., 128 (1967) 103–111.
2. G. T. Whyburn, Continuity of multifunctions, Proc. Nat. Acad. Sci., 54 (1965) 1494–1501.

A DIRICHLET PROBLEM

CLARENCE R. EDSTROM, Air Force Institute of Technology

We consider the Dirichlet problem

$$u_{xx} + u_{yy} = 0 \tag{1}$$

with

$$u(x,0) = u(0,y) = u(1,y) = 0,\ u(x,1) = x. \tag{2}$$

Using the standard technique of separation of variables as given in Churchill [**1**, page 34], we find the solution is

$$u(x,y) = (2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1} \sinh n\pi y \sin n\pi x}{n \sinh n\pi}. \tag{3}$$

Next we consider the Dirichlet problem

$$U_{xx} + U_{yy} = 0 \tag{4}$$

with

$$U(x,0) = U(0,y) = 0,\ U(x,1) = x,\ U(1,y) = y. \tag{5}$$

The solution of this problem is

$$U(x,y) = (2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}[\sinh n\pi y \sin n\pi x + \sinh n\pi x \sin n\pi y]}{n \sinh n\pi}. \tag{6}$$

Another solution of this problem is

$$U(x,y) = xy. \tag{7}$$

Thus by the uniqueness of the solution of the Dirichlet problem we have

$$(2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}[\sinh n\pi y \sin n\pi x + \sinh n\pi x \sin n\pi y]}{n \sinh n\pi} = xy. \tag{8}$$

Now if in (3) we interchange x and y and then add the result to (3), we obtain

$$u(x,y) + u(y,x) = (2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}[\sinh n\pi y \sin n\pi x + \sinh n\pi x \sin n\pi y]}{n \sinh n\pi}. \tag{9}$$

References

1. A. M. Bruckner and J. B. Bruckner, Darboux transformations, Trans. Amer. Math. Soc., 128 (1967) 103–111.
2. G. T. Whyburn, Continuity of multifunctions, Proc. Nat. Acad. Sci., 54 (1965) 1494–1501.

A DIRICHLET PROBLEM

CLARENCE R. EDSTROM, Air Force Institute of Technology

We consider the Dirichlet problem

$$u_{xx} + u_{yy} = 0 \tag{1}$$

with

$$u(x,0) = u(0,y) = u(1,y) = 0, \ u(x,1) = x. \tag{2}$$

Using the standard technique of separation of variables as given in Churchill [**1**, page 34], we find the solution is

$$u(x,y) = (2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1} \sinh n\pi y \sin n\pi x}{n \sinh n\pi}. \tag{3}$$

Next we consider the Dirichlet problem

$$U_{xx} + U_{yy} = 0 \tag{4}$$

with

$$U(x,0) = U(0,y) = 0, \ U(x,1) = x, \ U(1,y) = y. \tag{5}$$

The solution of this problem is

$$U(x,y) = (2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}[\sinh n\pi y \sin n\pi x + \sinh n\pi x \sin n\pi y]}{n \sinh n\pi}. \tag{6}$$

Another solution of this problem is

$$U(x,y) = xy. \tag{7}$$

Thus by the uniqueness of the solution of the Dirichlet problem we have

$$(2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}[\sinh n\pi y \sin n\pi x + \sinh n\pi x \sin n\pi y]}{n \sinh n\pi} = xy. \tag{8}$$

Now if in (3) we interchange x and y and then add the result to (3), we obtain

$$u(x,y) + u(y,x) = (2/\pi) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}[\sinh n\pi y \sin n\pi x + \sinh n\pi x \sin n\pi y]}{n \sinh n\pi}. \tag{9}$$

Substituting (8) into (9) yields

$$u(x,y) + u(y,x) = xy. \tag{10}$$

Next if we set $y = x$ in (10), we obtain

$$u(x,x) = \frac{x^2}{2}. \tag{11}$$

Hence combining (3) with $y = x$ and (11) produces

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1} \sinh n\pi x \sin n\pi x}{n \sinh n\pi} = \frac{\pi x^2}{4}, \quad |x| < 1. \tag{12}$$

Reference

1. R. V. Churchill, Fourier Series and Boundary Value Problems, 2nd ed., McGraw-Hill, New York, 1963.

THE ELLIPSE IN EIGHTEENTH CENTURY SUNDIAL DESIGN

W. W. DOLAN, Linfield College, McMinnville, Oregon

1. Introduction. Sundials served a vital purpose in human affairs up until about the end of the 18th century. Not until then were clocks and watches reliable enough to be depended upon without frequent correction by the sundial on the town hall or village church. Watches, moreover, were too expensive to be available to most people, so that the public sundial was often the only timekeeping device present. This accounts for the vast number of sundials on walls, towers and pedestals in Europe; Zinner [**1**] has recorded more than 6000 of them extant in 3000 communities. They were much less common in America because clocks were beginning to appear more frequently in the colonial period.

Because the designing of a good sundial was a practical matter of considerable importance, the art was taught in schools and described in hundreds of textbooks, manuals and treatises in the three centuries preceding 1800. To understand the process properly required, and still requires, a good understanding of geometry, spherical trigonometry and fundamentals of astronomy. The purpose of the present comment is to explore one group of 18th century methods of design utilizing the geometry of the ellipse. For simplicity, attention will be restricted to horizontal dials, commonly mounted on pedestals, though the relationships are readily adapted to dials on vertical walls or even on obliquely inclined surfaces.

The introduction of clocks had already led to the universal use of 24 equal hours in the day, as distinguished from the "seasonal" hours of ancient times [**2**] which divided the variable period between sunrise and sunset into twelve equal parts, an arrangement clearly unsuited to mechanical clocks. Equal hours, in turn, are con-

Substituting (8) into (9) yields

$$u(x,y)+u(y,x) = xy. \tag{10}$$

Next if we set $y = x$ in (10), we obtain

$$u(x,x) = \frac{x^2}{2}. \tag{11}$$

Hence combining (3) with $y = x$ and (11) produces

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}\sinh n\pi x \sin n\pi x}{n \sinh n\pi} = \frac{\pi x^2}{4}, \quad |x| < 1. \tag{12}$$

Reference

1. R. V. Churchill, Fourier Series and Boundary Value Problems, 2nd ed., McGraw-Hill, New York, 1963.

THE ELLIPSE IN EIGHTEENTH CENTURY SUNDIAL DESIGN

W. W. DOLAN, Linfield College, McMinnville, Oregon

1. Introduction. Sundials served a vital purpose in human affairs up until about the end of the 18th century. Not until then were clocks and watches reliable enough to be depended upon without frequent correction by the sundial on the town hall or village church. Watches, moreover, were too expensive to be available to most people, so that the public sundial was often the only timekeeping device present. This accounts for the vast number of sundials on walls, towers and pedestals in Europe; Zinner [**1**] has recorded more than 6000 of them extant in 3000 communities. They were much less common in America because clocks were beginning to appear more frequently in the colonial period.

Because the designing of a good sundial was a practical matter of considerable importance, the art was taught in schools and described in hundreds of textbooks, manuals and treatises in the three centuries preceding 1800. To understand the process properly required, and still requires, a good understanding of geometry, spherical trigonometry and fundamentals of astronomy. The purpose of the present comment is to explore one group of 18th century methods of design utilizing the geometry of the ellipse. For simplicity, attention will be restricted to horizontal dials, commonly mounted on pedestals, though the relationships are readily adapted to dials on vertical walls or even on obliquely inclined surfaces.

The introduction of clocks had already led to the universal use of 24 equal hours in the day, as distinguished from the "seasonal" hours of ancient times [**2**] which divided the variable period between sunrise and sunset into twelve equal parts, an arrangement clearly unsuited to mechanical clocks. Equal hours, in turn, are conveniently calibrated on a sundial only when the gnomon, or shadow caster, is parallel to the earth's axis. The gnomon may be a rod, or the structurally more stable thin plate with inclined edge; in either case the edge casting the shadow (often called the *style*) must lie in the plane of the local meridian and be inclined to the horizontal plane at an angle equal to the latitude (Figure 1). With this arrangement the shadow line depends only on the hour angle of the sun, independent of its changing declination

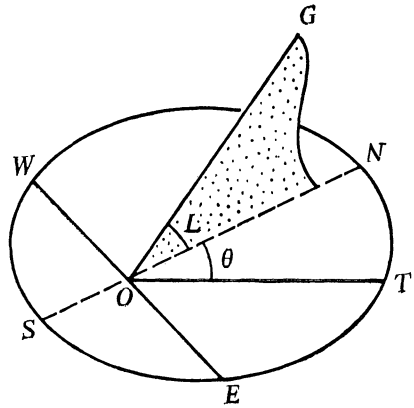


FIG. 1. Schematic view of horizontal sundial. *NS*, local north-south line or noon line; *EW*, east-west or 6-o'clock line; *OG*, style or edge of gnomon; *L*, local latitude; *OT*, shadow of *OG*.

as the seasons progress. We ignore here the small annual cyclic variation due to the "equation of time."

2. The design problem. The task of the designer consists of determining the angles formed on the face of the dial between the successive hour lines and the noon line at the base of the gnomon. It depends on the solution of a simple problem in spherical trigonometry, a technique rarely studied today. The result is the following formula, where θ is the desired angle between noon line and hour line, H is the hour angle of the sun, and L is the local latitude:

$$\tan\theta = \sin L \tan H. \tag{1}$$

The writers of the old treatises liked to work out geometric constructions for the solution of this formula on the dial plane. In presenting these they rarely offered proofs, and often seemed more concerned with ingenious original devices than with simplicity.

Typical of this approach is a method proposed in 1720 by M. Ozanam [3], though it is difficult to say whether it was original with him. The instructions (adapted slightly for this presentation) begin with drawing the semicircle (Figure 2) with indicated horizontal and vertical radii. Divide one quadrant into 15° arcs, representing the successive hour angles, at points K, L, M, N, P; this is presumably to be done by classical ruler-and-compass construction, not with a protractor. Draw perpendiculars from these points to the vertical radius OA at points C, D, E, F, G. Construct angle

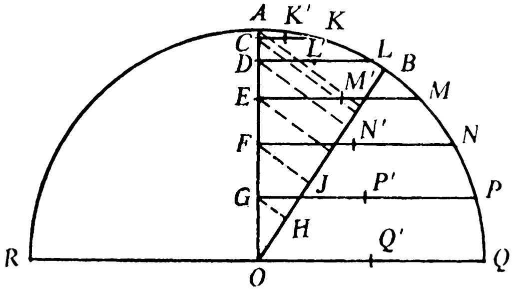


FIG. 2. Layout of sundial face as described in text (after Ozanam).

AOB equal to the latitude, and perpendiculars such as GH from the division points on the vertical radius to OB. On line CK, lay off CK' equal to GH, similarly $DL' = FJ$, etc. Finally, radii from O to K', L', M', N', P' are the required hour lines, and their reflections in the other quadrant complete the 12-hour span from 6 a.m. at R to 6 p.m. at Q. Hour lines before 6 a.m. or after 6 p.m. are extensions of those already drawn; for example, the 5 a.m. line is an extension of that for 5 p.m. The gnomon may be visualized as the sector OAB revolved 90° about OA when OA is oriented from south to north.

To correlate this construction with (1) is a straightforward exercise in plane trigonometry, taken for granted by Ozanam in his book. The present point of interest however, is the locus outlined by points K', L', M', N', P', Q'. These catch the eye as strongly suggesting an ellipse, though the author said nothing about it. The reader may wish to check this out in any of several ways before continuing.

In 1784 James Ferguson [4] published a method (perhaps not for the first time, since his work went through several editions) reminiscent of the one above. Though differing in parts of the procedure, it developed the same points K', L', etc., to determine the hour lines. It is hardly remarkable that Ferguson did not refer to Ozanam or anyone else as a source, for such credits rarely appear in these writings. It does seem remarkable that he did not say anything about his own 1767 publication [5] in which he had pointedly used the same ellipse as a basic part of his design. To compare the two, I have abbreviated Ferguson's procedure and simplified the drawing without, I hope, violating any essential feature. This method of drawing an ellipse occasionally appears in modern texts.

Draw the right triangle PQR (Figure 3) containing the latitude as angle P. Draw concentric semicircles with radii equal to QR and PQ, and the indicated horizontal and vertical diameters. Divide each semicircle into 15° arcs; draw horizontal lines through the points of division on the outer circle, and vertical lines through those of the inner circle. The intersections of corresponding pairs of lines in this grid are points D, E, F, etc. Ferguson concluded with these words: "[These] points are in the elliptical curve, and it is to be drawn through them, by hand... And right lines drawn from the center A through these points in the ellipsis, will be the true hour lines for a horizontal dial."

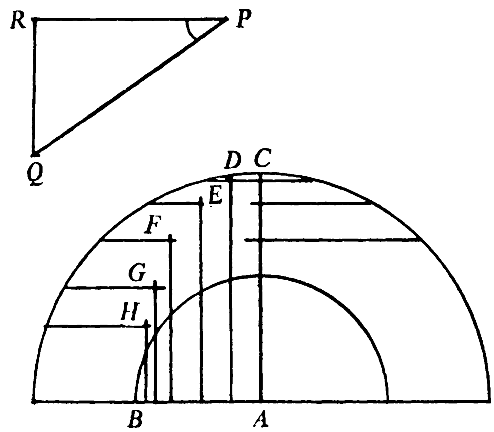


FIG. 3. Layout of sundial face as described in text (after Ferguson).

3. Comment. Ferguson said nothing as to why one should draw the whole ellipse, since only the selected points are significant for the hour lines. One guess is that he was intrigued by the presence of this curve and wanted to make the most of it. Again, he gave no hint of proof that the hour lines are correct in this construction.

From our point of view we see the relations nicely through the presence of the two auxiliary circles, which suggest the familiar parametric equations of the ellipse. If the diameters illustrated are taken as x and y axis and the hour angle H (measured as before from the vertical y-axis) as parameter, the equations of the ellipse are $x = AB \sin H$, $y = AC \cos H$. It follows immediately that any one of the angles θ is determined by the relation

$$\tan\theta = \frac{x}{y} = \frac{AB\sin H}{AC\cos H} = \sin L \tan H$$

as in (1). Thus the proof turns out to be remarkably simple for a relatively involved construction.

It is also easy to see that the eccentricity of the ellipse is cos L, a circumstance that ties in with a final unifying observation. One of the simplest types of sundial is the equatorial, which consists of a circle with radii 15° apart to mark the hours, and the gnomon a peg perpendicular to the plane of the circle at its center. This dial, although it has other disadvantages, has the virtue of being "universal," i.e., useful at any point of the earth, provided it is locally oriented so that the gnomon is parallel to the earth's axis. This requires the dial plane to be inclined to the horizontal at an angle equal to the colatitude (Figure 4).

If now the entire circle of the equatorial dial, together with its hourly radii, is projected parallel to the earth's axis on the horizontal plane, the projection is clearly an ellipse. The major axis is $DD'' = DD' \csc L$, and the minor axis, in the east-west line, is equal to the circle's diameter DD'. This is exactly the ellipse of Ferguson's construction.

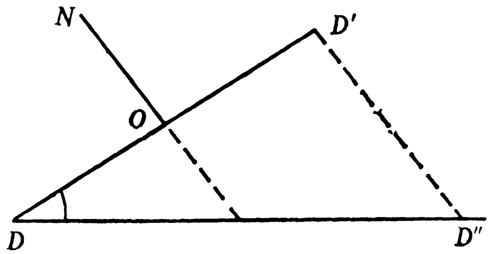


FIG. 4. Polar projection of equatorial sundial on horizontal plane. DD', diameter of circular dial, viewed from west edge; ON, direction of earth's axis; DD'', major axis of horizontal ellipse.

Ferguson did not mention anything of this sort. It seems that he must have had it in mind, however, since he included a paragraph on how to make a south inclining dial by an appropriate change in one axis of the ellipse. An inclining dial is one that "leans forward" toward the south from a south-facing vertical position, and its geometry involves considerations like those of Figure 4.

If one begins with the equatorial dial and works backward through the projection of Figure 4 and the parametric equations of the ellipse, he may establish the fundamental equation (1) without any spherical trigonometry at all, perhaps a crowning indignity for that once-popular branch of applied mathematics.

References

1. E. Zinner, Alte Sonnenuhren an Europäischen Gebäude, Wiesbaden, 1964.
2. W. W. Dolan, Early sundials and the discovery of the conic sections, this MAGAZINE, 45 (1972) 8–12.
3. M. Ozanam, Traité de Gnomonique, Paris, 1720, p. 72.
4. J. Ferguson, Lectures on Select Subjects, 6th ed., London, 1784, p. 330.
5. ———, A new method of constructing sundials, Philos. Trans., 57 (1767) 389.

A CHARACTERIZATION OF 0-SEQUENCES

J. L. BROWN, JR., Pennsylvania State University

I. Introduction. Let $\{f_i\}_1^\infty$ be a nondecreasing sequence of positive integers; we say $\{f_i\}$ is *complete* ([**1**], [**2**]) if every positive integer can be written as a sum of distinct terms from the sequence $\{f_i\}$. In [**3**], Brown and Weiss define a nondecreasing sequence of positive integers $\{f_i\}_1^\infty$ to be an *n-sequence* ($n \geqq 0$) if $\{f_i\}$ possesses the following properties:

(i) $\{f_i\}$ is complete.

(ii) If n arbitrary terms are removed from $\{f_i\}$, the sequence of remaining terms is complete (independently of which n terms are removed).

(iii) If $n+1$ arbitrary terms are removed from $\{f_i\}$, the sequence of remaining terms is incomplete (that is, there exists at least one positive integer which cannot be represented as a sum of distinct terms from the deleted sequence).

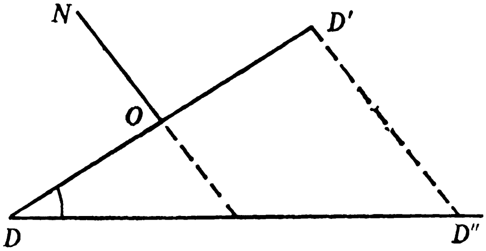


FIG. 4. Polar projection of equatorial sundial on horizontal plane. DD', diameter of circular dial, viewed from west edge; ON, direction of earth's axis; DD'', major axis of horizontal ellipse.

Ferguson did not mention anything of this sort. It seems that he must have had it in mind, however, since he included a paragraph on how to make a south inclining dial by an appropriate change in one axis of the ellipse. An inclining dial is one that "leans forward" toward the south from a south-facing vertical position, and its geometry involves considerations like those of Figure 4.

If one begins with the equatorial dial and works backward through the projection of Figure 4 and the parametric equations of the ellipse, he may establish the fundamental equation (1) without any spherical trigonometry at all, perhaps a crowning indignity for that once-popular branch of applied mathematics.

References

1. E. Zinner, Alte Sonnenuhren an Europäischen Gebäude, Wiesbaden, 1964.
2. W. W. Dolan, Early sundials and the discovery of the conic sections, this MAGAZINE, 45 (1972) 8–12.
3. M. Ozanam, Traité de Gnomonique, Paris, 1720, p. 72.
4. J. Ferguson, Lectures on Select Subjects, 6th ed., London, 1784, p. 330.
5. ———, A new method of constructing sundials, Philos. Trans., 57 (1767) 389.

A CHARACTERIZATION OF 0-SEQUENCES

J. L. BROWN, JR., Pennsylvania State University

I. Introduction. Let $\{f_i\}_1^\infty$ be a nondecreasing sequence of positive integers; we say $\{f_i\}$ is *complete* ([**1**], [**2**]) if every positive integer can be written as a sum of distinct terms from the sequence $\{f_i\}$. In [**3**], Brown and Weiss define a nondecreasing sequence of positive integers $\{f_i\}_1^\infty$ to be an *n-sequence* ($n \geqq 0$) if $\{f_i\}$ possesses the following properties:

(i) $\{f_i\}$ is complete.

(ii) If n arbitrary terms are removed from $\{f_i\}$, the sequence of remaining terms is complete (independently of which n terms are removed).

(iii) If $n+1$ arbitrary terms are removed from $\{f_i\}$, the sequence of remaining terms is incomplete (that is, there exists at least one positive integer which cannot be represented as a sum of distinct terms from the deleted sequence).

Brown and Weiss [3] then give a composite necessary and sufficient condition for $\{f_i\}$ to be a 1-sequence and go on to prove that there exist no n-sequences for $n \geqq 2$. The problem of characterizing 0-sequences, sequences which are complete but which are rendered incomplete by the removal of an arbitrary term, was left open. It is the purpose of the present paper to give a simple complete characterization of the class of all 0-sequences.

2. Derivation of conditions. We quote first a lemma ([3], 558) which will be of repeated use:

LEMMA 1. *Let* $\{f_i\}_1^\infty$ *be a nondecreasing sequence of positive integers with* $f_1 = 1$. *Then* $\{f_i\}$ *is complete if and only if*

(1) $$f_{p+1} \leqq 1 + \sum_1^p f_i \text{ for all } p \geqq 1.$$

We assume henceforth that the sequence $\{f_i\}_1^\infty$ is always a nondecreasing sequence of positive integers. Then if $\{f_i\}$ is a 0-sequence, $\{f_i\}$ must be complete by definition; hence $f_1 = 1$ and condition (1) is clearly a necessary condition. The following lemma establishes another necessary condition:

LEMMA 2. *If* $\{f_i\}$ *is a* 0-*sequence, then*

(2) $$f_{p+1} > 1 + \sum_1^{p-1} f_i \text{ for infinitely many values of } p \geqq 1.$$

(Interpret $\sum_m^n$ as zero when $n < m$).

Proof. If the lemma is assumed false, then $\exists\, N \geqq 10$ such that

(3) $$f_{p+1} \leqq 1 + \sum_1^{p-1} f_i \quad \text{for all } p \geqq N.$$

Let $\{f_i\}'_N$ denote the sequence $\{f_i\}$ after the single term f_N has been deleted. Then $f_{p+1} \leqq 1 + \sum_1^p f_i$ for $1 \leqq p \leqq N-2$ since $\{f_i\}$ is complete. Moreover, by (3), for any $k \geqq 1$, we have

$$f_{N+k} \leqq 1 + \sum_1^{N+k-2} f_i = 1 + \sum_1^{N-1} f_i + \sum_N^{N+k-2} f_i \leqq 1 + \sum_1^{N-1} f_i + \sum_{N+1}^{N+k-1} f_i.$$

These inequalities show that each term of $\{f_i\}'_N$ is less than or equal to one plus the sum of all preceding terms; that is, $\{f_i\}'_N$ is complete by Lemma 1, thus contradicting our hypothesis that $\{f_i\}$ is a 0-sequence. Therefore, condition (2) is necessary.

The next lemma gives a sufficient condition for any complete sequence $\{f_i\}$ to be a 0-sequence.

LEMMA 3. *If* $\{f_i\}$ *is complete and*

(4) $$f_{p+1} > 1 + \sum_1^{p-1} f_i \quad \text{for all } p \geqq 1,$$

then $\{f_i\}$ *is a* 0-*sequence.*

Proof. It remains to show that $\{f_i\}$ is rendered incomplete by removal of an arbitrary term. Clearly $f_1 = 1$ and condition (4) for $p = 1$ along with completeness of $\{f_i\}$ implies $f_2 = 2$. Now, assume f_k with $k \geqq 1$ is deleted; then the deleted sequence $\{f_i\}'_k$ is incomplete since by (4),

$$f_{k+1} > f_{k+1} - 1 > \sum_1^{k-1} f_i,$$

which in turn implies that the integer $f_{k+1} - 1$ can have no representation as a sum of distinct terms from $\{f_i\}'_k$. Since k was arbitrary, we conclude $\{f_i\}$ is a 0-sequence.

To see condition (2) is not sufficient for a complete sequence to be a 0–sequence, we note that the sequence $1,1,1,3,3,3,11,14,25,39,54,\cdots$, is complete and satisfies (2) for *all* $p \geqq 6$, but it is easily seen from Lemma 1 that we may delete a 1 from the sequence without destroying completeness. On the other hand, condition (4) is not necessary for a complete sequence to be a 0-sequence. For the sequence

$$\underbrace{1, 1, 1, \cdots, 1,}_{n \text{ terms } (n > 1)} \; n+1,\, n+2,\, 2n+3,\, 3n+5,\, 5n+8, \cdots$$

does not satisfy (4) for $p \leqq n - 1$; yet it is complete and deletion of any single term renders it incomplete (hence a 0-sequence).

Thus Lemmas 2 and 3, which are often easy to apply, do not afford the total picture—for a complete sequence $\{f_i\}$, the condition, $f_{p+1} > 1 + \Sigma_1^p f_i$, must hold for infinitely many $p \geqq 1$ if $\{f_i\}$ is to be a 0-sequence, while if the condition holds for all p, then $\{f_i\}$ is a 0-sequence. A necessary and sufficient condition lies somewhere between these two alternatives and is given by the following characterization:

THEOREM. $\{f_i\}_1^\infty$ *(nondecreasing positive integers) is a* 0*-sequence if and only if the following two conditions are satisfied*:

(i) $f_{p+1} \leqq 1 + \Sigma_1^p f_i$ *for all* $p \geqq 1$.

(ii) *For each* $n \geqq 1$, $\exists$ *at least one* $k = k(n)$ *such that* $k \geqq n$ *and*

$$f_{k+1} + f_n > 1 + \sum_1^k f_i. \tag{5}$$

Proof. Since a 0-sequence is required to be complete, $f_1 = 1$, and Lemma 1 shows condition (i) of the theorem to be both necessary and sufficient for the 0-sequence to be complete. Now, let $n \geqq 1$ be given. If (ii) is satisfied then $\exists\, k = k(n)$ such that

$$f_{k+1} > f_{k+1} - 1 > \left(\sum_1^k f_i \right) - f_n$$

and it is clear that the integer $f_{k+1} - 1$ can have no admissible representation in terms of the deleted sequence $\{f_i\}'_n$. Conversely, assume $\{f_i\}$ is a 0-sequence and consider the deleted sequence $\{f_i\}'_n$ for fixed $n \geqq 1$. This latter sequence must be incomplete; consequently, by Lemma 1, $\exists$ at least one integer $k = k(n)$ with $k \geqq n$ such that

$$f_{k+1} > 1 + \sum_1^k f_i. \tag{6}$$

Adding f_n to both sides of (6) yields

$$f_{k+1} + f_n > 1 + \sum_1^k f_i,$$

so that condition (ii) is fulfilled.

3. Remarks and examples. We note that condition (ii) of the theorem does imply the necessary condition of Lemma 2. For, with each given $n \geqq 1$, $\exists\, k = k(n)$ such that $k \geqq n$ and $f_{k+1} + f_n > 1 + \Sigma_1^k f_i = 1 + \Sigma_1^{k-1} f_i + (f_k - f_n) + f_n$. Thus $f_{k+1} > 1 + \Sigma_1^{k-1} f_i + (f_k - f_n) \geqq 1 + \Sigma_1^{k-1} f_i$. Since $k = k(n)$ is required to be not less than n, it is clear that (2) of Lemma 2 is satisfied.

Example 1. The Fibonacci sequence defined by $U_1 = 1$, $U_2 = 2$, and $U_{n+2} = U_{n+1} + U_n$ for $n \geqq 1$ also satisfies the obvious recurrence,

$$U_{n+1} = 2 + \sum_1^{n-1} U_i, \quad \text{for } n \geqq 1. \tag{7}$$

This sequence is complete by Lemma 1, since $U_1 = 1$, and from (7), noting $U_n \geqq 1$ for $n \geqq 1$, we have $U_{n+1} \leqq 1 + \Sigma_1^n U_i$. Also from (7),

$$U_{n+1} > 1 + \sum_1^{n-1} U_i \quad \text{for all } n \geqq 1,$$

so that Lemma 3 shows $\{U_i\}$ is a 0-sequence.

Example 2. The most useful 0-sequence is the sequence of powers of 2 which leads to binary representations of the integers. Define

$$V_{n+1} = 2^n \text{ for } n \geqq 0. \text{ Then } V_1 = 1,\ V_2 = 2$$

and

$$V_{n+1} = 1 + \sum_1^n V_i \quad \text{for } n \geqq 0. \tag{8}$$

Lemma 1 again insures completeness and since (8) implies

$$V_{n+1} > 1 + \sum_1^{n-1} V_i \quad \text{for } n \geqq 1,$$

the fact that $\{V_i\}$ is a 0-sequence follows from Lemma 3. Other examples of 0-sequences which do not satisfy the sufficient condition of Lemma 3 are easily constructed using the criterion of the theorem.

References

1. V. E. Hoggatt and C. King, Problem E 1424, Amer. Math. Monthly, 67 (1960) 593.
2. J. L. Brown, Jr., Note on complete sequences of integers, Amer. Math. Monthly, 68 (1961) 557–560.
3. T. C. Brown and M. L. Weiss, On N-sequences, this MAGAZINE, 44 (1971) 89–92.

THE SOLUTION OF A SIMPLE GAME

DANIEL I. A. COHEN, Harvard University

The game of Lose Tic-Tac-Toe is played on the same three-by-three grid as the regular version. The moves are made in the usual fashion with X going first. The first player to score three in a row, column, or diagonal loses. In the following we determine the optimal strategy and prove that it is unique.

For convenience we consider the board to be the matrix $\|(i,j)\|$, $(i,j = 1,2,3)$. The terms center, corner, and side designate the usual squares. Every noncenter square (a,b) has as "antipode" the square $(4-a, 4-b)$. We will adopt the chess format for annotating moves.

THEOREM 1. *X can guarantee at least a draw by playing his first move in the center and subsequently playing in the antipode of each of O's successive moves.*

Proof. Following this strategy obviously prevents X from losing through the center; before X will play three along a given side, O must place three along its antipodal side.

We shall now show that this simple strategy is the best that X can hope for, and in fact, that any deviation from it can be converted by O into a win.

LEMMA. *There is no drawn final position with an X in the center and two antipodal O's.*

Proof. The two antipodal O's break up the perimeter squares into two sets of three. Since there will be four X's on the perimeter in the final position, they will either be split 2–2 or 3–1. In either case, it is easily seen that two of the X's will be antipodal and X will lose through the center.

THEOREM 2. *If X plays his first move in the center, but does not play antipodal to O's first two moves, he can be forced to lose.*

Proof. Case 1. O moves in $(1,1)$, but X does not move into $(3,3)$. O can now win by playing there himself. For no matter where X goes on his second and third moves, O can still occupy a noncorner square on his third move. Then there is only one losing square for O on his next move; but there will be two to choose from. By the above lemma, if O is not forced to complete a line himself, X must.

References

1. V. E. Hoggatt and C. King, Problem E 1424, Amer. Math. Monthly, 67 (1960) 593.
2. J. L. Brown, Jr., Note on complete sequences of integers, Amer. Math. Monthly, 68 (1961) 557–560.
3. T. C. Brown and M. L. Weiss, On N-sequences, this MAGAZINE, 44 (1971) 89–92.

THE SOLUTION OF A SIMPLE GAME

DANIEL I. A. COHEN, Harvard University

The game of Lose Tic-Tac-Toe is played on the same three-by-three grid as the regular version. The moves are made in the usual fashion with X going first. The first player to score three in a row, column, or diagonal loses. In the following we determine the optimal strategy and prove that it is unique.

For convenience we consider the board to be the matrix $\| (i,j) \|$, $(i,j = 1,2,3)$. The terms center, corner, and side designate the usual squares. Every noncenter square (a,b) has as "antipode" the square $(4-a, 4-b)$. We will adopt the chess format for annotating moves.

THEOREM 1. *X can guarantee at least a draw by playing his first move in the center and subsequently playing in the antipode of each of O's successive moves.*

Proof. Following this strategy obviously prevents X from losing through the center; before X will play three along a given side, O must place three along its antipodal side.

We shall now show that this simple strategy is the best that X can hope for, and in fact, that any deviation from it can be converted by O into a win.

LEMMA. *There is no drawn final position with an X in the center and two antipodal O's.*

Proof. The two antipodal O's break up the perimeter squares into two sets of three. Since there will be four X's on the perimeter in the final position, they will either be split 2–2 or 3–1. In either case, it is easily seen that two of the X's will be antipodal and X will lose through the center.

THEOREM 2. *If X plays his first move in the center, but does not play antipodal to O's first two moves, he can be forced to lose.*

Proof. Case 1. O moves in $(1,1)$, but X does not move into $(3,3)$. O can now win by playing there himself. For no matter where X goes on his second and third moves, O can still occupy a noncorner square on his third move. Then there is only one losing square for O on his next move; but there will be two to choose from. By the above lemma, if O is not forced to complete a line himself, X must.

Case 2. X moves in (3,3) his second move but does not answer O's move 2. (1,2) with 3.(3,2). O can now win by playing in (3,2) himself. Again with only one losing square and two to choose from, he can still be sure of making a nonlosing fourth move. By the above lemma, if O does not lose then X does. (See Diagram 1.)

O	O	
	X	
	O	X

DIAGRAM 1

We now complete the proof of the uniqueness of the strategy given in Theorem 1 by demonstrating that X's first move must, indeed, be in the center—albeit an unlikely expedient, as it is the best first move in the usual game of Win Tic-Tac-Toe. It would be nice if this proof could be of the form "If X doesn't play his first move in the center then O will, and guarantee himself a win." Unfortunately, this is not the case, for if the game proceeds as follows:

	X	O
1.	(2,1)	(2,2)
2.	(2,3)	

O must lose. He has only two nonisomorphic moves, (1,1) and (1,2) and X's crushing responses are given in Diagram 2.

O		
X	O	X
	X	

	O	
X	O	X
X		

DIAGRAM 2

In either case O loses. Therefore we will prove that O can win by analyzing both of X's nonisomorphic first moves, (1,1) and (2,1).

THEOREM 3. *After X's opening move of* 1.(2,1), O *can force a win by answering* 1. (1,2).

Proof. Case 1. X does not respond 2.(3,2).
Here O can play in (3,2) on his second move and the position after two and a half rounds will look like Diagram 3 with two more X's on the board.

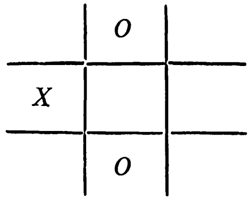


DIAGRAM 3

If either of these X's is in the center, then O will win by our lemma. Therefore, either both X's are along the third column or else another X is placed in the first column. Either way one of these two columns must contain two X's. Without loss of generality, let us assume it is the first column. The third square in that column will guarantee an open square for O, which X himself cannot fill. To win, O need only occupy the one square in the third column which is not in the same row as the open square in the first column. X must now fill the third column. O responds by filling the open square, causing X to play in the center and lose.

Case 2. X plays 2.(3,2). O plays 2. (3,3) and leaves the board as in Diagram 4. Were X to play

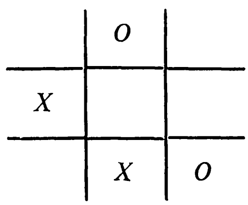


DIAGRAM 4

in (1,1) or (3,1), O would win by playing in (2,2). X and O then split (1,3) and (2,3): X loses along the first column. If X plays 3.(1,3) [or (2,3)], O can win by playing 3. (2,3) [or (1,3)], returning the first column onus to X. Lastly, if X plays 3.(2,2), O can play in his own antipode and win by the lemma.

THEOREM 4. *If X plays* 1.(1,1) *then O can guarantee a win by playing* 1. (2,3).

Proof. Case 1. X does not play 2. (3,2). Here the board looks like Diagram 5 after O plays 2. (3,2), with one more X. The position

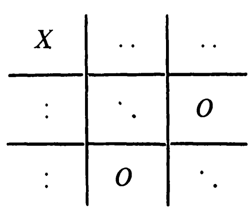


DIAGRAM 5

has three lines along which X can lose and so he can make at most three more moves without having to fill one of these lines. Unfortunately, X must occupy four more squares before he can draw. And so he will lose.

Case 2. X plays 2.(3,2). This case is just a transposition of the game

	X	O
1.	(2,1)	(1,2)
2.	(3,3)	

and we have shown in Theorem 3 that O can win from this beginning.

In summary, we observe that the game heavily favors the second player. This is no surprise since the standard version favors the first player. However, that the correct first move for X is the center is unintuitive. The analysis of this game is more involved than that of the regular version but the result is the same, namely, both games are forced draws. The generalization of the game to higher dimensions seems harder than the still unsolved generalizations of standard Tic-Tac-Toe.

ON FERMAT'S PROBLEM ON THE SURFACE OF A SPHERE

E. J. COCKAYNE, University of Victoria

1. Introduction. Fermat's problem in the plane may be stated as follows: Given 3 distinct points A,B,C in the plane, find the point(s) P which minimizes the sum of distances $PA + PB + PC$. The solution is well known: P is unique. If triangle ABC has an angle $\geqq 2\pi/3$ radians, then P is the vertex of this angle, otherwise P is the point at which the sides of the triangle subtend angles of $2\pi/3$.

This note discusses Fermat's problem on the surface of a sphere. As in the plane, the results have application to the construction of minimal length networks on the sphere (see [1], [2]). The radius is assumed to be unity. The notation UV will be used for both the geodesic (smaller arc of great circle) joining the points U,V and for the length of this geodesic; $X\{ABC\}$ and $\min\{ABC\}$ will denote $XA+XB+XC$ and a point minimizing $XA+XB+XC$ respectively. The special case in which A,B,C are equally spaced on a great circle illustrates that in general $\min\{ABC\}$ is not unique. For brevity, some of the proofs are outlined only and simpler ones are omitted.

2. Two basic theorems. Suppose that A,B,C are not all on the same great circle. Then the closed curve γ, formed by the geodesics AB, BC, CA, divides the surface into two unequal areas. By ΔABC we shall mean the set of points which comprise the smaller area including the boundary γ.

THEOREM 1. *For A, B, C not on a great circle,* $\min\{ABC\}\in\Delta ABC$.

has three lines along which X can lose and so he can make at most three more moves without having to fill one of these lines. Unfortunately, X must occupy four more squares before he can draw. And so he will lose.

Case 2. X plays 2.(3,2). This case is just a transposition of the game

	X	O
1.	(2,1)	(1,2)
2.	(3,3)	

and we have shown in Theorem 3 that O can win from this beginning.

In summary, we observe that the game heavily favors the second player. This is no surprise since the standard version favors the first player. However, that the correct first move for X is the center is unintuitive. The analysis of this game is more involved than that of the regular version but the result is the same, namely, both games are forced draws. The generalization of the game to higher dimensions seems harder than the still unsolved generalizations of standard Tic-Tac-Toe.

ON FERMAT'S PROBLEM ON THE SURFACE OF A SPHERE

E. J. COCKAYNE, University of Victoria

1. Introduction. Fermat's problem in the plane may be stated as follows: Given 3 distinct points A, B, C in the plane, find the point(s) P which minimizes the sum of distances $PA + PB + PC$. The solution is well known: P is unique. If triangle ABC has an angle $\geqq 2\pi/3$ radians, then P is the vertex of this angle, otherwise P is the point at which the sides of the triangle subtend angles of $2\pi/3$.

This note discusses Fermat's problem on the surface of a sphere. As in the plane, the results have application to the construction of minimal length networks on the sphere (see [1], [2]). The radius is assumed to be unity. The notation UV will be used for both the geodesic (smaller arc of great circle) joining the points U, V and for the length of this geodesic; $X\{ABC\}$ and $\min\{ABC\}$ will denote $XA + XB + XC$ and a point minimizing $XA + XB + XC$ respectively. The special case in which A, B, C are equally spaced on a great circle illustrates that in general $\min\{ABC\}$ is not unique. For brevity, some of the proofs are outlined only and simpler ones are omitted.

2. Two basic theorems. Suppose that A, B, C are not all on the same great circle. Then the closed curve γ, formed by the geodesics AB, BC, CA, divides the surface into two unequal areas. By ΔABC we shall mean the set of points which comprise the smaller area including the boundary γ.

THEOREM 1. *For A, B, C not on a great circle,* $\min\{ABC\} \in \Delta ABC$.

The following elegant proof of Theorem 1 is due to M. G. Greening [4]. It replaces a cumbersome proof of the author!

Proof. There are two possibilities for $P = \min\{ABC\}$ exterior to γ.

(i) None of AP, BP, CP intersect γ again. Then the sum of the angles at A, B, C, P is 8π which equals the sum of the angles contained by the four spherical triangles APB, APC, BPC, ABC. This however means that at least one of these triangles has angles sum $\leqq 2\pi$; impossible.

(ii) Suppose that AP intersects BC at D. Then

$$\begin{aligned} D\{ABC\} &= AD + BC \\ &< AD + BP + PC \quad (\text{by } \Delta \text{ inequality}) \\ &< AP + BP + PC = P\{ABC\}. \end{aligned}$$

If the three angles at Q between the geodesics QA, QB, QC are each $2\pi/3$, we say that Q is a $2\pi/3$-point (of A, B, C).

THEOREM 2. (i) *Either* $\min\{ABC\} = A, B$ *or* C
or $\min\{ABC\}$ *is a* $2\pi/3$*-point of* A, B, C.
(ii) *If* $\angle ABC$ *of* ΔABC *is less than* $2\pi/3$, *then* $\min\{ABC\} \neq B$.

Proof. This is a special case of the results in Section 4 of [1]. Solutions of the two types of Theorem 2(i) will be termed vertex minima and $2\pi/3$-point minima respectively. The subsequent sections decide for large classes of triples A, B, C whether they have vertex minima, $2\pi/3$-point minima or both.

3. Solution for equilateral triangles. In this section we use the following formulae for a right-angled spherical triangle with sides a, b, c and angles A, B, $C = \pi/2$ (e.g., see [3]).

$$\begin{aligned} \cos c &= \cos a \cos b \\ \tan a &= \tan A \sin b \qquad (1) \\ \sin a &= \sin A \sin c \end{aligned}$$

Let ABC be equilateral of side $2x$. Necessarily $x \leqq \pi/3$. Certainly, the triangle has one $2\pi/3$-point, the point Q symmetrically placed with respect to A, B, C; Q is the only $2\pi/3$-point if $0 < x \leqq \sin^{-1}\sqrt{2/3}$. However, for $\sin^{-1}\sqrt{2/3} < x \leqq \pi/3$, there are exactly three more unsymmetric $2\pi/3$-points Q_1, Q_2, Q_3, one on each altitude and if Q, Q_1 are on the altitude BX where $QX = q, Q_1X = q_1$ then q, q_1 are roots of $\sin q = \tan x/\sqrt{3}$, $q + q_1 = \pi$ and $q_1 > q$.

The unsymmetric $2\pi/3$-points are not minima of the problem. To prove this we consider the function $P\{ABC\}$ along the geodesic BX and show that the function has a maximum turning value at Q_1. Let P be on BX distance p from X. Then from (1) $AP = \cos^{-1}(\cos x \cos p)$ hence

$$f(p) = P\{ABC\} = 2\cos^{-1}(\cos x \cos p) + h - p$$

where h is the altitude.

$$\therefore \qquad f'(p) = \frac{2\sin p \cos x}{\sqrt{(1-\cos^2 x\cos^2 p)}} - 1$$

from which we deduce

$$f'(p) = 0 \quad \text{when } \sin p = \pm\frac{\tan x}{\sqrt{3}}.$$

Since we are only interested in values of p between 0 and $h(<\pi)$, the negative sign may be ignored and we see that $f(p)$ has critical values when $P = Q$ and $P = Q_1$. Further calculation gives

$$f''(p) = \frac{\sin^2 x \cos x \cos p}{(1-\cos^2 x \cos^2 p)^{3/2}}.$$

Thus $f''(p) \gtrless 0$ according as $\cos p \gtrless 0$. Since $q + q_1 = \pi$ and $q < q_1$, $\cos q_1 < 0$. Therefore $f''(q_1) < 0$ and $f(p)$ has a maximum turning value at Q_1 as asserted.

This result together with Theorems 1 and 2(i) imply that $\min\{ABC\} = Q$ or $\min\{ABC\} = A, B, C$ or both. For $x < \sin^{-1}\sqrt{2/3}$, the angles of the triangle are $< 2\pi/3$, hence by Theorem 2(ii), there is no minimum at A, B or C and for this range of values of x, Q is the unique solution. To find the solution for the remaining values of x, we consider the function

$$F(x) = Q\{ABC\} - B\{ABC\}$$

on the domain $[\sin^{-1}\sqrt{2/3}, \pi/3]$. Using (1)

$$F(x) = 3\sin^{-1}(2\sin x/\sqrt{3}) - 4x.$$

Direct calculation shows that $F(\sin^{-1}\sqrt{2/3}) < 0$ and $F(\pi/3) > 0$. Now $F(x) > 0$ if and only if $\sin^{-1}(2\sin x/\sqrt{3}) > 4x/3$. Each side of this inequality is less than $\pi/2$, hence $F(x) > 0$ if and only if $2\sin x/\sqrt{3} > \sin(4x/3)$. Writing $x = 3\theta$ and $\lambda = \cos\theta$ we may reduce this to

$$g(\lambda) = 4\sqrt{3}\lambda^3 - 4\lambda^2 - 2\sqrt{3}\lambda + 1 < 0.$$

Application of the intermediate value theorem shows that $g(\lambda)$ has precisely one zero between those values of λ corresponding to $x = \sin^{-1}\sqrt{2/3}$ and $x = \pi/3$. Hence $F(x)$ has precisely one zero say x^* in this interval. Then

for $\sin^{-1}\sqrt{2/3} < x < x^*$, $\min\{ABC\} = Q$;

for $x = x^*$, $\min\{ABC\} = A, B, C$ and Q;

for $x^* < x \leqq \pi/3$, $\min\{ABC\} = A, B, C$.

4. Solution for isosceles triangles. Similar methods may be used to establish the following solution for an isosceles triangle.

THEOREM 3. *Let ΔABC be isosceles with base $2x$ and altitude h. Then*

(i) *If $x > \pi/3$, ABC has only vertex minima.*

(ii) *For each x in $(0, \pi/3]$, there exists $h^*(x)$ in $(\sin^{-1}(\tan x/\sqrt{3}), \pi)$ such that solutions are as follows:* —

If $0 < h \leqq \sin^{-1}(\tan x/\sqrt{3})$, unique vertex minimum.
If $\sin^{-1}(\tan x/\sqrt{3}) < h < h^*(x)$, unique $2\pi/3$-point minimum.
If $h = h^*(x)$, one or two vertex minima and precisely one $2\pi/3$-point minimum.
If $h^*(x) < h \leqq \pi$, one or two vertex minima.

Further if $x \lesseqqgtr x^*$ (see Section 3), then $h^*(x) \lesseqqgtr$ the altitude of the equilateral triangle of side $2x$.

5. General triangle. The following two theorems enable us to deduce solutions for large classes of triangles having unequal sides from Theorem 3. We prove the first and leave the other to the reader.

THEOREM 4. *Let Q be a $2\pi/3$-point minimum of the isosceles triangle $A'B'C'$. Suppose that A, B, C are distinct from Q in the closed geodesic segments QA', QB', QC' respectively and such that at least one of $A \neq A'$, $B \neq B'$, $C \neq C'$ is true. Then Q is the unique minimum point for A, B, C.*

Proof. Assume that $A \neq A'$, i.e., $QA < QA'$ and suppose contrary to the theorem that there exists P of ΔABC such that $P\{ABC\} \leqq Q\{ABC\}$. Then

$$(AA' + AP) + (BB' + BP) + (CC' + CP) \leqq (AA' + AQ) + (BB' + BQ) + (CC' + CQ).$$

Applying the triangle inequality on the left we deduce

$$A'P + B'P + C'P \leqq A'Q + B'Q + C'Q. \tag{2}$$

If P is not on the great circle defined by QA, then $AA' + AP > A'P$ and (2) is strict, contrary to the minimum property of Q. If P is on this great circle, either (2) is strict giving a similar contradiction or (2) has equality in which case $\Delta A'B'C'$ has 2 nonvertex minima contrary to Theorem 3.

THEOREM 5. *Let A be a vertex minimum of the isosceles triangle $AB'C'$ and let B, C be distinct from A in the closed geodesic segments AB', AC' respectively such that either $B \neq B'$ or $C \neq C'$. Then A is the unique minimum point for A, B, C.*

References

1. E. J. Cockayne, On the Steiner problem, Canad. Math. Bull., 3 (1967) 431–450.
2. Z. A. Melzak, On the problem of Steiner, Canad. Math. Bull., 4 (1961) 143–148.
3. I. Todhunter, Spherical Trigonometry, Macmillan, London, 1929.
4. M. G. Greening, Solution to problem E2233, Amer. Math. Monthly, 4 (1970) and 3 (1971).

ABSTRACT MÖBIUS INVERSION

ROBERT SPIRA, Michigan State University

We give an exposition of the theory of Möbius inversion as developed by Morgan Ward and others (see references in Rota [**1**]). The theorems are given in a slightly different form in order to cover all previously known special cases.

A nonempty set P is said to be *partially ordered* by a relation $\leqq$ on P, if for all x, y and z in P,

(a) $x \leqq x$,

(b) $x \leqq y$ and $y \leqq x$ imply $y = x$,

(c) $x \leqq y$ and $y \leqq z$ imply $x \leqq z$.

We write $x < y$ if $x \leqq y$ and $x \neq y$. By a *segment* $[x, y]$ we mean $\{z \mid x \leqq z$ and $z \leqq y\}$. We define $(x, y]$ as $\{z \mid x < z$ and $z \leqq y\}$, and similarly define $[x, y)$ and (x, y). A segment is not necessarily linearly ordered. For example, let P be the positive integers under the partial order "divides"; the segment $[1, 6]$ is not linearly ordered.

To motivate the definitions we note that the Möbius μ function is determined by (a) $\mu(1) = 1$, (b) $\sum_{d|n} \mu(n) = 0$ if $n > 1$. In other words the value at $n > 1$ is determined by the values in the segment $[1, n)$. The central idea is to define a function $\mu(x, y)$ whose value on $[x, y]$ will be a sum of values determined in $[x, y)$.

We assume that the partially ordered set P is *locally finite*, i.e., that every segment $[x, y]$ is a finite set.

Let R be a ring with identity. We say a function f from $P \times P$ to R is an *incidence function* if $f(x, y) = 0$ unless $x \leqq y$. For two such incidence functions, f and g, we define the composition

$$(f * g)(x, y) = \begin{cases} \sum_{x \leqq z \leqq y} f(x, z) \cdot g(z, y), & \text{if } x \leqq y, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly $f * g$ is also an incidence function. We write $(f * g)(x, y)$ as $f * g(x, y)$. Defining $(f + g)(x, y)$ as $f(x, y) + g(x, y)$, it is now easy to verify:

THEOREM 1. *The incidence functions under $+$ and $*$ form a ring with identity δ, the Kronecker delta.*

Since segments are not necessarily linearly ordered some care must be taken in the proof of the associativity of $*$. Otherwise, the verification is perfectly straightforward. We call this ring, the *incidence ring* (over P and R). Next, we define three functions of this ring:

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x \leqq y, \\ 0 & \text{otherwise;} \end{cases} \qquad \mu(x, y) = \begin{cases} 1, & \text{if } x = y, \\ -\sum_{x \leqq z < y} \mu(x, z), & \text{if } x < y, \\ 0, & \text{otherwise,} \end{cases}$$

$$\mu_1(x,y) = \begin{cases} 1 & , \text{ if } x = y, \\ -\sum_{x<z\leqq y}\mu_1(z,y), & \text{if } x < y, \\ 0 & , \text{ otherwise.} \end{cases}$$

The function μ_1 is only temporarily introduced. We show shortly that $\mu = \mu_1$.

LEMMA 1. $\mu * \zeta = \delta = \zeta * \mu_1$.

Proof. If $x \nleqq y$, both $\mu * \zeta$ and δ are 0. For $x = y$, we have $\mu * \zeta(x,x) = \mu(x,x)$. $\zeta(x,x) = 1 \cdot 1 = 1 = \delta(x,x)$. If $x < y$, we have

$$\begin{aligned} \mu * \zeta(x,y) &= \textstyle\sum_{x\leqq z\leqq y}\mu(x,z)\zeta(z,y) = \sum_{x\leqq z\leqq y}\mu(x,z) \\ &= 0 = \delta(x,y). \end{aligned}$$

Similarly $\zeta * \mu_1 = \delta$.

LEMMA 2. $\mu = \mu_1$.

Proof. $\mu = \mu * \delta = \mu * (\zeta * \mu_1) = (\mu * \zeta) * \mu_1 = \delta * \mu_1 = \mu_1$.
Note that if $x \leqq y$,

$$\textstyle\sum_{x\leqq z\leqq y}\mu(z,y) = \sum_{x\leqq z\leqq y}\mu(x,z) = \delta(x,y).$$

We say that P is *left-finite* if for all y in P the set $P = \{x \mid x \leqq y\}$ is finite. If P is left-finite it has no descending chains. If P has a least element, it is left-finite.

THEOREM 2. *If P is left-finite, then f is a function from P to R and $g(y) = \sum_{x\leqq y} f(x)$ if and only if g is a function from P to R and*

$$f(y) = \sum_{x\leqq y} g(x)\mu(x,y).$$

Rota [**1**] gave the "only if" part of this theorem, assuming that there is a $p \in P$ such that $f(y) = 0$ unless $p \leqq y$. This assumption is not enough to have the theorem cover previously known special cases (Example 2 below). The two assumptions are independent.

Proof. Since P is left-finite, then sum $\sum_{x\leqq y} f(x)$ is a well defined finite sum so indeed g is a well defined function from P to R. Further,

$$\begin{aligned} \sum_{x\leqq y} g(x)\mu(x,y) &= \sum_{x\leqq y} \Big(\sum_{z\leqq x} f(z)\Big)\mu(x,y) \\ &= \sum_{x\leqq y} \Big(\sum_{z\leqq x} f(z)\zeta(z,x)\Big)\mu(x,y) \\ &= \sum_{x\leqq y} \Big(\sum_{z\leqq y} f(z)\zeta(z,x)\mu(x,y)\Big) \\ &= \sum_{z\leqq y} f(z) \Big(\sum_{x\leqq y} \zeta(z,x)\mu(x,y)\Big) \end{aligned}$$

$$= \sum_{z \leqq y} f(z) \left(\sum_{z \leqq x \leqq y} \zeta(z,x)\mu(x,y)\right)$$

$$= \sum_{z \leqq y} f(z)\zeta * \mu(z,y) = \sum_{z \leqq y} f(z)\delta(z,y) = f(y).$$

The other half of the theorem is proved similarly.

Finally, we show the connection with the other Möbius inversion formulas.

Example 1. Let P be the set of positive integers with the partial order of "divides". We show first that if $m \mid n$, then $\mu(m,n) = \mu(1,n/m)$. Let $n = km$. If $k = 1$, we have $\mu(m,m) = \mu(1,1) = 1$. Assume as an induction hypothesis that the result holds for $1, 2, \cdots, k-1$, where $k > 1$. Since m now properly divides n, 1 also properly divides n/m, so

$$\mu(m,n) = -\sum_{\substack{d \mid n \\ m \mid d \\ d \neq n}} \mu(m,d) = -\sum_{\substack{j \mid k \\ j \neq k}} \mu(m,jm)$$

$$= -\sum_{\substack{j \mid k \\ j \neq k}} \mu(1,j) = \mu(1,k) = \mu(1,n/m).$$

Finally, since the Möbius μ function is determined by $\mu(1) = 1$ and $\Sigma_{d \mid n}\mu(d) = 0$, for $n > 1$, from the definition of $\mu(1,n)$ we clearly have $\mu(n) = \mu(1,n)$. Since P has a least element, it is left finite and we can apply Theorem 2. The first condition then translates as $g(n) = \Sigma_{d \mid n} f(d)$ and the second as

$$f(n) = \Sigma_{d \mid n} g(d)\mu(d,n) = \Sigma_{d \mid n} g(d)\mu(1,n/d) = \Sigma_{d \mid n} g(d)\mu(n/d),$$

the usual Möbius inversion formulas.

Example 2. We take P to be the real numbers $\geqq 1$, and easily verify that the "divides" relation: $x \mid y$ if and only if y/x is a positive integer, is indeed a partial order. There is no least element in P, in the sense of the "divides" relation, so Rota's condition is not sufficient to carry the usual second inversion formula. We next can easily prove: x and y are in P and $x \mid y$ is equivalent to: y is in P and

$$x \in \{y, y/2, y/3, \cdots, y/[y]\}.$$

As in Example 1, we show that if $x \mid y$ then $\mu(x,y) = \mu(1,y/x)$, and $\mu(1,y/x) = \mu(y/x)$. The formulas in Theorem 2 then translate as:

$$\Sigma_{x \mid y} f(x) = \Sigma_{j=1}^{[y]} f(y/j) = \Sigma_{j \leqq y} f(y/j),$$

and similarly

$$\Sigma_{x \mid y} g(x)\mu(x,y) = \Sigma_{j \leqq y} g(y/j)\mu(j),$$

yielding the usual second Möbius inversion formulas.

Example 3. Let P be the set of positive integers under the partial order "is divisible by", and let P_N be $P \cap [1,N]$ under the same partial order. As above,

$\mu(jn, n) = \mu(j, 1) = \mu(j)$, and the formulas are

$$g(n) = \sum_{\substack{n|d \\ d \leqq N}} f(d) = \sum_{j=1}^{[N/n]} f(jn),$$

$$f(n) = \sum_{j=1}^{[N/n]} g(jn)\mu(j).$$

If now $f_x(n)$ is defined on P depending on a parameter x in R, then $g(n)$ will also depend on the parameter, say $g_x(n)$, and the above formulas will simply be the same except having subscripts x on f and g. As a special case, if F is a function from R to R, then we could define $f_x(n) = F(nx)$, where $nx = x + x + \cdots + x$ (n times). Putting $n = 1$ in the formulas gives

$$g_x^{(N)}(1) = \sum_{j=1}^{N} f_x(j) = \sum_{j=1}^{N} F(jx), \quad \text{and}$$

$$F(x) = \sum_{j=1}^{N} g_x^{(N)}(j)\mu(j),$$

where we have indicated the dependence on N. Under sufficient convergence hypotheses, $g_x^{(N)}(1)$ will tend to some $g(x)$, and $g_x^{(N)}(j)$ will tend to $g(jx)$, and we will obtain

$$G(x) = \sum_{j=1}^{\infty} F(jx) \text{ implies } F(x) = \sum_{j=1}^{\infty} G(jx)\mu(j),$$

and the converse will be obtained similarly, giving the other common case of Möbius inversion. For further results on this case, see Satyanarayana [3].

References

1. Gian-Carlo Rota, On the foundations of combinatorial theory I, Z. Warscheinlichkeitstheorie und Verw. Gebiete, 2 (1964) 340–368, MR 30 # 4688.

2. T. M. Apostol, A characteristic property of the Möbius function, Amer. Math. Monthly, 72(1965) 279–282, MR 31 # 2193.

3. U. V. Satyanarayana, On the inversion property of the Möbius μ-function II, Math. Gaz., 49 (1965) 171–178, MR 32 # 1156.

4. F. Dress, Formule d'inversion et fonction de Möbius sur les ensembles ordonnés, Sem. Delange-Pisot, No. 2, 11 pp.

5. E. M. Žmud, The group theoretic Möbius function, Zap. Meh.–Mat. Fak. i Har'kov. Mat. Obač., 28(1961) 73—96, MR 40 # 2647.

6. A. A. Mullin, On Möbius' function and related matters, Amer. Math. Monthly, 74(1967) 1100–1102, MR 37 # 143.

7. D. E. Daykin, Generalized Möbius inversion formulae, Quart. J. Math. Oxford Ser., 15(1964) 349–354, MR 30 # 4709.

8. T. M. Apostol, Möbius functions of order k, Pacific J. Math., 32(1970) 21–27, MR 40 # 7212.

9. V. S. Albis, A note on generalized Möbius μ-functions, Rev. Colombiana Mat., 2(1968) 6–11, MR 38 # 5699.

10. T. N. Stanciu, Cardinal function of Möbius, Stud. Cerc. Mat., 19(1967) 583–584, MR 40 # 4196

11. R. Wiegandt, On the general theory of Möbius inversion formula and Möbius product, Acta. Sci. Math. (Szeged), 20(1959) 164–180, MR 21 # 7176.

CONTINUOUS, EXACTLY K-TO-ONE FUNCTIONS ON R

B. R. WENNER, University of Missouri—Kansas City

Any linear function $f(x) = ax + b$ is a continuous one-to-one function on the real numbers R; in this note we consider the question of the existence of continuous, exactly k-to-one functions on R for $k > 1$ (f is said to be exactly k-to-one if for every y in the range of f there are exactly k distinct points $x_1, \cdots, x_k$ such that $i(x_1) = \cdots = f(x_k) = y$). The following results answer the question completely.

THEOREM 1. *There exists a continuous, exactly k-to-one function on R if and only if k is odd.*

THEOREM 2. *There exists a continuous, exactly k-to-one function from a non-degenerate closed interval into R if and only if $k = 1$.*

In the proofs of Theorems 1 and 2 we shall need only the following basic topological properties of R; they are immediate consequences of the fact that a continuous function preserves connectedness and compactness, respectively.

Let $a < b$, and let f be a continuous function from $[a, b]$ into R. Then

(1) for each y between $f(a)$ and $f(b)$ there exists $x \in (a, b)$ for which $f(x) = y$ [**1**, p. 67], and (2) there exist d and e in $[a, b]$ for which $f(d) = \sup f([a, b])$, and $f(e) = \inf f([a, b])$ [**1**, p. 86].

An immediate consequence is the following:

(3) Assume $a < b$, f is a continuous function from $[a, b]$ into R, and $f(a) = f(b)$. Then every value between $\inf f([a, b])$ and $\sup f([a, b])$ is assumed at least twice by f on $[a, b]$.

To see this we note that if $\sup f([a, b]) > c = f(a)$, then by (2) this value is assumed at some member d of (a, b), so by (1) every value between c and $f(d)$ is taken on at least once in each of (a, d) and (d, b). A similar argument shows that every value between $\inf f([a, b])$ and c is assumed at least twice; this proves (3), since the value c is taken on by at least the two endpoints a and b.

Theorem 1 will follow from (4) and (5) below. We first prove

(4) If k is even, then there does not exist a continuous function on R which is exactly k-to-one.

We assume the existence of such a function, where $k = 2m$, and hope to arrive at a contradiction. Let c be any member of the range of f; then $f^{-1}(c) = \{a_1, \cdots, a_k\}$ for $a_1 < \cdots < a_k$. For each $i = 1, \cdots, k-1$ it follows from (1) that either (i) $f(x) > c$ for all $x \in (a_i, a_{i+1})$, or (ii) $f(x) < c$ for all $x \in (a_i, a_{i+1})$. Let $i(1), \cdots, i(r)$ be the indices for which (i) holds, and let $d = \min\{\sup f([a_{i(j)}, a_{i(j)+1}]) : j = 1, \cdots, r\}$; then, by (3), each value between c and d (note that $d > c$) is assumed at least twice on each interval $(a_{i(j)}, a_{i(j)+1})$ for a total of at least $2r$ times, so $r \leqq m$ (recall that f fs $2m$-to-one). By the same reasoning, if $i(r+1), \cdots, i(r+s)$ are the indices for which (ii) holds, we have $s \leqq m$. But $r + s = 2m-1$, so either $r = m$ and $s = m-1$ or $r = m-1$ and $s = m$. Without loss of generality we may assume the former.

If $x < a_1$ or $x > a_k$, then $f(x) < c$ by (1), since every value in (c,d) is assumed $2r = k$ times between a_1 and a_k. Hence if $p = \sup f([a_1, a_k])$, then (by (2))

$$f^{-1}(p) = \{b_1, \cdots, b_k\}, \text{ where } a_1 = b_0 < \cdots < b_k < b_{k+1} = a_k.$$

Let $q = \max\{\inf f([b_i, b_{i+1}]): i = 0, 1, \cdots, k\}$; then $q < p$, and every value in (q, p) is taken on at least twice in each of the $k-1$ intervals $(b_1, b_2), \cdots, (b_{k-1}, b_k)$, by (3), and at least once in each of (a_1, b_1) and (b_k, a_k), by (1). Hence each of these values is assumed at $2(k-1)+2 = 2k$ different points, which contradicts the assumption that f was k-to-one.

(5) If k is odd, then there exists a continuous, exactly k-to-one function on R.

Let $k = 2m+1$, and for all $n = 0, +1, +2, \cdots$ we define $A_n = [kn, kn+(m+1)]$ and $B_n = [kn+(m+1), k(n+1)]$. Then $R = \bigcup \{A_n \cup B_n : n = 0, \pm 1, \pm 2, \cdots\}$, so we define the piecewise linear function f on R by

$$f(x) = \begin{cases} x - 2mn & \text{if } x \in A_n, \text{ and} \\ -x + (k+1)(n+1) & \text{if } x \in B_n. \end{cases}$$

If $x = kn + (m+1)$, then $f(x)$ is given as the value $n + m + 1$ by the formula for each of A_n and B_n, and if $x = kn$ then $f(x) = n$ on each of B_{n-1} and A_n, so f is a continuous function on R. We note further that $\sup f(A_n) = \sup f(B_n) = n + m + 1$, $\inf f(A_n) = n$, and $\inf f(B_n) = n + 1$.

If $y \in R$ then there exists an integer n for which $n < y < n+1$ or $y = n$. If the first holds, then the value y is assumed exactly once on each of the k intervals $A_{n-m}, B_{n-m}, A_{n-m+1}, \cdots, B_{n-1}, A_n$ by (1), and is assumed on no other A_i or B_i (for example, $\sup f(A_i) \leqq n < y$ for all $i < n-m$). But if $y = n$, then y is assumed by the two numbers $(kn - km - m) \in A_{n-m-1} \cap B_{n-m-1}$ and $kn \in B_{n-1} \cap A_n$, and once each on the $2m-1 = k-2$ intervals $A_{n-m}, B_{n-m}, \cdots, A_{n-2}, B_{n-2}, A_{n-1}$, but on no other A_i or B_i. Hence y has exactly k distinct preimages in either case, so f is exactly k-to-one.

For Theorem 2, the proof of (4) can be used without alteration to dispose of even k, so we now assume $f: [a,b] \to R$ is continuous and exactly k-to-one for k odd and $\geqq 3$. The value $c = \sup f([a,b])$ is assumed at exactly k distinct points $\{b_1, \cdots, b_k\}$, where $a \leqq b_1 < \cdots < b_k \leqq b$, so the value $d = \max\{\inf f([b_i, b_{i+1}]): i = 1, \cdots, k-1\}$ is less than c. By (3), any value in (d, c) is assumed at least twice in each of the $k-1$ intervals (b_i, b_{i+1}) for a total of at least $2(k-1)$ times; but $2(k-1) > k$ for $k \geqq 3$, which contradicts the assumption that f is k-to-one.

We conclude this communication with the remark that the situation regarding mappings on open intervals is precisely that described in Theorem 1. Indeed, if $f: (a,b) \to R$ is continuous then there exists a homeomorphism $h: R \to (a,b)$ [**1**, p. 93], and $f \circ h: R \to R$ is continuous and is exactly k-to-one if and only if f is.

Reference

1. T. O. Moore, Elementary General Topology, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

GEOMETRIC INTERPRETATIONS OF SOME CLASSICAL INEQUALITIES

JOSEPH L. ERCOLANO, Baruch College, CUNY

For a, b any positive real numbers, consider the following classical means relating them:

A.M.	$= \frac{1}{2}(a+b)$	(arithmetic mean)
G.M.	$= \sqrt{ab}$	(geometric mean)
H.M.	$= 2ab/(a+b)$	(harmonic mean)
R.M.	$= \sqrt{\frac{1}{2}(a^2+b^2)}$	(root-mean square)

Let $\max(a,b)$ and $\min(a,b)$ denote respectively the maximum and minimum values of a and b. Then it is well known [1,2] that the following set of inequalities relating these six quantities is true:

$$\max(a,b) \geqq \text{R.M.} \geqq \text{A.M.} \geqq \text{G.M.} \geqq \text{H.M.} \geqq \min(a,b).$$

The purpose of this note is to present a geometric interpretation of these inequalities. We will do this, with straightedge-and-compass constructions, in two different settings.

Suppose $a > b > 0$. In Figure 1, the lengths of AC and BC are a and b, respectively, while in Figure 2, these are the lengths of AB and BC. Then $2ab/(a+b)$, $\sqrt{ab}$, $\frac{1}{2}(a+b)$, and $\sqrt{\frac{1}{2}(a^2+b^2)}$ are the lengths of EC, DC, OC, and FC respectively in Figure 1, and of ED, BD, OC, and BF respectively in Figure 2. In Figure 2, $a = b$ is permissible.

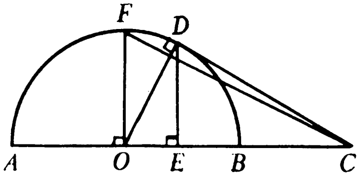


FIG. 1

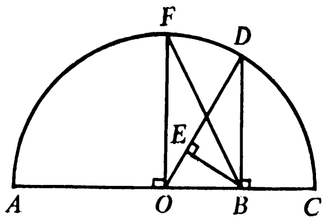


FIG. 2.

References

1. E. Beckenbach and R. Bellman, An Introduction to Inequalities, Random House, New York, 1961.

2. D. S. Mitrinovic, Elementary Inequalities, Noordhoff, Groningen, The Netherlands, 1964.

BOOK REVIEWS

EDITED BY D. ELIZABETH KENNEDY, University of Victoria

Materials intended for review should be sent to: Professor D. Elizabeth Kennedy, Department of Mathematics, University of Victoria, British Columbia, Canada.

Reviews of texts at the freshman-sophomore level based upon classroom experience will be welcomed by the Book Review Editor.

A boldface capital **C** *in the margin indicates a classroom review.*

Integral Transforms in Applied Mathematics. By John Miles. Cambridge University Press, (Macmillan of Canada), 1971. 97 pp. $6.95.

Anyone with a basic knowledge of ordinary and partial differential equations and complex variable theory will find in this book a lucid, nonrigorous introduction to integral transform theory and its applications.

The first chapter presents integral transform theory in a style which, according to the author, follows the line set down by Lord Rayleigh. Thus formal, rather than rigorous, derivations are given. Further chapters deal in the same style with the Laplace transform, Fourier transforms, Hankel transforms and finite Fourier transforms. Each chapter contains worked examples and suggested exercises showing uses of the transform in some applications. Dr. Miles draws on areas in which he has done authoritative research to present examples in fluid mechanics, heat conduction, electric circuits, mechanical vibration and wave motion. The examples are treated in such a way that previous knowledge of these fields is not essential.

The book includes, in Appendix 2, a short table of transform pairs, which are sufficient for the examples used. Further tables and texts are referred to in an annotated bibliography.

P. VAN DEN DRIESSCHE, *University of Victoria.*

Guide Book to Departments in the Mathematical Sciences. Fifth Edition, 1972. The Mathematical Association of America, Washington, D. C. 135 pp. $1.30.

This Guidebook to Departments in the Mathematical Sciences is intended to provide in summary form information about the location, size, staff, library facilities, course offerings, and special features of departments in mathematical sciences in four year colleges and universities in the United States and Canada. Its purpose is to assist prospective students in these countries and abroad, in high school, junior college or in college, and their counselors and parents in obtaining comparable information about many institutions so that the selection of a proposed place of study may be narrowed down to a few from which more detailed information may be sought on an individual basis.

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles Valley College

ASSOCIATE EDITOR, J. S. FRAME, Michigan State University

Readers of this department are invited to submit for solution problems believed to be new that may arise in study, in research, or in extra-academic situations. Problems may be submitted from any branch of mathematics and ranging in subject content from that accessible to the talented high school student to problems challenging to the professional mathematician. Proposals should be accompanied by solutions, when available, and by any information that will assist the editor. Ordinarily, problems in well-known textbooks should not be submitted.

The asterisk (*) will be placed by the problem number to indicate that the proposer did not supply a solution. Readers' solutions are solicited for all problems proposed. Proposers' solutions may not be "best possible" and solutions by others will be given preference.

Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink and exactly the size desired for reproduction.

Send all communications for this department to Robert E. Horton, Los Angeles Valley College, 5800 Fulton Avenue, Van Nuys, California 91401.

To be considered for publication, solutions should be mailed before January 15, 1973.

PROPOSALS

838. *Proposed by Stanley Rabinowitz, Far Rockaway, New York.*

Mr. Jones makes n trips a day to his bank to remove money from his account. On the first trip he withdrew $1/n^2$ percent of the account. On the next trip he withdrew $2/n^2$ percent of the balance. On the kth trip he withdrew k/n^2 percent of the balance left at that time. This continued until he had no money left in the bank. Show that the time he removed the largest amount of money was on his last trip of the tenth day.

839. *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey.*

Given three boxes each containing w white balls and r red balls identical in shape. Take a ball from the first box and put it in the second box, then take a ball from the second box and put it in the third, and finally take a ball from the third box and put it in the first. Find the probability that the boxes have their original contents as to color.

840. *Proposed by Charles W. Trigg, San Diego, California.*

Show that in the system of numeration with base five, the set of palindromic triangular numbers is infinite.

841. *Proposed by Murray S. Klamkin, Ford Motor Company, Dearborn, Michigan.*

Solve the following generalization of Clairaut's equation:

$$y = xp + F(p)\{1 + \sqrt{1 + xG(p)}\}$$

where $p = dy/dx$.

842. *Proposed by Kenneth Fogarty and Erwin Just, Bronx Community College, New York.*

Prove that there exists an infinite set of quadratic monic polynomials, f, such that the four roots of $f[f(x)] = 0$ are distinct integers.

843. *Proposed by Leon Bankoff, Los Angeles, California*

The center w_0 of a circle whose radius is p_0 lies at the vertex of a parabola and is tangent to the latus-rectum AB at the focus F. Touching the latus-rectum and centered on the parabola is a sequence of successively tangent circles $(w_1), (w_2), \cdots (w_n)$ of radii $p_1, p_2 \cdots p_n$, the first of which is tangent to (w_0). Find a general expression for the calculation of p_n in terms of p_0.

844. *Proposed by Gregory Wulczyn, Bucknell University, Pennsylvania.*

Show that there is an infinite number of sets of three consecutive integers such that the sum of the square of the first, twice the square of the second, and three times the square of the third is a square.

QUICKIES

From time to time this department will publish problems which may be solved by laborious methods, but which with the proper insight may be disposed of with dispatch. Readers are urged to submit their favorite problems of this type, together with the elegant solution and the source, if known.

Q547. Let f be a continuous function such that

$$0 < A \leqq f(x) \leqq B$$

for $0 \leqq x \leqq 1$. Prove that

$$AB \int_0^1 dx/f(x) \leqq A + B - \int_0^1 f(x)dx$$

[*Submitted by E. F. Schmeichel*]

Q548. Let T denote the irrationals in the interval $I = [0, 1]$. Prove that both T and I are uncountable.

[*Submitted by Warren Page*]

Q549. Find a sequence $\{a_n\}$ of real numbers such that $\{a_n\}$ converges to zero and $\{na_n\}$ converges to a transcendental number.

[*Submitted by Michael Jones*]

Q550. Find the differential equation of the family of coplanar circles having a constant radius r.

[*Submitted by C. S. Venkataraman*]

Q551. Prove that the diophantine equation

$$5^x + 2 = 17^y$$

has no solutions.

[*Submitted by Erwin Just*]

(Answers on pages 239–240.)

SOLUTIONS

Errata: The name of Murray S. Klamkin was omitted from the list of Also Solvers of Problem 801 (March 1972, page 109).

In the solution of Problem 802 (March 1972, page 109) the following misprints occurred:

(1) The first line of the statement of Lemma 1 reads "If $(a,x) \equiv 1, a^n \equiv b^n \pmod{x}$ and $a \equiv b \pmod{x} \cdots$". The assertion "$a \equiv b \pmod{x}$" should have been "$a \not\equiv b \pmod{x}$".

(2) The fourth line in the proof of Lemma 1 reads

$$\text{“}\equiv (b^n)^c \equiv (b^n)^c(b^{\phi(x)})^d \equiv b^{cn+a\phi(x)}\ldots\text{”}.$$

The term "$b^{cn+a\phi(x)}$" should have been "$b^{cn+d\phi(x)}$".

(3) The sixth line in the proof of Lemma 2 begins with "it follows that $x^n + y^n \equiv z \pmod{x}$". The assertion "$x^n + y^n \equiv z \pmod{x}$" should have been "$x^n + y^n \equiv z^n \pmod{x}$".

An Alphametic

810. [November, 1971] *Proposed by Zalman Usiskin, University of Michigan.*

Solve the following cryptarithm (in base 10):

$$\begin{array}{r} S\ I\ X \\ \times \quad T\ W\ O \\ \hline \text{— — — —} \\ \text{— — — —}\quad \\ \text{— — — —}\quad\quad \\ \hline T\ W\ E\ L\ V\ E \end{array}$$

Solution by J. A. H. Hunter, Toronto, Canada.

This may be solved without undue figuring by algebraical means. We set $(TW) = x$, $(O) = y$, $(ELVE) = z$, where obviously $x \geqq 23$. By quick trial it is found that $SIX = 987$ or 986.

Then, $10000x + z = 987(10x + y)$ or $986(10x + y)$.

Say $SIX = 987$, with $z = 987y - 130x$. Testing for $y = 6, 5, 4$ and bearing in mind that duplications are not permitted and that the initial and fourth digits of z are the same, it is found that no solution exists.

Then $SIX = 986$, with $z = 986y - 140x$. The letters O and E cannot stand for the same digit, hence y must be odd. Testing for $y = 7, 5$ a solution comes with $y = 5,\ x = 34,\ z = 0170$.

Hence we have: $986 \cdot 345 = 340170$.

Comment by the proposer.

Please note ERROR in statement of Problem 810.

The problem should read:

$$\begin{array}{rrrrrr} & & & S & I & X \\ \times & & & T & W & O \\ \hline & & - & - & - & - \\ & & - & - & - & \\ - & - & - & - & & \\ \hline T & W & E & L & V & E \end{array}$$

The error is in the second line of the partial product. There should be only three dashes, not four. This changes the entire nature of the problem.

Under these conditions the solution would be:

$$\begin{array}{rrrrrr} & & & 9 & 6 & 5 \\ \times & & & 2 & 1 & 8 \\ \hline 2 & 1 & 0 & 3 & 7 & 0 \end{array}$$

Also solved by Merrill Barnebey, University of Wisconsin at LaCrosse; Virginia Bolton, Oxford, Georgia; Richard L. Breisch, Pennsylvania State University; Charles M. Campanile, North Bellmore, New York; Arlyn Colby, Barron, Wisconsin; M. G. Greening, University of New South Wales, Australia; Lee Halverson, Clintonville, Wisconsin; Sidney Kravitz, Dover, New Jersey; Lebanon Valley College Problem Seminar; Norman T. Lyon, Caledonia, Ohio; Paul P. McCarty, La Salle College, Pennsylvania; William Nuesslein, SUNY-Geneseo; Regina Parker, Lew Wallace High School, Gary, Indiana; Karl Rubin, Woodrow Wilson High School, Washington, D. C.; E. F. Schmeichel, Itasca, Illinois; Paul Sugarman, Massachusetts Institute of Technology; Kenneth M. Wilke, Topeka, Kansas; Mark Winholtz, Luther College, Iowa; and the proposer.

A Problem of Visibility

811. [November, 1971] *Proposed by James Bookey, Mount Senario College, Wisconsin.*

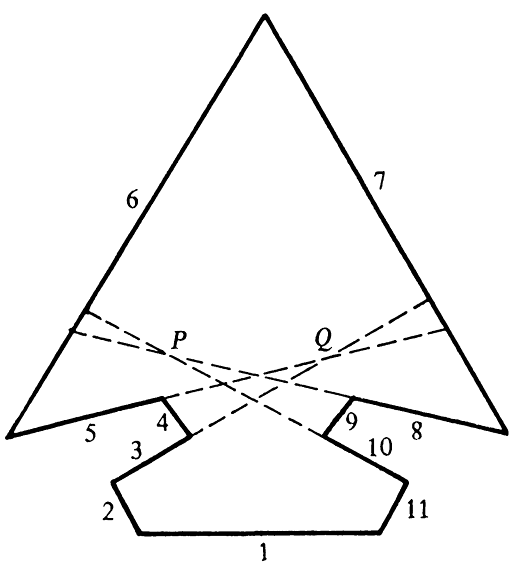
Find a domain bounded by a simple closed polygon such that for each two sides of the polygon there is an interior point from which these two sides are visible, but such that there is no interior point from which all sides are visible.

I. *Solution by Norman G. Gunderson, University of Rochester.*

Divide each side of an equilateral triangle T of altitude 4 into four congruent segments. Label the vertices and points of division A_1 through A_{12} cyclically starting at any vertex of T. Erect equilateral triangles on $\overline{A_1A_2}$, $\overline{A_5A_6}$, and $\overline{A_9A_{10}}$, externally to T, and delete these segments. The resulting polygon has the desired property since any specific pair of sides is visible from some point inside T within 1 of one of the sides of T, but for all sides to be visible from a point P, P would have to be within 1 of all three sides of T.

II. *Solution by Michael Goldberg, Washington, D.C.*

A solution is shown as the arrow-shaped polygon in which the sides are numbered from 1 to 11. The sides 3 and 5 intersect in the point Q. The sides 8 and 10 intersect in the point P. Sides 3 and 5 are simultaneously visible only from points in the small region to the right of Q. Sides 8 and 10 are simultaneously visible only from points in the small region to the left of P. Hence, there are no points from which sides 3, 5, 8 and 10 are simultaneously visible. By trial, it is verified that for each pair of sides, there is a point from which these sides are entirely visible.



Also solved by Richard L. Breisch, Pennsylvania State University; Romae J. Cormier, De Kalb, Illinois; Robert G. Griswold, University of Hawaii; H. S. Hahn, West Georgia College; Michael Legacy, State University College of Arts and Science, Plattsburg, New York; Carl Rubin, Woodrow Wilson High School, Washington, D. C.; E. F. Schmeichel, Itasca, Illinois; Jim Tattersall, Attleboro, Massachusetts; William Nuesslein, SUNY, Geneseo; Barry Woods, SUNY, Plattsburg; Kenneth L. Yocom, South Dakota State University; and the proposer.

Order-Isomorphic Groups

812. [*November*, 1971] *Proposed by Donald P. Minassian, Butler University, Indiana.*

Two fully ordered groups G and H are order-isomorphic if there is an isomorphism f from G to H which preserves orderings: $a > b$ in G if and only if $f(a) > f(b)$ in H. Let R be the additive group of real numbers under the usual ordering. Show that no proper extension, and no proper subgroup of R is order-isomorphic to R: we assume that the extension and subgroup are ordered to preserve the ordering of R.

Solution by William Nuesslein, State University of New York at Geneseo.

This problem consists of two parts.
(i) If $f:(R,+) \to H - (R,+)$ is order isomorphic, then $H = R$.
(ii) If $g: G \to (R,+)$ is order isomorphic and $G \subseteq R$, then $G = R$.

Proof of (i). It is easy to show that f^{-1} is also order isomorphic. For $\varepsilon > 0$,

$$f(a) < f(b) < f(a) + \varepsilon \text{ if and only if } a < b < a + f^{-1}(\varepsilon).$$

Letting $\delta = f^{-1}(\varepsilon) > 0$ proves that f is uniformly continuous on R.

By induction, $f(m/n) = (m/n)f(1)$ for integers m, n with $n \neq 0$. Thus $f(x) = kx$ on the rationals, $k = f(1)$. By continuity $f(x) = kx$ on R. This function is onto and so $H = R$.

Part (ii) follows by applying part (i) to g^{-1} which is also order isomorphic.

Also solved by H. S. Hahn, West Georgia College; Richard Kerns, Hamburg, Germany; E. F. Schmeichel, Itasca, Illinois; and the proposer. One unsigned correct solution was received.

A Polynomial Property

813. [November, 1971] *Proposed by L. Carlitz and R. A. Scoville, Duke University.*

Show that any polynomial with real coefficients can be written as a difference of two real monotone increasing polynomials.

Solution by K. L. Yocom, South Dakota State University.

First suppose $g(x) = ax^n$. The case when $a = 0$ is trivial and the cases when $n = 0$ or n is odd are handled by writing

$$g(x) = (a + |a| + 1)x^n - (|a| + 1)x^n$$

which expresses $g(x)$ as the difference of two monotone increasing polynomials. Now suppose n is even, say $n = 2k$, $k \geqq 1$, and $a \neq 0$. Then for $a > 0$ write

$$ax^{2k} = (ax^{2k+1} + 2kax) - (ax^{2k+1} - ax^{2k} + 2kax).$$

The polynomials in parentheses are monotone increasing as can be seen by examining their derivatives. Now for $a < 0$, write

$$ax^{2k} = (-ax^{2k+1} + ax^{2k} - 2kax) - (-ax^{2k+1} - 2kax)$$

which again expresses $g(x)$ as the difference of two monotone increasing polynomials. For the general polynomial $g(x) = (a_n x^n\colon 0 \leqq n \leqq m)$ simply apply the above to each individual term $a_n x^n$.

Also solved by Bernard C. Anderson, Henry Ford Community College, Michigan; J. C. Binz, Bern, Switzerland; James Bookey, Mount Senario College, Wisconsin; James E. Brewer, Naples High School, Naples, Florida; Richard A. Gibbs, Fort Lewis College, Colorado; Michael Goldberg, Washington, D. C.; M. G. Greening, University of New South Wales, Australia; H. S. Hahn, West Georgia College; Burrell W. Helton, San Marcos, Texas; Murray S, Klamkin, Ford Motor Company; Arturo Kung-Fu, Hamburg, Germany; Carl P. MacCarty, La Salle College, Pennsylvania; William Nuesslein, SUNY Geneseo; Warren Page, New York City Community College; D. C. Pfaff, University of Nevada, Reno; I. H. Nagaraga Rao, Andrha University, India; Lawrence A. Ringenberg, Eastern Illinois University; Rina Rubenfeld, New York City Community College; E. F. Schmeichel, Itasca, Illinois; Michael Stolnicki, Oakland Community College, Michigan; Burnet R. Toskey, Seattle University; and the proposers.

Tangents of Inverse Functions

814. [November, 1971] *Proposed by Marlow Sholander, Case Western Reserve University.*

For what values of a is the graph of a^x tangent to the graph of $\log_a x$?

Solution by Vaclav Konecny, Jarvis Christian College, Hawkins, Texas.

$y = \log_a x$ $(a > 0,\ a \neq 1,\ x > 0)$ is the inverse function to $y = a^x$. From the symmetry of the graphs about the straight line $y = x$ we conclude that they may have one common tangent $y = x$ or perpendicular to that at a point on $y = x$. Therefore at a point of tangency $y^1 = a \ln a = +1$ and $x = a^x$. The solutions of these equations are $a = e^{1/e}$, $(x = e)$ and $a = e^{-e}$, $(x = e^{-1})$ for $+$ and $-$ signs, respectively. The graph of a^x is tangent to the graph of $\log_a x$ for $a = e^{1/e}$ and $a = e^{-e}$.

Also solved by Gladwin E. Bartel, La Junta, Colorado; Merl J. Biggin, Washington, D. C.; Virginia F. Bolton, Oxford, Georgia; Richard L. Breisch, Pennsylvania State University; Romae J. Cormier, De Kalb, Indiana; Fred Dodd and Leon Mattics (Jointly), University, of South Alabama; Robert N. Eckert, Culver Military Academy, Indiana; R. Fourneau, Université de Liège, Belgium; Richard A. Gibbs, Fort Lewis College, Colorado; Michael Goldberg, Washington, D. C.; M. G. Greening, University of New South Wales, Australia; Robert G. Griswold, University of Hawaii at Hilo; H. S. Hahn, West Georgia College; Heiko Harborth, Braunschweig, Germany; John M. Howell, Littlerock, California; Paul Ilacqua, Santa Clara University; Murray S. Klamkin, Ford Motor Company; Lew Kowarski, Morgan State College, Maryland; Vaclav Konecny, Hawkins, Texas; William Nuesslein, SUNY-Geneseo; C. Stanley Ogilvy, Hamilton College, New York; Joseph O'Rourke, Saint Joseph's College, Pennsylvania; Warren Page, New York City Community College; C. B. A. Peck, State College, Pennsylvania; E. F. Schmeichel, Itasca, Illinois; Jim Tattersall, Attleboro, Massachusetts; Wolf R. Umbach, Rottorf, Germany; K. L. Yocom, South Dakota State University; and the proposer.

A Vector Relationship

815. [November, 1971] *Proposed by Sidney H. L. Kung, Jacksonville University, Florida.*

Given two complex numbers z_1 and z_2 whose sum is z. Let the angle between the vectors $0z_1$ and $0z$ be designated by θ. Let $0z_3$ and z_1z_3 intersect at z_3 such that:

(a) angle $z_20z_3 = \theta =$ angle $0z_1z_3$,

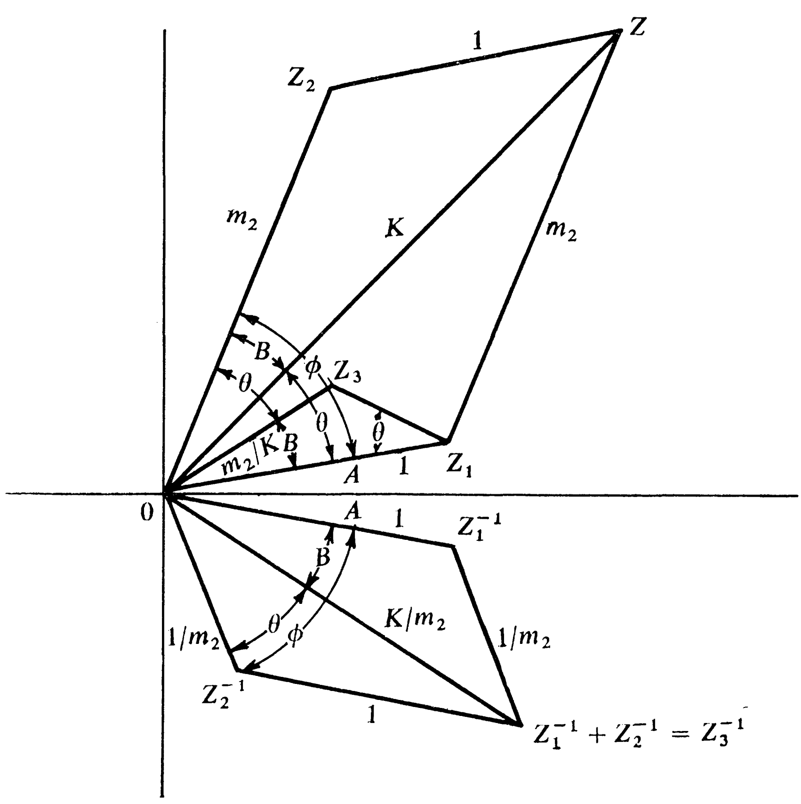
(b) the vector $0z_3$ lies within the angle ϕ subtended by $0z_1$ and $0z_2$; $0 < \phi < \pi$.

Prove that $z_3 = (z_1^{-1} + z_2^{-1})^{-1}$.

Solution by Michael Goldberg, Washington, D.C.

The construction described in the problem is a good graphical method that could be used by electrical engineers to determine the impedance and phase angle of two given impedances in parallel.

The validity of this construction is shown in the accompanying Argand diagram. The complex number z_1 has the amplitude angle A. Without loss of generality, its modulus is taken as unity. The number z_2 has the amplitude $(A + \phi)$ and its modulus is m_2. Then z_1^{-1} has the amplitude $-A$ and unit modulus; while z_2^{-1} has the amplitude $-(A + \phi)$ and its modulus is $1/m_2$. Hence, the parallelogram $0z_1^{-1}z_3^{-1}z_2^{-1}$ is similar to the parallelogram $0z_2zz_1$, and the corresponding angles are equal. Note that $B = \phi - \theta$, and that all the triangles in the figure are similar. Hence, the amplitude of z_3^{-1} is $-(A + B)$ and, therefore, the amplitude of z_3 is $(A + B)$. Then the modulus of z_3^{-1} is k/m_2, and the modulus of z_3 is m_2/k.



Also solved by Fred Dodd, University of South Alabama; Richard A. Gibbs, Fort Lewis College, Colorado; M. G. Greening, University of New South Wales, Australia; H. S. Hahn, West Georgia College; Vladimir F. Ivanoff, San Carlos, California; Murray S. Klamkin, Ford Motor Company; Vaclav Konecny, Jarvis Christian College, Texas; E. F. Schmeichel, Itasca, Illinois; Wolf R. Umbach, Rottorf, Germany; and the proposer.

Centroid of an Ellipse

816. [November, 1971] *Proposed by Murray S. Klamkin, Ford Scientific Laboratory, Dearborn, Michigan.*

Show that no equilateral triangle which is either inscribed in or circumscribed about a noncircular ellipse can have its centroid coincide with the center of the ellipse.

Solution by Leon Bankoff, Los Angeles, California.

If an equilateral triangle and a *circumscribed* ellipse were to share the same centroid, the ellipse and the circumcircle of the triangle would be concentric. Consequently the four intersections of the circle and the ellipse would be vertices of a rectangle. Since the vertices of an equilateral triangle cannot lie on three vertices of a rectangle, the initial assumption regarding a common centroid is untenable.

The same assumption for the *inscribed* ellipse would mean that each chord of contact of the ellipse would be bisected by a corresponding internal angle bisector of the tangential equilateral triangle. This could occur only if each vertex of the triangle lay on an extended principal axis of the ellipse. This, in turn, would necessitate two vertices on one axis—an obvious impossibility for a circumscribed triangle.

Also solved by Stanley R. Clemens, Illinois State University; Michael Goldberg, Washington, D. C.; M. G. Greening, University of New South Wales, Australia; H. S. Hahn, West Georgia College; Vladimir F. Ivanoff, San Carlos, California; Vaclav Konecny, Jarvis Christian College, Texas; Lew Kowarski, Morgan State College, Maryland; E. F. Schmeichel, Itasca, Illinois; E. P. Starke, Plainfield, New Jersey; and the proposer.

Comment on Problem 790

790. [March, 1971 and January, 1972] *Proposed by Stanley Rabinowitz, Far Rockaway, New York*

(1) Find all triangles ABC such that the median to side a, the angle bisector of angle B and the altitude to side c are concurrent.

(2) Find all such triangles with integral sides.

Comment by John P. Hoyt, Indiana University of Pennsylvania.

On page 50 of the January, 1972 issue of this MAGAZINE is the statement: "It is known that all triangles having integer sides must contain one or another of the angles 60°, 90°, or 120°." This is obviously not a true statement for it would imply that the only isosceles triangle with integer sides would be an equilateral triangle.

It is obvious that the isosceles triangle whose sides are 1, $a+1$, and $a+1$ for any positive integer a does not have a 60°, a 90°, or a 120° angle.

Comment on Problem 797

797. [May, 1971] *Proposed by Frank J. Papp, University of Lethbridge, Alberta, Canada.*

Determine the critical points and relative extrema, if any, of the two functions $f(x_1,\cdots,x_n)$ and $g(x_1,\cdots,x_n)$ for $n = 1,2,3\cdots$ where

$$f(x_1,\cdots,x_n) = \det(a_{ij})g(x_1,\cdots,x_n) = \det(b_{ij})$$

with

$$a_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ or if } i = j = 1 \\ 1 + x_{i-1} & \text{if } i = j = 2,3\cdots n+1 \end{cases}$$

$$b_{ij} = \begin{cases} 1 & \text{if } i \neq j \\ 1 + x_i & \text{if } i = j = 1,2,\cdots,n. \end{cases}$$

Comment on Solution [Mar.-Apr. 1972] *by Derrill J. Bordelon, Naval Underwater Systems Center, Newport, Rhode Island.*

The Hessian of f, in this case,

$$f = \prod_{i=0}^{n} x_i \sum_{i=0}^{n} x_i^{-1}, \quad x_0 = 1$$

at a critical point is neither negative nor positive definite does not *alone* imply that f fails to possess a relative extremum at that critical point. However, the existence of a positive and also a negative eigenvalue of the Hessian of f at the critical point does imply that the critical point is not a relative extremum point. Since at the critical point $x_i = (1-n)$, $i = 1,2,\cdots n$; $\partial^2 f/\partial x_i \partial x_j = -(1-n)^{n-3}$ for $i \neq j$, $n > 1$, etc., det (Hessian of f) $= (1-n)^{n(n-3)+1} < 0$ and the eigenvalues, λ, of the Hessian of f are $\lambda = (1-n)^{n-3}$, of multiplicity $n-1$, and $\lambda = (1-n)^{n-2}$ of multiplicity 1; the Hessian of f at the critical point is indefinite and the critical point is not a relative extremum point.

Since the Hessian of g at the critical point $x_i = \pm\infty$, $i = 1,2,\cdots,n$ is the zero matrix, and the n-dimensional Taylor expansion of g at the critical points is identically zero; the critical points are absolute not relative extremum points [*Calculus*, Vol II, T. M. Apostol, second edition, Blaisdell, Waltham, Mass., 1969 p. 304].

Comment on Q505

Q505. [January, November, 1971] Solve the differential equation:

$$(x-a)(x-b)y'' + 2(2x-a-b)y' + 2y = 0.$$

[*Submitted by Gregory Wulczyn*]

Comment by Murray S. Klamkin, Ford Motor Company.

The still more general equation

$$(x-a)(x-b)y''+n(2x-a-b)y' = F(x), \ (n \ arbitrary)$$

can also be solved easily by first noting the factorization

$$\{(x-a)D+n\}\{(x-b)D+n-1\}y = F(x).$$

Then by the exponential shift theorem,

$$y = (x-b)^{1-n}\int \frac{(x-b)^{n-2}dx}{(x-a)^n}\int F(x)(x-a)^{n-1}dx.$$

On letting $F(x) = 0$, we find on comparison with my previous comment (Nov.–Dec., 1971) that

$$\text{(1)} \qquad \int \frac{(x-b)^{n-2}dx}{(x-a)^n} = A'\left\{\frac{x-b}{x-a}\right\}^{n-1} + B'$$

which at first glance is somewhat surprising since one would expect a series by expanding out $\{(x-a)+(a-b)\}^{n-2}$ (here $A' = 1/(n-1)(b-a)$). This leads to the summation

$$\sum_{r=0}^{n}\frac{\binom{n}{r}}{(r+1)}\left\{\frac{a-b}{x-a}\right\}^{r+1} = \frac{1}{n+1}\left\{\left(\frac{x-b}{x-a}\right)^{n+1}-1\right\}.$$

A further extension to an rth order equation is given by the following: If $S_i (i=0,1,\cdots,r)$ denote the elementary symmetric functions of $x-a_i$ $(i=1,2,\cdots,r)$, i.e.,

$$\prod_{i=1}^{r}(\lambda+x-a_i) \equiv \sum_{i=0}^{r} S_i\lambda^{r-i},$$

then the solution of the differential equation

$$\text{(2)} \qquad \sum_{j=0}^{r} j!\binom{n}{j} S_{r-j}D^{r-j}y = 0$$

is given by

$$y = \sum_{i=1}^{r} A_i(x-a_i)^{r-n-1} \qquad (A_i \text{ arbitrary constants}).$$

The latter follows since it can be shown by induction that (2) factorizes into

$$\{(x-a_i)D+n\}\ \{(x-a_2)D+n-1\}\cdots$$
$$\{(x-a_r)D+n-r+1\}y = 0$$

for any ordering of the a_i's. The nonhomogeneous equation corresponding to (2) can also be solved by quadratures by means of the exponential shift theorem. Also, corresponding to (1) for $r = 3$, we have

$$\int \frac{(x-c)^{n-3}dx}{(x-b)^{n-1}} \int \frac{(x-b)^{n-2}dx}{(x-a)^n}$$
$$= A\left\{\frac{x-c}{x-a}\right\}^{n-2} + B\left\{\frac{x-c}{x-b}\right\}^{n-2} + C.$$

There are analogous results for $r > 3$.

Comment on Q530

Q530. [November, 1971] Suppose $a-1$ and $a+1$ are twin primes larger than 10. Prove that $a^3 - 4a$ is divisible by 120.

[*Submitted by George E. Andrews*]

Comment by A. L. Andrew, La Trobe University, Victoria, Australia.

A different, and perhaps simpler, argument used for a similar problem by Keith Burns (*Mathematical Gazette*, vol. 55 (1971) p. 51), yields the following stronger result. "If $a-1$ and $a+1$ are odd numbers not divisible by 5, and $a > 2$, then $a^3 - 4a$ is divisible by 240."

Proof. One of the consecutive even numbers $a-2$, a, $a+2$ is divisible by 3 and at least one by 4. One of $a-2$, $a-1$, a, $a+1$, $a+2$ is divisible by 5. Hence $a(a-2)(a+2)$ is divisible by $2\cdot 2\cdot 4\cdot 3\cdot 5 = 240$.

Similar comments were submitted by Clayton W. Dodge, University of Maine; and L. R. Nyhoff and Paul Boonstra (jointly), Calvin College, Michigan.

Comment on Q536

Q536. [January, 1972] Show that the square roots of three distinct prime numbers cannot be terms of a common geometric progression.

[*Submitted by Murray S. Klamkin*]

Comment by William Wernick, City College of New York.

If three terms are in geometric progression then the product of the first and last must equal the square of the second, thus in this case $\sqrt{ac} = b$ or $b^2 = ac$ which is clearly impossible with distinct primes.

ANSWERS

A547. Observe that

$$\frac{(f-A)(f-B)}{f} \leqq 0$$

in the closed unit interval. Integrating both sides of this inequality from 0 to 1 yields the desired result.

A548. Every countable set of the real line has Lebesgue measure zero. Since both sets T and I have measure one, neither is countable.

A549. Let $\{a_n\}$ be defined by $a_1 = a_2 = 0$; for $3 \leqq n$, $a_n =$ length of one side of a regular polygon of n sides inscribed in a circle of radius 1. Then $na =$ length of the perimeter of the polygon. Then $\{a_n\} \to 0$ and $\{na_n\} \to 2\pi$.

A550. For any curve the radius of curvature is $\rho = (1 + y_1^2)^{3/2}/y_2$ where $y_1 = dy/dx$ and $y_2 = d^2y/dx^2$. If the curve is a circle, $\rho =$ radius. Therefore the $(1 + y_1^2)^{3/2}/y_2 = r$ is the required differential equation.

A551. When viewed modulo 3, the equation asserts that $(-1)^x - 1 = (-1)^y (\text{mod } 3)$ from which it is evident that y must be even. On the other hand, when viewed modulo 5, we obtain $2^y \equiv 2 \pmod 5$ or $2^{y-1} \equiv \pmod 5$. The latter congruence requires that $y - 1$ be divisible by 4 which implies that y is odd. This contradiction yields the desired conclusion.

(Quickies on pages 229-230.)

ANNOUNCEMENT OF LESTER R. FORD AWARDS

At its meeting on January 27, 1965, in Denver, Colorado, the Board of Governors authorized a number of awards, to be named after Lester R. Ford, Sr., to authors of expository articles published in the MONTHLY and the MATHEMATICS MAGAZINE. A maximum of six awards will be made annually; each award is in the amount of $100. The articles are to be selected by a subcommittee of the Committee on Publications appointed for this purpose.

The 1972 recipients of these Awards, selected by a committee consisting of E. F. Beckenbach, Chairman, Marvin Marcus, and D. E. Richmond, were announced by First Vice-President Dorothy Bernstein at the business meeting of the Association on August 29, 1972, at Dartmouth College. The recipients of the Ford Awards for articles published in 1971 were the following:

G. D. Chakerian and L. H. Lange, Geometric Extremum Problems, this MAGAZINE, 44 (1971) 57–69.

P. M. Cohn, Rings of Fractions, MONTHLY, 78 (1971) 596–615.

F. Cunningham, The Kakeya Problem for Simply Connected and for Star-shaped Sets, MONTHLY, 78 (1971) 114–129.

W. J. Ellison, Waring's Problem, MONTHLY, 78 (1971) 10–36.

Leon Henkin, Mathematical Foundations for Mathematics, MONTHLY, 78 (1971) 463–487.

Victor Klee, What is a Convex Set?, MONTHLY, 78 (1971) 616–631.

HENRY L. ALDER, *Secretary*

A548. Every countable set of the real line has Lebesgue measure zero. Since both sets T and I have measure one, neither is countable.

A549. Let $\{a_n\}$ be defined by $a_1 = a_2 = 0$; for $3 \leqq n$, a_n = length of one side of a regular polygon of n sides inscribed in a circle of radius 1. Then na = length of the perimeter of the polygon. Then $\{a_n\} \to 0$ and $\{na_n\} \to 2\pi$.

A550. For any curve the radius of curvature is $\rho = (1 + y_1^2)^{3/2}/y_2$ where $y_1 = dy/dx$ and $y_2 = d^2y/dx^2$. If the curve is a circle, ρ = radius. Therefore the $(1 + y_1^2)^{3/2}/y_2 = r$ is the required differential equation.

A551. When viewed modulo 3, the equation asserts that $(-1)^x - 1 = (-1)^y (\text{mod } 3)$ from which it is evident that y must be even. On the other hand, when viewed modulo 5, we obtain $2^y \equiv 2 \pmod 5$ or $2^{y-1} \equiv (\text{mod } 5)$. The latter congruence requires that $y - 1$ be divisible by 4 which implies that y is odd. This contradiction yields the desired conclusion.

(Quickies on pages 229-230.)

ANNOUNCEMENT OF LESTER R. FORD AWARDS

At its meeting on January 27, 1965, in Denver, Colorado, the Board of Governors authorized a number of awards, to be named after Lester R. Ford, Sr., to authors of expository articles published in the MONTHLY and the MATHEMATICS MAGAZINE. A maximum of six awards will be made annually; each award is in the amount of $100. The articles are to be selected by a subcommittee of the Committee on Publications appointed for this purpose.

The 1972 recipients of these Awards, selected by a committee consisting of E. F. Beckenbach, Chairman, Marvin Marcus, and D. E. Richmond, were announced by First Vice-President Dorothy Bernstein at the business meeting of the Association on August 29, 1972, at Dartmouth College. The recipients of the Ford Awards for articles published in 1971 were the following:

G. D. Chakerian and L. H. Lange, Geometric Extremum Problems, this MAGAZINE, 44 (1971) 57–69.

P. M. Cohn, Rings of Fractions, MONTHLY, 78 (1971) 596–615.

F. Cunningham, The Kakeya Problem for Simply Connected and for Star-shaped Sets, MONTHLY, 78 (1971) 114–129.

W. J. Ellison, Waring's Problem, MONTHLY, 78 (1971) 10–36.

Leon Henkin, Mathematical Foundations for Mathematics, MONTHLY, 78 (1971) 463–487.

Victor Klee, What is a Convex Set?, MONTHLY, 78 (1971) 616–631.

HENRY L. ALDER, *Secretary*

MATHEMATICS MAGAZINE, Vol. 45, No. 4, SEPTEMBER–OCTOBER 1972