ARTICLES BY TITLE

INDEX TO VOLUME 41 JANUARY THROUGH DECEMBER, 1968

ARTICLES BY AUTHOR

BOOK REVIEWS

EDITED BY DMITRI THORO, San Jose State College
BRIEF MENTION 40–41, 210–211, 282–283
Names of authors are in ordinary type; those of reviewers in capitals and small capitals.

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles Valley College

PROPOSALS

SOLUTIONS

Quickies and Answers

The page on which Quickies appear is in the parentheses following the number of these problems: the page on which the Answers appear is in boldface. 420, 421, 422, 423, 424, (50) **(41–42)**; 425, 426, 427, 428, 429, (102) **(94)**; 430, 431, 432, 433, 434, (166) **(133)**; 435, 436, 437, 438, 439, (223–224) **(191)**; 440, 441, 442, 443, (295) **(247-8)**.

MATHEMATICS MAGAZINE

CONTENTS

VOLUME 41 • NOVEMBER 1968 • NUMBER 5

MATHEMATICS MAGAZINE

ROY DUBISCH, *EDITOR*

EDITORIAL CORRESPONDENCE should be sent to the Editor-Elect, S. A. JENNINGS, Department of Mathematics, University of Victoria, Victoria, British Columbia, Canada. Articles should be typewritten and double-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.

The MATHEMATICS MAGAZINE is published by the Mathematical Association of America at Washington, D. C., bi-monthly except July–August. Ordinary subscriptions are: 1 year \$4.00 Members of the Mathematical Association of America and of Mu Alpha Theta may subscribe at the special rate of 2 years for \$6.00. Single copies are 80¢.

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

THE PROBLEM OF THE SHORTEST NETWORK JOINING n POINTS

RICHARD F. DEMAR, University of California, Davis and University of Cincinnati

We consider the following problem:

Given n points in the plane, find the shortest network joining them.

By a network joining n points, we mean a connected set consisting of arcs and containing the given points. The given points will be called *given vertices* of the network and any other point where two or more arcs intersect will be called *added vertices*. An arc whose endpoints are vertices (either given or added) and containing no other vertex will be called an *edge*.

This problem is interesting for several reasons. It is an example of a minimization problem which can best be attacked by elementary geometric methods rather than by calculus. It is unsolved in the sense that no practical algorithm is known for finding the solution in the general case although certain special cases have been considered for centuries. It is also interesting because of its possible applications; e.g., construction of a highway system or power lines joining n towns or a pipeline system joining oil wells to a collection station. Professor S. K. Stein pointed out to the author that it is Plateau's (soap bubble) problem in two dimensions.

In the trivial case $n=2$, the solution is, of course, the line segment joining the points. If $n=3$, the problem is already not entirely trivial. It is equivalent to the following, known as Fermat's problem: Given three points A, B, and C, find the point P which minimizes the sum of the distances AP, BP, and CP. We shall show in a moment how to find P and that P is unique. That this problem is equivalent to the given problem can be shown as follows: Let P be the solution to Fermat's problem and let N be the network consisting of AP, BP, and CP. Let N_1 be the shortest network joining A, B, and C. Then there exists a path p_1 in N_1 from A to B and a path p_2 from A to C. These paths may have points in common. Let Q be the vertex where the paths first separate. Unless the arcs from A to Q, from Q to B, and from Q to C are line segments, then N_1 could be shortened by replacing them by line segments. Then unless $Q=P$, this network is still longer than N. Thus, we obtain a contradiction unless $N_1=N$; so the problems are equivalent.

Fermat's problem has received attention from numerous people through the years and a number of methods have been given for solving it [**5**; **6**; **7**; **9**; **10**]. A particularly simple and beautiful solution was given by J. E. Hofmann in 1929 [**5**; **3**, p. 21]. It is the purpose of this paper to show how the method of Hofmann can be generalized to apply to certain cases of the given problem with $n>3$. This generalization was given by Cockayne [**1**] and was obtained independently by the author. Melzak [**8**] showed that there exists a Euclidean construction for the shortest network in the general case. We show an explicit construction for candidates for shortest network in certain cases. First, we show Hofmann's solution.

We shall use the notation $\overline{AB}$ for the length of the line segment AB. Let A, B, and C be three given points. Let $\angle C$ be the largest angle in ΔABC. Let Q be any

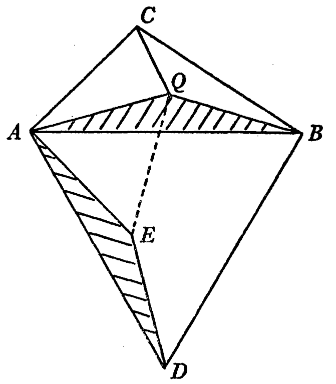


FIG. 1

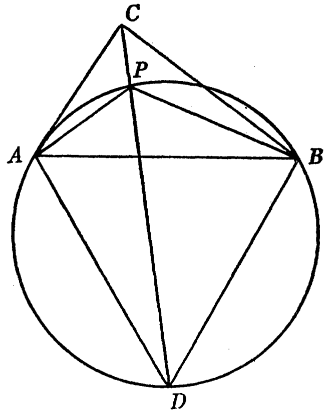


FIG. 2

point (other than A or B). Construct an equilateral triangle on AB with vertex D on the opposite side of AB from C (Figure 1). Construct point E such that $\Delta AED \cong \Delta AQB$, i.e., rotate ΔAQB about A until AB coincides with AD. Then $\angle EAQ = 60°$ and $\overline{AQ} = \overline{AE}$; so ΔAQE is equilateral. Therefore, $\overline{AQ} = \overline{QE}$ and since $\overline{QB} = \overline{ED}$, the sum $\overline{AQ} + \overline{BQ} + \overline{CQ}$ is the length of the polygonal path $CQED$. The choice of the point Q which minimizes the length of this polygonal path is the choice which makes it the line segment CD. Thus P lies on this line segment CD. To find where, circumscribe a circle about ΔABD (Figure 2). Suppose $\angle C < 120°$. Then this circle intersects CD inside ΔABC. Let P be this point of intersection. Draw AP and BP. Then $\angle APB = \frac{1}{2}\overset{\frown}{ADB} = 120°$ and since D bisects $\overset{\frown}{ADB}$, $\angle APD = 60°$ so that $\angle APC = \angle BPC = 120°$. Letting E on CD be such that $\overline{PE} = \overline{AP}$, we again have $\Delta APB \cong \Delta AED$; so $\overline{ED} = \overline{PB}$ and $\overline{CP} + \overline{AP} + \overline{BP} = \overline{CD}$. Thus P is the required point. From the fact that the angles APB, BPC, and CPA are all 120°, it follows that if we had used a different side on which to construct the equilateral triangle, we would have obtained the same point P, so that P is unique. From our construction, if $\angle C = 120°$, then $P = C$.

If $\angle C > 120°$ and Q is any point other than C, extend AC and rotate ΔQCB about the point C until CB lies along the extension of AC, i.e., construct $\Delta CDE \cong \Delta CBQ$ (even if these triangles are degenerate) (Figure 3). Then $\angle QCE = \angle BCD < 60°$ and $\overline{QC} = \overline{CE}$, so that $\overline{QE} < \overline{QC}$. Therefore, since $\overline{DE} = \overline{QB}$, we have $\overline{AQ} + \overline{QC} + \overline{QB} > \overline{AQ} + \overline{QE} + \overline{ED} \geq \overline{AD} = \overline{AC} + \overline{CB}$. Thus $P = C$ in this case also. This solves our problem for $n = 3$.

In the general problem of n points, in order that a network be of minimum length, it is obvious that all edges must be line segments. Also, every added vertex must be an endpoint of exactly three edges intersecting at 120° angles. If an added vertex V were an endpoint of only one edge, the edge could be omitted and if it were an endpoint of only two edges, these two line segments could be replaced by one line segment shortening the network. If it were an endpoint of three or more edges and if any angle formed by a pair of these edges were less than 120°, then this pair of edges is not the shortest network joining these three vertices which implies that the given network could be shortened. This must be the case if V is an endpoint of more than three edges. If a given vertex is an endpoint of more than one edge, these edges must intersect in an angle of at least 120° by the same considerations. Thus we have the following necessary condition in order that a network be of minimum length:

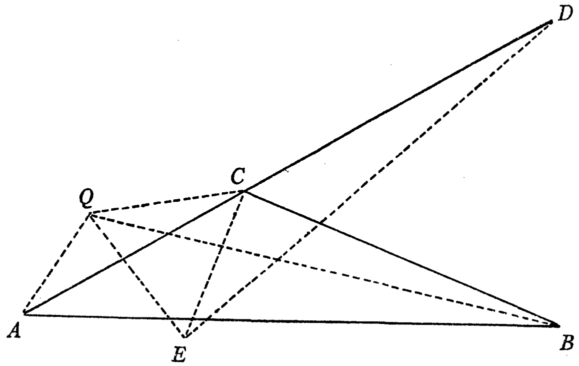


FIG. 3

Condition 1: Every edge is a line segment, every added vertex is an endpoint of exactly three edges which intersect at angles of 120°, and if a given vertex is an endpoint of more than one edge, these edges make an angle of at least 120°.

We also note that the shortest network cannot contain a loop since one edge could then be omitted without disconnecting the network. Such a network is called a tree. By Euler's formula, the number of edges in a tree with m vertices is $m-1$. Suppose that the shortest network joining n points has k added vertices. Each of these added vertices is an endpoint of exactly three edges and at most $k-1$ of these edges join two added vertices, so the total number of edges in the network is at least $3k-(k-1)=2k+1$. All but at most $k-1$ of the edges have one or two given points as endpoints. Therefore $n \geqq (2k+1)-(k-1)$ or $k \leqq n-2$. Thus the number of added vertices never exceeds $n-2$.

Now consider the problem with $n=4$. Let A, B, C, and D be the given points. Choose two pairs of points, say A, B and C, D. Since $n=4$, the number of added vertices is at most 2. Thus, we want to find points P and Q (where one or both might coincide with given vertices) such that $\overline{AP}+\overline{BP}+\overline{PQ}+\overline{QC}+\overline{QD}$ is minimum. Let P' and Q' be any two points, and form the network, AP', BP', $P'Q'$, $Q'D$, $Q'C$ (Figure 4). Construct equilateral triangles ABE and DCG. Then construct $\Delta AP''E \cong \Delta AP'B$ and $\Delta DQ''G \cong \Delta DQ'C$. Then $\overline{AP'}=\overline{AP''}$ and $\angle P'AP''=60°$. Therefore $\Delta AP'P''$ is equilateral and $P''P'=\overline{AP'}$. Similarly $\overline{Q'Q''}=\overline{DQ'}$. Also, we have $\overline{BP'}=\overline{EP''}$ and $\overline{Q''G}=\overline{Q'C}$. Therefore $\overline{AP'}+\overline{BP'}+\overline{P'Q'}+\overline{Q'C}+\overline{Q'D}=\overline{EP''}+\overline{P''P'}+\overline{P'Q'}+\overline{Q'Q''}+\overline{Q''G}$ which is the length of a polygonal path joining E and G. Therefore the shortest network is obtained by choosing P and Q so that this polygonal path coincides with the line segment EG. This can be done by circumscribing circles about the triangles ABE and CDG and taking P and Q to be the points of intersection of these circles with the line segment EG. It is easily verified as before that the length of the resulting network is $\overline{EG}$.

However, the resulting network may not be the one of minimum length; it

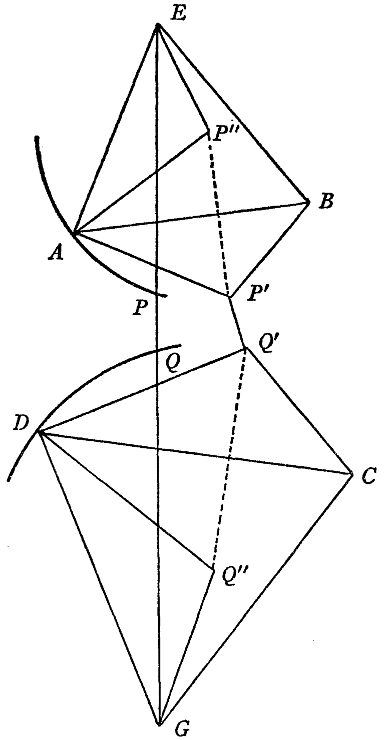


FIG. 4

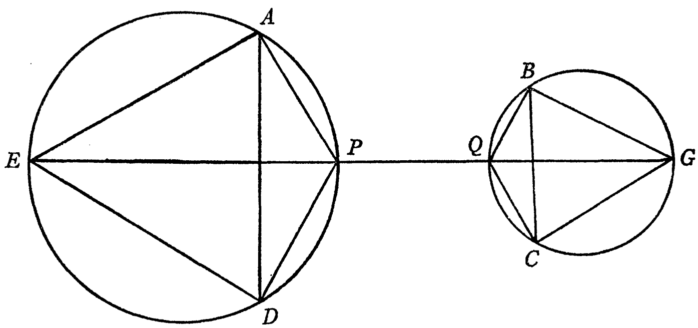


FIG. 5

may be just a local minimum. Remember we paired the points A, B and C, D. We could just as well have paired them A, D and B, C and have made a similar construction. This would have led to a different network also satisfying Condition 1 (Figure 5). Thus all pairings of the given points must be considered and the shortest path obtained for all of these pairings is the solution.

Just as in the case of three points with one angle greater than 120°, for a given pairing of the points the construction may not be possible. For example, with the pairing A, B and C, D, the line segment EG may not intersect AB.

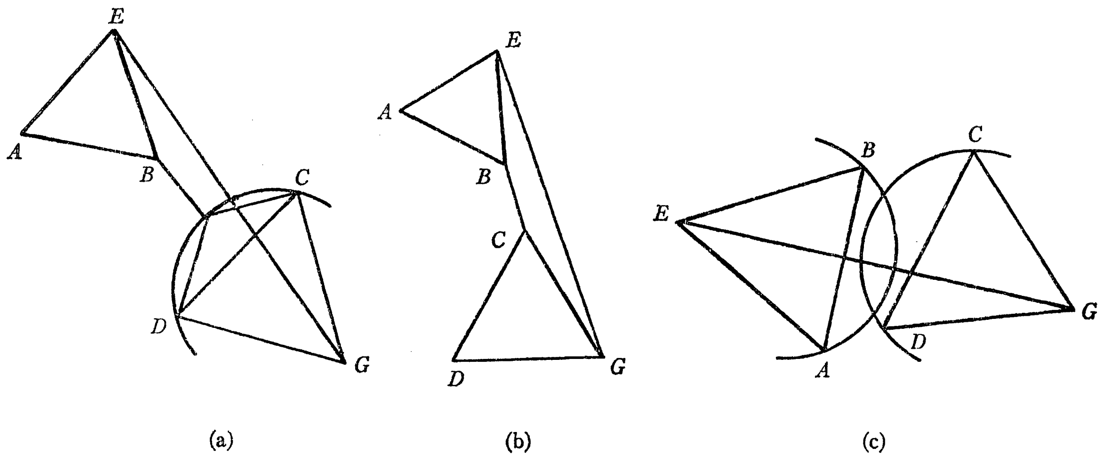


FIG. 6

(Figure 6 (a)). In this case, the shortest network for this pairing is the line segment AB together with the shortest network joining either A or B to C and D. If EG does not intersect either AB or CD (Figure 6 (b)), then the shortest network for the given pairing consists of AB, CD, and the shortest line segment joining these. A third thing which can render the construction impossible is that the intersection of EG with the circumscribed circle about ΔDCG may lie inside the circle about ΔABE (Figure 6 (c)). This is the case, for example, in Figure 5, for the pairing A, C and B, D. In this case there is no network satisfying Condition 1 for this pairing of the points.

The shortest network joining any four points can now be found by determining the length of the network for each pairing of the points for which a network satisfying Condition 1 exists and comparing these. Then the shortest network can be constructed by the method shown.

For $n=5$, we proceed similarly. Let A, B, C, D, and E be the given points. Choose one point, say E, and pair the other four, say A, B and C, D. Construct equilateral triangles ABK and CDL (Figure 7). Construct equilateral triangle KLM. Draw EM. Let P be the intersection of the circumscribed circle about ΔKLM with EM. Draw KP and LP. Let R and T be the intersections of KP and LP with the circumscribed circles about ΔABK and ΔCDL, respectively. Draw AR, BR, CT, and DT. Then the shortest network for the given pairing of points consists of AR, BR, RP, CT, DT, TP, and PE. This can be shown in the same way as before. Its length is $\overline{EM}$. Again, for a given pairing, there may not exist a network satisfying Condition 1, or it may exist but have fewer than three added vertices, so that the construction will not look like Figure 7. However, the network, if it exists, can be found by this method of construction.

The same procedure can be used for any n. A way of looking at what we are doing which makes it easier to use in practice is that each time we pair two points, we replace these two points by one—the third vertex of an equilateral triangle of which the pair form the other two vertices—and then work with the

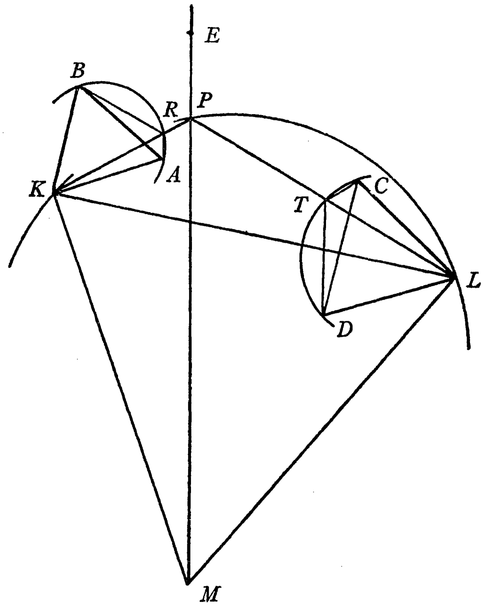


FIG. 7

resulting set which has fewer elements, eventually obtaining a network joining the given points and satisfying Condition 1. However, no way is known of determining which one (or ones) of the networks satisfying Condition 1 gives an absolute minimum length except to find the lengths and compare them. This is impractical except for small n. For example, if $n=7$, Gilbert and Pollak [4] give the number of ways in which networks satisfying Condition 1 might be formed as 62,370. Thus, the problem of a practical algorithm for finding the shortest network is still open. Some progress on this problem has been made by Cockayne and by Gilbert and Pollak. The interested reader should consult these two papers.

That there can be more than one network of minimum length has been shown by having the given points vertices of a square [2, p. 361]. In this example, one network is simply a rotation of the other through 90°. The following example shows that there may be two networks which are quite different, both of which are of minimum length. Let A, B, C, and D be vertices of an isosceles trapezoid with base angles 60°, lower base $\overline{AD}=2\sqrt{3}/3$ and sides of length $2(1-\sqrt{3}/3)$ (Figure 8). Then the upper base BC is of length $2(2\sqrt{3}/3-1)$. Therefore, the network consisting of AB, BC, and CD is of length 2. On the other hand constructing equilateral triangles ADE and BCG on the lower and upper bases, respectively, we have $\overline{AG}=\overline{AE}=\overline{AD}=2\sqrt{3}/3$, so that $(\overline{EG}/2)^2+(\sqrt{3}/3)^2=(2\sqrt{3}/3)^2$, so $\overline{EG}=2$. But the length of the network consisting of BP, CP, PQ, QA, and QD is of length equal to $\overline{EG}=2$. Since these are the only two networks joining these points which satisfy Condition 1, they are both of minimum length.

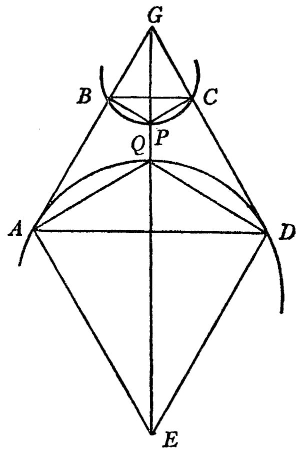


FIG. 8

The author wishes to thank Professor G. D. Chakerian for pointing out a number of references.

References

1. E. J. Cockayne, On the Steiner problem, Canad. Math. Bull., 10 (1967) 431–450.
2. R. Courant and H. Robbins, What is Mathematics?, Oxford University Press, 1948.
3. H. S. M. Coxeter, Introduction to Geometry, Wiley, New York, 1961.
4. E. N. Gilbert and H. O. Pollak, Steiner minimal trees, SIAM Rev., 16 (1968) 1–29.
5. J. E. Hofmann, Elementare Losung einer minimumsaufgabe, Math. Z., 60 (1929) 22–23.
6. D. Jackson, Relating to a minimum problem, Amer. Math. Monthly, 24 (1917) 42.
7. R. Johnson, Relating to a problem in minima discussed by Professor D. Jackson, Amer. Math. Monthly, 24 (1917) 243–244.
8. Z. A. Melzak, On the problem of Steiner, Canad. Math. Bull., 4 (1961) 143–148.
9. J. Steiner, Gesammelte Werke, Vol. 2, Riemer, Berlin, 1882.
10. E. B. Wilson, Relating to infinitesimal methods in geometry, Amer. Math. Monthly, 24 (1917) 241–243.

BOUNDS ON THE LOGARITHMIC DERIVATIVE OF SOLUTIONS OF SECOND ORDER DIFFERENTIAL EQUATIONS

ANDRÉ RONVEAUX, Université de Montréal

1. Introduction. It is well known that the logarithmic derivative $(y'/y) = (1/t)$ of the solution of a second order linear differential homogeneous equation satisfies a first order nonlinear (Riccati) equation, [1]. The aim of this paper is to obtain bounds on the t function, transforming first the Riccati equation. We also illustrate these results, giving approximations on the elementary functions $\tanh x$ and $\tan x$.

2. Derivation of the bounds. Let us consider the differential problem

$$y''(x) - p(x)y(x) = 0 \tag{1}$$

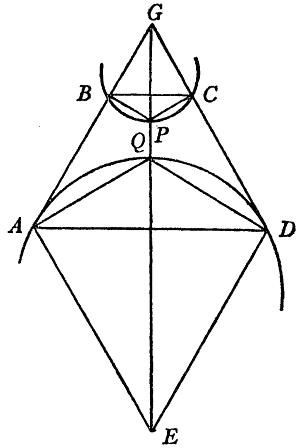


FIG. 8

The author wishes to thank Professor G. D. Chakerian for pointing out a number of references.

References

1. E. J. Cockayne, On the Steiner problem, Canad. Math. Bull., 10 (1967) 431–450.
2. R. Courant and H. Robbins, What is Mathematics?, Oxford University Press, 1948.
3. H. S. M. Coxeter, Introduction to Geometry, Wiley, New York, 1961.
4. E. N. Gilbert and H. O. Pollak, Steiner minimal trees, SIAM Rev., 16 (1968) 1–29.
5. J. E. Hofmann, Elementare Losung einer minimumsaufgabe, Math. Z., 60 (1929) 22–23.
6. D. Jackson, Relating to a minimum problem, Amer. Math. Monthly, 24 (1917) 42.
7. R. Johnson, Relating to a problem in minima discussed by Professor D. Jackson, Amer. Math. Monthly, 24 (1917) 243–244.
8. Z. A. Melzak, On the problem of Steiner, Canad. Math. Bull., 4 (1961) 143–148.
9. J. Steiner, Gesammelte Werke, Vol. 2, Riemer, Berlin, 1882.
10. E. B. Wilson, Relating to infinitesimal methods in geometry, Amer. Math. Monthly, 24 (1917) 241–243.

BOUNDS ON THE LOGARITHMIC DERIVATIVE OF SOLUTIONS OF SECOND ORDER DIFFERENTIAL EQUATIONS

ANDRÉ RONVEAUX, Université de Montréal

1. Introduction. It is well known that the logarithmic derivative $(y'/y) = (1/t)$ of the solution of a second order linear differential homogeneous equation satisfies a first order nonlinear (Riccati) equation, [1]. The aim of this paper is to obtain bounds on the t function, transforming first the Riccati equation. We also illustrate these results, giving approximations on the elementary functions tanh x and tan x.

2. Derivation of the bounds. Let us consider the differential problem

$$y''(x) - p(x)y(x) = 0 \tag{1}$$

with $y(0)=0$ and $y'(0)=\alpha\neq 0$.

We assume that $p(x)$ is a continuous function of x in the interval $[0, \infty)$ and also that $p(x)$ is not negative in that interval.

The inverse of the logarithmic derivative $t=(y/y')$ satisfies

$$t'(x) = 1 - p(x)t^2(x) \tag{2}$$

with $t(0)=0$.

The change of variable $T(x)=(x[x-t(x)])/(t(x))$ gives

$$T'(x) = -\frac{T^2(x)}{x^2} + p(x)x^2 \tag{3}$$

with $T(0)=0$.

The first approximation on $T(x)$ gives, by (3), the upper bound

$$T(x) < \int_0^x p(s)s^2ds = R(x), \tag{4}$$

from which we deduce the lower bound on the function $t(x)$

$$\frac{x^2}{x+R(x)} < t(x). \tag{5}$$

Now, the first iteration in equation (3) gives the lower bound on $T(x)$

$$R(x) - \int_0^x \frac{R^2(s)}{s^2}\,ds < T(x), \tag{6}$$

and for $t(x)$ the upper bound

$$t(x) < \frac{x^2}{x + R(x) - \int_0^x \frac{R^2(s)}{s^2}\,ds}. \tag{7}$$

Summarizing, we conclude that

$$R(x) - \int_0^x \frac{R^2(s)}{s^2}\,ds < T(x) < R(x)$$

and

$$\frac{x^2}{R(x)+x} < t(x) < \frac{x^2}{R(x) + x - \int_0^x \frac{R^2(s)}{s^2}\,ds}. \tag{8}$$

3. Range of validity and comments. It is interesting to compare the bounds given by (8) with the bounds derived directly from equation (2). The same technique as before (first approximation and first iteration), gives immediately from equation (2)

$$x - R(x) < t < x. \tag{9}$$

If we compare this with equation (8) we can see the improvements:

1. Equation (9) gives the lower bound $x-R(x)$ and equation (8) $x/(1+(R(x)/x))$. The lower bound of (8) is valid for all x and the lower bound of (9) is meaningless for $R(x)>x$, because we know *a priori* that $t(x)$ is never negative. (In equation (2), $t(x)=0$ for $x>0$ gives a contradiction for the sign of the derivative.)

From $x-R(x)<x/(1+(R(x)/x))$, we conclude, therefore, that the lower bound of (8) is better than the lower bound of (9).

2. Equation (9) gives $t<x$ for all x and (8) gives

$$t<\frac{x}{1+\dfrac{R(x)}{x}-\dfrac{1}{x}\displaystyle\int_0^x\frac{R^2(s)}{s^2}ds}\cdot$$

It is clear that for x small, the upper bound (8) is better than the upper bound (9). But while the upper bound (8) is only valid for x smaller than the first zero x_1 (different from the origin) of $x+R(x)-\int_0^x(R^2(s)/s^2)ds$, the approximation $t<x$ will become better after $x=x_0$ defined by $R(x_0)=\int_0^{x_0}(R^2(s)/s^2)ds$.

Summarizing:

Lower bound (8) is always better than lower bound (9)
Upper bound (8) is better than upper bound (9) in $x<x_0$
Upper bound (9) is better than upper bound (8) in $x_0<x$
Upper bound (8) is only valid for $x<x_1$.

3. It is obvious that the change of variable in equation (2) does not depend on the sign of $p(x)$. But if $p(x)$ is not positive, the oscillatory character of the solution of equation (1) gives infinities in the solution of Riccati's equation. Therefore, the lower bound is now limited to values of x smaller than the first pole of the solution. From the negative value of the lower bound, the upper bound is not available after iteration, but only one other lower bound.

4. Approximations on functions. We can illustrate the preceding results by particularizing the function $p(x)$.

(1) For $p(x)=1$, we obtain $t=\tanh x$ and the approximations give

$$\frac{3x}{3+x^2}<\tanh x<\frac{3x}{3+x^2-\dfrac{x^4}{15}}\cdot \tag{10}$$

The upper bound is valid for $x<[(15+\sqrt{405})/2]^{1/2}$, but obviously not very good for $x>1+(x^2/3)-(x^4/45)$. The lower bound, valid for all x, is known as the two term's expansion of $\tanh x$ in a continued fraction [2]. It is easy to verify that these bounds are good approximations of the $\tanh x$ function.

(2) For $p(x)=-1$, we obtain $t=\tan x$ and the two lower bounds now give

$$\frac{3x}{3-x^2}<\frac{3x}{3-x^2-\dfrac{x^4}{15}}<\tan x\quad\text{for } x<\frac{\pi}{2}\cdot$$

This approximation is in fact very good because the first zero of $3-x^2-(x^4/15)$ is very close to $\pi/2$.

References

1. H. T. Davis, Introduction to Nonlinear Differential and Integral Equations, Dover, New York, 1962, Chapter 3.
2. M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Nat. Bur. Standards, Washington D. C., 1964.

SOME REMARKS ON MACLAURIN'S CONIC CONSTRUCTION

C. R. WYLIE, JR., University of Utah

Of the various procedures for locating additional points on the conic determined by five given points, Maclaurin's construction is probably the simplest. As illustrated in Figure 1, the construction is carried out as follows: Let P_0, P_1, P_2, P_3, P_4 be five points, no three of which are collinear, and let any three of the points, say P_0, P_1, P_2, be selected. Let Q_3 and Q_3' be the points in which P_1P_3 and P_2P_3 intersect P_0P_2 and P_0P_1, respectively. Similarly, let Q_4 and Q_4' be the points in which P_1P_4 and P_2P_4 intersect P_0P_2 and P_0P_1, respectively. Finally, let V be the intersection of Q_3Q_3' and Q_4Q_4', let Q be the intersection of P_0P_2 and an arbitrary line, l, on P_1, and let Q' be the intersection of VQ and P_0P_1. Then the intersection, P, of P_2Q' and $l(=P_1Q)$ is the second intersection of l and the conic determined by the five given points.

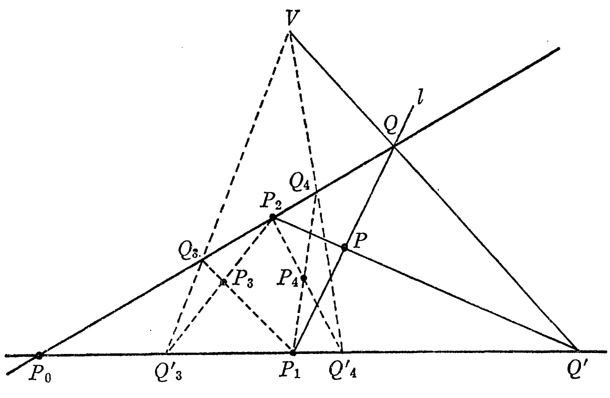


FIG. 1

Although Maclaurin's construction is really a theorem of projective geometry [1], it leads directly to a number of interesting results in elementary analytic geometry, and it is the purpose of this note to indicate some of these. In particular, we shall obtain a surprisingly simple, yet apparently new, geometrical criterion for the type of conic determined by five given points of the euclidean plane.

This approximation is in fact very good because the first zero of $3-x^2-(x^4/15)$ is very close to $\pi/2$.

References

1. H. T. Davis, Introduction to Nonlinear Differential and Integral Equations, Dover, New York, 1962, Chapter 3.

2. M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Nat. Bur. Standards, Washington D. C., 1964.

SOME REMARKS ON MACLAURIN'S CONIC CONSTRUCTION

C. R. WYLIE, JR., University of Utah

Of the various procedures for locating additional points on the conic determined by five given points, Maclaurin's construction is probably the simplest. As illustrated in Figure 1, the construction is carried out as follows: Let P_0, P_1, P_2, P_3, P_4 be five points, no three of which are collinear, and let any three of the points, say P_0, P_1, P_2, be selected. Let Q_3 and Q_3' be the points in which P_1P_3 and P_2P_3 intersect P_0P_2 and P_0P_1, respectively. Similarly, let Q_4 and Q_4' be the points in which P_1P_4 and P_2P_4 intersect P_0P_2 and P_0P_1, respectively. Finally, let V be the intersection of Q_3Q_3' and Q_4Q_4', let Q be the intersection of P_0P_2 and an arbitrary line, l, on P_1, and let Q' be the intersection of VQ and P_0P_1. Then the intersection, P, of P_2Q' and $l(=P_1Q)$ is the second intersection of l and the conic determined by the five given points.

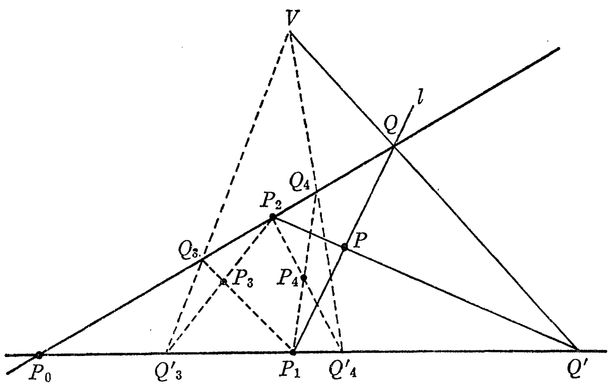


FIG. 1

Although Maclaurin's construction is really a theorem of projective geometry [1], it leads directly to a number of interesting results in elementary analytic geometry, and it is the purpose of this note to indicate some of these. In particular, we shall obtain a surprisingly simple, yet apparently new, geometrical criterion for the type of conic determined by five given points of the euclidean plane.

Let a rectangular coordinate system be chosen so that P_0 is the origin, P_1 is the point $(a, 0)$, and P_2 is the point (b, c). In this coordinate system let the coordinates of V be (u, v). Given an arbitrary point P: (x, y), it is then a simple matter to find the coordinates of the associated points Q and Q'. The condition that P is a point of the locus is then the condition that Q and Q' be collinear with V, which an easy calculation shows to be

$$c^2vx^2 + (ac^2 - 2bcv)xy + (b^2v + acu - abv - abc)y^2 - ac^2vx + (2abcv - ac^2u)y = 0. \tag{1}$$

The conic defined by (1) will be an ellipse, a parabola, or a hyperbola according as the discriminant of its second degree terms, namely,

$$ac^2(ac^2 - 4cuv + 4bv^2) \tag{2}$$

is less than, equal to, or greater than zero. With u and v regarded as variables, the equation obtained by setting this discriminant equal to zero defines the hyperbola

$$4bv^2 - 4cuv + ac^2 = 0 \tag{H}$$

whose asymptotes are the lines P_0P_1 and P_0P_2. Since it is easy to verify that the line P_1P_2 is tangent to H at the point $[(a+b)/2, c/2]$, it follows that in all cases H is located so that one of its branches lies in the interior of $\angle P_1P_0P_2$. Moreover, H can easily be constructed, point-by-point, by the following simple procedure: if R is an arbitrary point on the asymptote P_0P_1 and if S is the point on the asymptote P_0P_2 such that

$$(P_0S)(P_0R) = (P_0P_1)(P_0P_2) = a\sqrt{b^2 + c^2},$$

then the segment RS is tangent to H at the midpoint of the segment. The hyperbola associated with a particular set of points P_0, P_1, P_2 is shown in Figure 2.

It is clear from (2) that when V: (u, v) is outside H the discriminant is positive and when V: (u, v) is inside H the discriminant is negative. Hence we have obtained one possible criterion for the nature of the conic determined by five general points:

If P_0, P_1, P_2, P_3, P_4 are five points in the euclidean plane, no three of which are collinear, and if V is determined from P_0, P_1, P_2, P_3, P_4 as described above, then the conic determined by these points is an ellipse, a parabola, or a hyperbola according as V lies inside, on, or outside the hyperbola, H, determined by P_0, P_1, P_2.

An even simpler criterion that does not explicitly involve the hyperbola H can be obtained as follows: let the measure of $\angle P_1P_0P_2$ be 2γ. Then the vertices of H are the intersections of H and the line $v = (\tan \gamma)u$, namely

$$(\pm\sqrt{ad}\cos^2\gamma,\ \pm\sqrt{ad}\sin\gamma\cos\gamma)$$

where $d = (P_0P_2) = \sqrt{(b^2+c^2)}$. Hence the length of the semitransverse and semiconjugate axes of H are $\sqrt{ad}\cos\gamma$ and $\sqrt{ad}\sin\gamma$, respectively. It follows that the eccentricity of H is $\sec\gamma$ and that the distance from the center of H to either

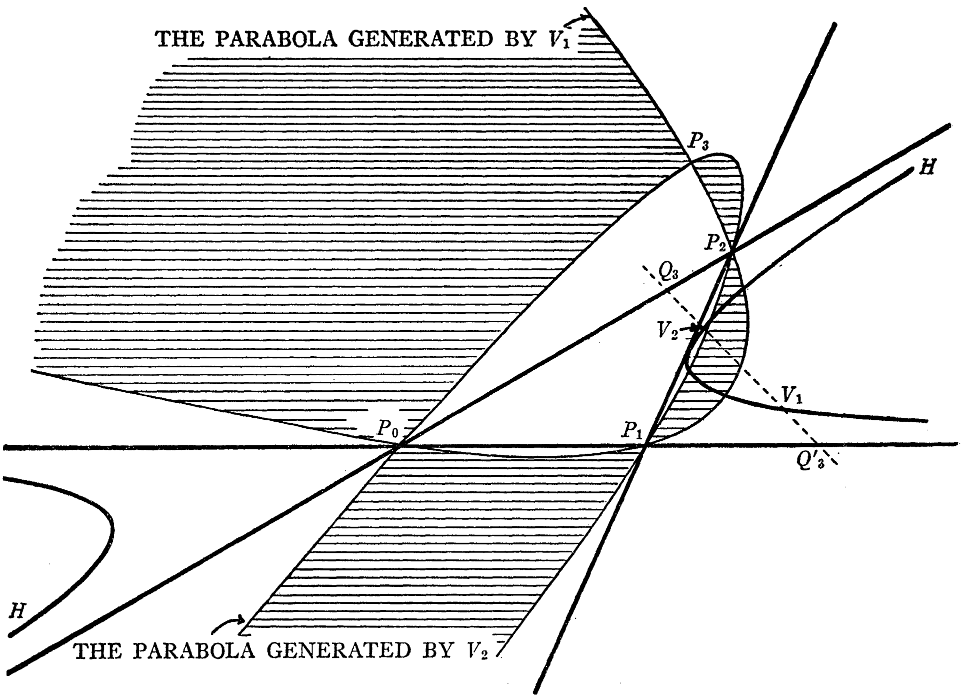


FIG. 2

of the foci is $\sqrt{ad}$. Since $\sqrt{ad}$ can easily be found, using the familiar construction for the mean proportional between two given lengths, and since $\sqrt{ad}\cos\gamma$ can then be found by projecting $\sqrt{ad}$ through the known angle γ, it follows that the vertices, A and A', and the foci, F and F', of H can be located when P_0, P_1, and P_2 are given. Hence, applying the focal-distance definition of a hyperbola to our first criterion, we have a second criterion for the nature of the conic on five given points:

If P_0, P_1, P_2, P_3, P_4 are five points in the euclidean plane, no three of which are collinear, and if the points V, A, A', F, F' are determined from P_0, P_1, P_2, P_3, P_4 as described above, then the conic determined by these five points is an ellipse, a parabola, or a hyperbola according as

$$|(VF) - (VF')|$$

is greater than, equal to, or less than the distance (AA').

Since the nature of the conic on five given points is fixed by the points as a set, independent of any ordering, each of our criteria must hold independently of which of the points are chosen to play the roles of P_0, P_1, P_2. Taken by itself, this seems a rather a surprising property of a general set of five points.

Suppose now that at first we are given only the three points P_0, P_1, P_2. From what has been said, it is clear that these points suffice to fix uniquely a hyperbola, H, which can be used to determine the nature of the conic passing through P_0,

P_1, P_2 and any two additional points. If, however, we are given only one additional point, P_3, a unique conic is, of course, not determined since the point V is not fixed but is only restricted to the line Q_3Q_3' associated with P_3 (Figure 1). If the line Q_3Q_3' fails to intersect H, then among the conics of the pencil on P_0, P_1, P_2, P_3 there will be no ellipses and no parabolas. If the line Q_3Q_3' is tangent to H, then among the conics on P_0, P_1, P_2, P_3 there will be no ellipses but there will be one and only one parabola. Finally, if Q_3Q_3' intersects H in distinct points, the pencil of conics on P_0, P_1, P_2, P_3 will contain both ellipses and hyperbolas, as well as two different parabolas.

The line P_1P_2 is clearly a tangent to H. Also, the lines P_0P_1 and P_0P_2, being asymptotes of H are to be considered as being tangent to H 'at infinity.' Moreover it is easy to see from the definition of Maclaurin's construction, that if P_3 is on one of these lines, then Q_3Q_3' is that same line. (Of course, if this happens, three of the given points are collinear and any conic on the four points is composite. Thus P_0P_1, P_0P_2, and P_1P_2 together form the locus of points V for which Maclaurin's construction leads to a composite conic. This of course can be confirmed independently by considering the locus obtained by equating to zero the discriminant of the entire left member of (1). The three singular conics in the pencil on P_0, P_1, P_2, P_3 thus arise from the points V_1, V_2, V_3 in which the line Q_3Q_3' associated with P_3 meets P_0P_1, P_0P_2, and P_1P_2.) Furthermore, the points of the lines P_0P_1, P_0P_2, P_1P_2 are the only points in the euclidean plane for which the associated line Q_3Q_3' is tangent to H. In fact, if Q_3Q_3' is a tangent to H distinct from P_0P_1, P_0P_2, P_1P_2, then from the property used above to construct H it follows that

$$(P_0P_1)(P_0P_2) = (P_0Q_3)(P_0Q_3')$$

or

$$\frac{(P_0P_1)}{(P_0Q_3)} = \frac{(P_0Q_3')}{(P_0P_2)}$$

which shows that the lines P_1Q_3 and P_2Q_3', whose intersection should be the point P_3, are actually parallel. Thus, corresponding to a general line, Q_3Q_3', which is tangent to H, there is no point in the euclidean plane, but only a 'point at infinity.' Therefore: *in the euclidean plane there exists no set of four points with the property that just one parabola can be passed through them.*

The three lines P_0P_1, P_0P_2, and P_1P_2 divide the plane into seven regions. Four of these, the interior of $\Delta P_0P_1P_2$ and the three sectors vertical to the angles of $\Delta P_0P_1P_2$, have the property that if P_3 is chosen in one of them, then the associated line, Q_3Q_3', does not intersect H. The other three regions have the property that if P_3 is chosen in one of them, then the line Q_3Q_3' meets H in two points. In other words, if P_3 is chosen in one of the first four regions, the pencil of conics on P_0, P_1, P_2, P_3 contains only hyperbolas; but if P_3 is chosen in one of the other three regions, the pencil of conics on P_0, P_1, P_2, P_3 will contain both ellipses and hyperbolas, as well as two parabolas.

If, given P_0, P_1, P_2, a fourth point P_3 is chosen in one of the three regions where curves of all three types can be passed through the points P_0, P_1, P_2, P_3, the two parabolas which contain these four points can easily be constructed by Maclaurin's construction using, in turn, each of the points V_1 and V_2, in which the line Q_3Q_3' associated with P_3 intersects the hyperbola H. Obviously no point of either of these parabolas can lie within any of the four regions of the plane where a fourth point can lead only to hyperbolas. The other three regions are divided by the parabolas into subregions such that the unique conic through P_0, P_1, P_2, P_3 and a fifth point in a subregion lying outside one parabola and inside the other is always an ellipse while the unique conic through P_0, P_1, P_2, P_3 and a fifth point not in such a subregion is always a hyperbola. These observations are illustrated in Figure 2.

It is interesting to note that the direction of the axis of the parabola corresponding to a general point V, on the hyperbola H can easily be constructed. In fact, when the equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{3}$$

represents a parabola, the slope of its axis is $m = -(2A/B)$. Hence from (1), when V: (u, v) is on H and the associated conic is a parabola, the slope of its axis is

$$m = \frac{v}{v \cot 2\gamma - a/2}$$

where, as above, 2γ is the measure of $\angle P_1P_0P_2$. Thus: *the axis of the parabola corresponding to a general point, V, on H is parallel to the line joining the midpoint of the segment P_0P_1 to the point in which the line through V parallel to P_0P_1 meets the line P_0P_2.* An equivalent statement can, of course, be obtained by interchanging P_0 and P_1.

Using this property, we can give elementary proofs of several other interesting results. Since we now know the slope of the axis of the parabola corresponding to a general point, V of H, the tangent of the angle ϕ, between the axes of the parabolas corresponding to two points of H, say V_1: (u_1, v_1) and V_2: (u_2, v_2) can easily be determined, and we have

$$\tan \phi = \frac{2ac^2(v_1 - v_2)}{4d^2v_1v_2 - 2abc(v_1 + v_2) + a^2c^2}\cdot \tag{4}$$

Now, given a general point P: (x, y), the equation of the associated line QQ' can be determined without difficulty, and then the ordinates v_1 and v_2, of the intersections of this line and the hyperbola H can be found. Substituting these into (4), we obtain the equation of the locus of a point P which varies so that the axes of the two parabolas in the pencil of conics on P_0, P_1, P_2, and P_3 intersect at an angle of constant measure:

$$[c(x^2 + y^2) - acx + (ab - d^2)y]^2 \tan^2 \phi = 4cy(by - cx)(by - cx - ay + ac). \tag{5}$$

For all values of ϕ except 0 and $\pi/2$ this is a quartic curve tangent to the ideal line at the isotropic points I and J, and having nodes at P_0, P_1, and P_2. These curves are examples of trefoils, and all have the appearance of the particular curve shown in Figure 3. Moreover, any trefoil which touches the ideal line at I and J has an equation of the form (5). Hence: *through the three nodes and any fourth point P, on any trefoil which is tangent to the ideal line at I and J, two parabolas can be passed, and the axes of these parabolas intersect at an angle whose measure is constant for all points, P on the curve.*

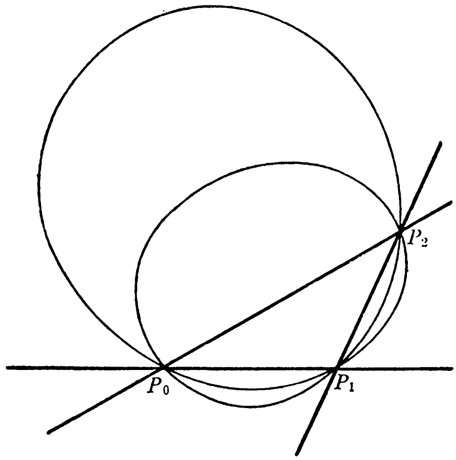


FIG. 3

If $\phi=0$, the quartic (5) reduces to the lines P_0P_1, P_0P_2, P_1P_2, and the ideal line. In this case, as we saw above, there is only one parabola in the pencil on P_0, P_1, P_2, and a fourth point, P, on the quadric. If P is on one of the lines P_0P_1, P_0P_2, or P_1P_2, the parabola is composite. If P is on the ideal line, but not on one of the other three lines, the parabola is proper. If $\phi=\pi/2$, the quartic (5) reduces to the circle on P_0, P_1, P_2 counted twice. Hence we have the following interesting result: *through any four points on a circle, two parabolas can be passed, and the axes of these parabolas are always perpendicular.*

It is interesting now to recall several properties of the general conic, Γ, defined by (3) and apply them to the conic (1) generated by Maclaurin's construction. In the first place, if Γ is an ellipse of eccentricity e whose semimajor and semiminor axes are α and β and if $\theta=\text{Tan}^{-1}\beta/\alpha$, then, as is well known,

$$\sin 2\theta = \frac{\sqrt{4AC - B^2}}{|A + C|} \quad \text{and} \quad e = \sqrt{1 - \tan^2\theta}.$$

On the other hand, if Γ is a hyperbola and if θ is the measure of the acute angles formed by its asymptotes, then

$$\tan 2\theta = \frac{\sqrt{B^2-4AC}}{|A + C|}$$

and the eccentricity of Γ is equal to $\sec\theta$ if Γ lies in the interior of the acute angles formed by its asymptotes, and is equal to $\csc\theta$ if Γ lies in the interior of the obtuse angles formed by its asymptotes.

Applying these considerations to the conic defined by (1), we find that the loci of points V: (u, v) which generate conics for which θ is a constant are the conics of the pencil

$$\lambda[(d^2 - ab)v + acu - abc]^2 = ac^2(ac^2 - 4cuv + 4bv^2) \tag{6}$$

for which $\lambda = -\sin^2 2\theta$ if $B^2 - 4AC \leqq 0$ and $\lambda = \tan^2 2\theta$ if $B^2 - 4AC \geqq 0$. The base members of the pencil are the hyperbola H, corresponding to

$$\lambda = 0, \quad \theta = 0, \quad \text{and} \quad e = 1$$

and the repeated line

$$l\colon [(d^2 - ab)v + acu - abc]^2 = 0$$

corresponding to

$$\lambda = \infty, \quad \theta = \frac{\pi}{4}, \quad \text{and} \quad e = \sqrt{2}.$$

If $\lambda < -1$, the conics of the pencil (6) are entirely imaginary, since the existence of a real point, V, on such a conic would imply that V generated a conic for which $\sin^2 2\theta > 1$. If $\lambda = -1$, so that $\theta = \pi/4$ and $e = 0$, the corresponding conic reduces to the single real point

$$V_C = \left(\frac{d^2 + ab}{2b}, \frac{ac}{2b}\right)$$

which is the point V which generates the unique circle on P_0, P_1, P_2. If $-1 < \lambda < -(c^2/d^2)$, the corresponding curves of the pencil (6) are ellipses, if $\lambda = -(c^2/d^2)$, the corresponding curve is a parabola; and if $-(c^2/d^2) < \lambda < 0$, the corresponding curves are hyperbolas. In each case, however, they are the loci of points, V, which generate ellipses of constant θ and hence constant e. If $\lambda = 0$, the corresponding curve, H, is the locus of points, V, which generate parabolas. If $\lambda > 0$, the corresponding curves are all hyperbolas and are the loci of points, V, which generate hyperbolas for which θ is constant. However, unlike the curves for which $\lambda < 0$, the hyperbolas for which $\lambda > 0$ are not the loci of points, V, which generate conics of constant eccentricity, although they do consist of arcs which have this property, the endpoints of these arcs being the points where the hyperbolas cross the lines P_0P_1, P_0P_2, and P_1P_2. From the formulas for the center of the general conic, Γ, namely

$$h = \frac{2CD - BE}{B^2 - 4AC}, \qquad k = \frac{2AE - BD}{B^2 - 4AC}$$

it follows that the center of the general number of the pencil (6) lies on the line joining the origin to the point V_C. Typical members of the pencil (6) are shown in Figure 4.

The line l is clearly the locus of points, V, which lead, via Maclaurin's construction, to rectangular hyperbolas. It intersects P_0P_1 in the point Q': $(b, 0)$

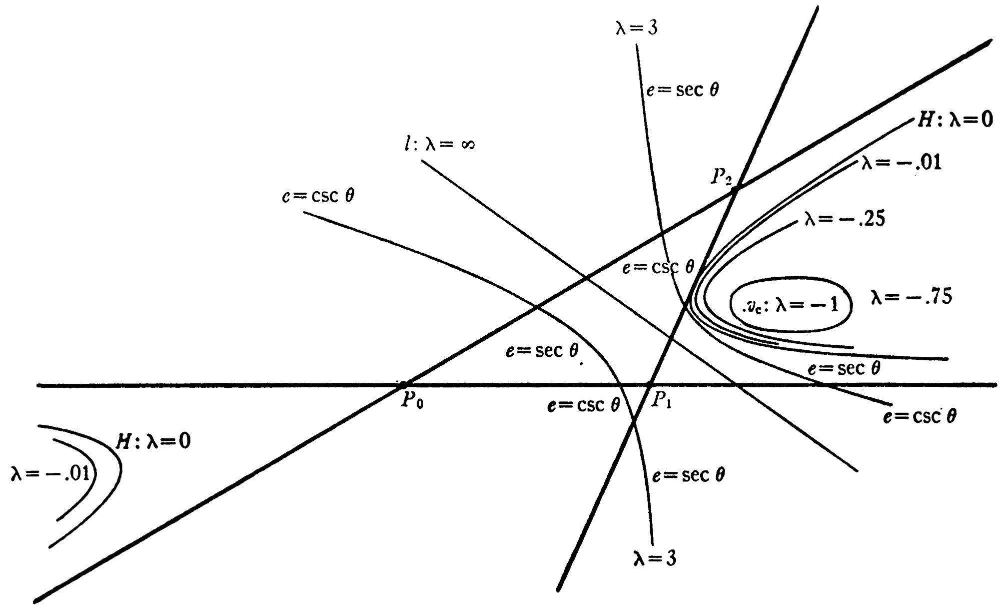


FIG. 4

and P_0P_2 in the point Q: $[(ab^2/d^2), (abc/d^2)]$ which are, respectively, the foot of the perpendicular from P_2 to P_0P_1 and the foot of the perpendicular from P_1 to P_0P_2. Hence the point, P, associated with the line l: QQ', i.e., the fourth point common to all conics generated by the points of l, is the intersection of two of the altitudes of $\Delta P_0P_1P_2$. Since this point must be independent of how the points P_0, P_1, P_2 are labeled, our observations have yielded still another proof of the fact that: *the altitudes of any triangle are concurrent.* Moreover, these observations also provide us with a new proof of the fact that: *if P is the point of concurrence of the altitudes of any triangle, then every conic on P and the vertices of the triangle is a rectangular hyperbola.*

Finally, from the well known formula for the inclination, δ of the axis of a conic, namely $\tan 2\delta = B/(C-A)$ we have for the conic defined by (1),

$$\tan 2\delta = \frac{ac^2 - 2bcv}{c^2v - b^2v + abv - acu + abc} \tag{7}$$

which, as δ varies, is the equation of the pencil of lines on the point

$$\left(\frac{d^2 + ab}{2b}, \frac{ac}{2b}\right).$$

Since this is the point V_C which generates the circle on P_0, P_1, P_2, i.e., since each point P on this circle corresponds to a unique line, QQ', on V_C, it follows that: *the axes of the conics of the pencil on P_0, P_1, P_2, and a fourth point P are all parallel if and only if P is on the circle determined by P_0, P_1, P_2, and when this is the case* *the directions of the axes of the conics are the (perpendicular) directions of the axes of the two parabolas on P_0, P_1, P_2, P.*

The last observation enables us to construct the directions of the axes of the conic on five given points. For the center of the pencil (7), namely V_C, can easily be constructed as the intersection of the lines parallel to P_0P_2 and P_0P_1 through the points on P_0P_1 and P_0P_2 which project into the midpoints of the segments P_0P_2 and P_0P_1, respectively. Then from the given five points the corresponding point V can be constructed and the line VV_C of the pencil (7) can be drawn. From the equation of the pencil it follows that the horizontal intercept of this line is

$$u = b - c \cot 2\delta$$

which shows that 2δ is the measure of the inclination angle of the line determined by P_2 and the horizontal intercept of the line VV_C.

Incidentally, these observations also allow us to construct the fourth point common to the conic determined by five given points and the circle determined by any three of the points. In fact, the point P corresponding to the line VV_C is the required point.

Reference

1. D. Pedoe, An Introduction to Projective Geometry, Pergamon Press, New York, 1963, p. 186.

A NOTE ON THE NETS OF RATIONALITY

ADAM Z. CZARNECKI, Northeastern Illinois State College

It seems that in some books on projective geometry, such as [1] and [2], a net of rationality is identified with the rational elements of an associative division ring. However, it is not explicitly emphasized that this identification of such a net is independent of the theorem of Pappus, or the various equivalent versions.

A construction of integral points and then of rational points on a projective line with a scale determined by P_0, P_1, P_∞ is given in [2]; the argument used is the statement that on such a line

$$P_x \cdot P_y = P_1 \quad \text{if and only if} \quad H(P_1P_{-1}, P_xP_y). \tag{1}$$

The purpose of this note is to establish the validity of this result (1) *without* the use of Pappus' Theorem.

Thus, it is supposed that there is a projective line l with a scale determined by three distinct fundamental points P_0, P_1, P_∞, with addition and multiplication of points as given in [1] or [2], such that:

(A) The algebra of points P_x on l, with P_∞ deleted, is isomorphic to an associative division ring D.

(B) D is infinite (otherwise the argument, although valid, is pointless). This *the directions of the axes of the conics are the (perpendicular) directions of the axes of the two parabolas on* P_0, P_1, P_2, P.

The last observation enables us to construct the directions of the axes of the conic on five given points. For the center of the pencil (7), namely V_C, can easily be constructed as the intersection of the lines parallel to P_0P_2 and P_0P_1 through the points on P_0P_1 and P_0P_2 which project into the midpoints of the segments P_0P_2 and P_0P_1, respectively. Then from the given five points the corresponding point V can be constructed and the line VV_C of the pencil (7) can be drawn. From the equation of the pencil it follows that the horizontal intercept of this line is

$$u = b - c \cot 2\delta$$

which shows that 2δ is the measure of the inclination angle of the line determined by P_2 and the horizontal intercept of the line VV_C.

Incidentally, these observations also allow us to construct the fourth point common to the conic determined by five given points and the circle determined by any three of the points. In fact, the point P corresponding to the line VV_C is the required point.

Reference

1. D. Pedoe, An Introduction to Projective Geometry, Pergamon Press, New York, 1963, p. 186.

A NOTE ON THE NETS OF RATIONALITY

ADAM Z. CZARNECKI, Northeastern Illinois State College

It seems that in some books on projective geometry, such as [1] and [2], a net of rationality is identified with the rational elements of an associative division ring. However, it is not explicitly emphasized that this identification of such a net is independent of the theorem of Pappus, or the various equivalent versions.

A construction of integral points and then of rational points on a projective line with a scale determined by P_0, P_1, P_∞ is given in [2]; the argument used is the statement that on such a line

$$P_x \cdot P_y = P_1 \quad \text{if and only if} \quad H(P_1P_{-1}, P_xP_y). \tag{1}$$

The purpose of this note is to establish the validity of this result (1) *without* the use of Pappus' Theorem.

Thus, it is supposed that there is a projective line l with a scale determined by three distinct fundamental points P_0, P_1, P_∞, with addition and multiplication of points as given in [1] or [2], such that:

(A) The algebra of points P_x on l, with P_∞ deleted, is isomorphic to an associative division ring D.

(B) D is infinite (otherwise the argument, although valid, is pointless). This utilizes incidence and extension axioms, and, in case of a plane, Desargues' Theorem; it is *not* assumed that multiplication in D is commutative, since this is equivalent to Pappus' Theorem.

Let $P_x \neq P_0, P_1, P_\infty$ be an arbitrary point of l; let $P_xW'' \cap l_\infty = I_x''$, $P_1I_x'' \cap l_0 = X''$, $A'X'' \cap l = P_y$.

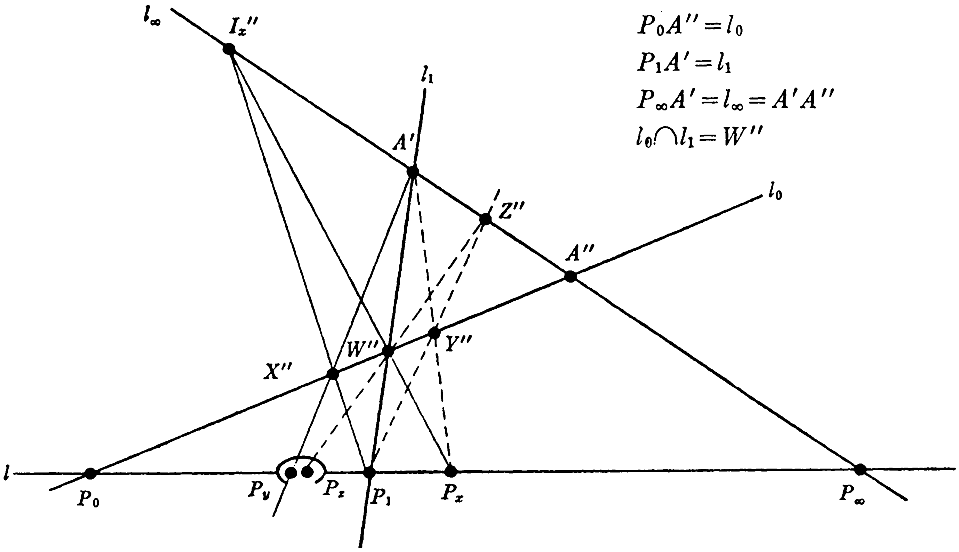


FIG. 1

From the quadrangle $A'W''X''I_x''$ it follows that $Q(P_0P_xP_1, P_\infty P_yP_1)$, and consequently

(a) $$P_x \cdot P_y = P_1 \quad \text{by [1]}.$$

Let $l_0 \cap P_xA' = Y''$, $l_\infty \cap P_1Y'' = Z''$, and $W''Z'' \cap l = P_z$. From the quadrangle $A'W''Z''Y''$ it is deduced that $Q(P_0P_zP_1, P_\infty P_xP_1)$, and hence that

(b) $$P_z \cdot P_x = P_1.$$

By (A), it follows from (a) and (b) that $P_y = P_z = P_{1/x}$, the multiplicative inverse of P_x. In particular, the points

(c) $$P_{1/x}, \quad W'', \quad Z''$$

are collinear. Next

$$\underset{\text{on } l}{P_0P_\infty P_1P_xP_{1/x}} \overset{W''}{\underset{\wedge}{=}} \underset{\text{on } l_\infty}{A''P_\infty A'I_x''Z''} \overset{P_1}{\underset{\wedge}{=}} \underset{\text{on } l_0}{A''P_0W''X''Y''} \overset{A'}{\underset{\wedge}{=}} \underset{\text{on } l}{P_\infty P_0P_1P_{1/x}P_x}.$$

Consequently there is a projectivity T on l given by

$$T\colon P_0P_\infty P_1P_xP_{1/x} \,\overline{\wedge}\, P_\infty P_0P_1P_{1/x}P_x.$$

Evidently $T \neq I$, the identity, since $T(P_0) = P_\infty \neq P_0$, and $T^2(P_x) = T[T(P_x)] = T(P_{1/x}) = P_x$, so that, since P_x is an arbitrary point of l, it follows that T is a *projective involution*.

Now P_1 is evidently an invariant point of T; moreover, for any pair $(P_x, P_{1/x})$ of T, there is another invariant point $P_{x'} \neq P_1$, where $H(P_xP_{1/x}, P_1P_{x'})$. (From the given assumptions it does not necessarily follow that there is a unique other invariant point, since this is a consequence of Pappus' Theorem.) Clearly $P_{x'} \neq P_\infty$, as $T(P_\infty) = P_0$. By (A), it follows that

$$\begin{aligned}(P_{x'} - P_1)\cdot(P_{x'} + P_1) &= (P_{x'} - P_1)\cdot P_{x'} + (P_{x'} - P_1)\cdot P_1 \\ &= P_{x'}\cdot P_{x'} - P_1\cdot P_{x'} + P_{x'}\cdot P_1 - P_1\cdot P_1 \\ &= P_{x'}\cdot P_{x'} - P_{x'} + P_{x'} - P_1 \\ &= P_{x'}\cdot P_{x'} - P_1 = P_0.\end{aligned}$$

since $P_{x'}\cdot P_{x'} = P_1$. As $P_{x'} \neq P_1$, P_∞, it follows that $P_{x'} + P_1 = P_0$, that is $P_{x'} = P_{-1}$. Therefore

(2) The involution $T: P_x \leftrightarrow P_{1/x}$ has exactly two distinct invariant points, namely P_1 and P_{-1}.

(3) For any pair $(P_x, P_{1/x})$ of T it follows that $H(P_xP_{1/x}, P_1P_{-1}) = H(P_1P_{-1}, P_xP_{1/x})$.

Finally, let P_y be a point on l such that $H(P_1P_{-1}, P_xP_y)$.

From the uniqueness of the fourth harmonic point it follows that $P_y = P_{1/x}$. This establishes the required result (1) without the use of Pappus' Theorem, as a consequence of axioms of alignment, extension, and the theorem of Desargues (if the dimension $n = 2$).

References

1. O. Veblen and J. W. Young, Projective Geometry, Ginn, Boston, 1918.
2. G. de B. Robinson, The Foundations of Geometry, University of Toronto Press, 1963.

AN INTERESTING METRIC SPACE

S. K. HILDEBRAND and HAROLD WILLIS MILNES, Texas Technological College

The intent of this paper is to present a metric for the set of elements in the x, y-plane which is uniquely different from the usual metric for the set. Prior to presenting the metric it is necessary to state the following definition.

1. DEFINITION. *ρ is said to be a metric for the set S if and only if ρ is a function from $S \times S$ into R (where R is the space of all real numbers) having the following properties for every x, y, z in S:*

(i) $\rho(x, y) \geqq 0$
(ii) $\rho(x, y) = 0$ *if and only if* $x = y$
(iii) $\rho(x, y) = \rho(y, x)$
(iv) $\rho(x, z) \leqq \rho(x, y) + \rho(y, z)$.

Now P_1 is evidently an invariant point of T; moreover, for any pair $(P_x, P_{1/x})$ of T, there is another invariant point $P_{x'} \neq P_1$, where $H(P_x P_{1/x}, P_1 P_{x'})$. (From the given assumptions it does not necessarily follow that there is a unique other invariant point, since this is a consequence of Pappus' Theorem.) Clearly $P_{x'} \neq P_\infty$, as $T(P_\infty) = P_0$. By (A), it follows that

$$\begin{aligned}(P_{x'} - P_1)\cdot(P_{x'} + P_1) &= (P_{x'} - P_1)\cdot P_{x'} + (P_{x'} - P_1)\cdot P_1\\ &= P_{x'}\cdot P_{x'} - P_1\cdot P_{x'} + P_{x'}\cdot P_1 - P_1\cdot P_1\\ &= P_{x'}\cdot P_{x'} - P_{x'} + P_{x'} - P_1\\ &= P_{x'}\cdot P_{x'} - P_1 = P_0.\end{aligned}$$

since $P_{x'}\cdot P_{x'} = P_1$. As $P_{x'} \neq P_1$, P_∞, it follows that $P_{x'} + P_1 = P_0$, that is $P_{x'} = P_{-1}$. Therefore

(2) The involution $T: P_x \leftrightarrow P_{1/x}$ has exactly two distinct invariant points, namely P_1 and P_{-1}.

(3) For any pair $(P_x, P_{1/x})$ of T it follows that $H(P_x P_{1/x}, P_1 P_{-1}) = H(P_1 P_{-1}, P_x P_{1/x})$.

Finally, let P_y be a point on l such that $H(P_1 P_{-1}, P_x P_y)$.

From the uniqueness of the fourth harmonic point it follows that $P_y = P_{1/x}$. This establishes the required result (1) without the use of Pappus' Theorem, as a consequence of axioms of alignment, extension, and the theorem of Desargues (if the dimension $n = 2$).

References

1. O. Veblen and J. W. Young, Projective Geometry, Ginn, Boston, 1918.
2. G. de B. Robinson, The Foundations of Geometry, University of Toronto Press, 1963.

AN INTERESTING METRIC SPACE

S. K. HILDEBRAND and HAROLD WILLIS MILNES, Texas Technological College

The intent of this paper is to present a metric for the set of elements in the x, y-plane which is uniquely different from the usual metric for the set. Prior to presenting the metric it is necessary to state the following definition.

1. DEFINITION. *ρ is said to be a metric for the set S if and only if ρ is a function from $S \times S$ into R (where R is the space of all real numbers) having the following properties for every x, y, z in S:*

(i) $\rho(x, y) \geqq 0$

(ii) $\rho(x, y) = 0$ *if and only if* $x = y$

(iii) $\rho(x, y) = \rho(y, x)$

(iv) $\rho(x, z) \leqq \rho(x, y) + \rho(y, z)$.

The following definition is presented so that the reader may easily compare the usual metric for the x, y-plane with the distinct one presented in this paper.

2. DEFINITION. *If x and y are elements of the x, y-plane then the usual metric, ρ_1, for the x, y-plane is given by $\rho_1(x, y) = \sqrt{((x_1-x_2)^2+(y_1-y_2)^2)}$ where $x=(x_1, x_2)$ and $y=(y_1, y_2)$.*

The fact that the usual metric for the x, y-plane satisfies Definition 1 and hence is actually a metric is well known and shall not be proved here. The following is a definition of a distinctly different metric.

3. DEFINITION. *If x and y are elements of the x, y-plane, then define ρ_2 as follows, (where $x=(x_1, x_2)$ and $y=(y_1, y_2)$):*

$\rho_2(x, y)=\frac{1}{2}$ if and only if $x_1=y_1$ and $x_2\neq y_2$ or if $x_1\neq y_1$ and $x_2=y_2$,

$\rho_2(x, y)=1$ if and only if $x_1\neq y_1$ and $x_2\neq y_2$

$\rho_2(x, y)=0$ if and only if $x_1=y_1$ and $x_2=y_2$.

4. THEOREM. *ρ_2 as defined in Definition 3 is a metric for the set of elements of the x, y-plane.*

Proof: It will suffice to prove that ρ_2 satisfies the criteria of Definition 1. It is apparent that the range of the function ρ_2 is the set, $\{0, 1/2, 1\}$, which is a subset of R. Therefore it is necessary only to check the additional four properties of Definition 1. Hence, if x, y, z are elements of the x, y-plane, where $x=(x_1, x_2)$, $y=(y_1, y_2)$, and $z=(z_1, z_2)$, then

(i) $\rho_2(x, y)\in\{0, \frac{1}{2}, 1\}$; hence $\rho_2(x, y)\geqq 0$.

(ii) If $\rho_2(x, y)=0$ then $x_1=y_1$ and $x_2=y_2$; hence $x=y$. If $x=y$, then $x_1=y_1$ and $x_2=y_2$; hence $\rho_2(x, y)=0$. Therefore $\rho_2(x, y)=0$ if and only if $x=y$.

(iii) It is apparent from the symmetry of the definition of ρ_2 that $\rho_2(x, y) =\rho_2(y, x)$.

(iv) If $\rho_2(x, z)=0$ then, since $\rho_2(x, y)\geqq 0$ and $\rho_2(y, z)\geqq 0$, $\rho_2(x, z)\leqq\rho_2(x, y) +\rho_2(y, z)$. If $\rho_2(x, z)=\frac{1}{2}$, then either $x_1=z_1$ and $x_2\neq z_2$ or $x_1\neq z_1$ and $x_2=z_2$. Assume $x_1=z_1$ and $x_2\neq z_2$. Therefore either $x_2\neq y_2$ and/or $y_2\neq z_2$. Hence $\rho_2(x, y) \in\{\frac{1}{2}, 1\}$ and/or $\rho_2(y, z)\in\{\frac{1}{2}, 1\}$. In any event it follows, that $\rho_2(x, y)+\rho_2(y, z) \geqq\frac{1}{2}$ and $\rho_2(x, z)\leqq\rho_2(x, y)+\rho_2(y, z)$. The proof is similar if one assumes $x_1\neq z_1$ and $x_2=z_2$. If $\rho_2(x, z)=1$, then $x_1\neq z_1$ and $x_2\neq z_2$. Hence $y_1\neq x_1$ and/or $y_1\neq z_1$ and also $y_2\neq x_2$ and/or $y_2\neq z_2$. It follows that $\rho_2(x, y)\in\{\frac{1}{2}, 1\}$ and $\rho_2(y, z) \in\{\frac{1}{2}, 1\}$. Therefore $\rho_2(x, z)\leqq\rho_2(x, y)+\rho_2(y, z)$.

Therefore ρ_2 is a metric for the set consisting of the points of the x, y-plane.

Now to examine the unique properties of ρ_2. Suppose we have an ordinary road map on which the usual coordinates of the x, y-plane have been superimposed with the x-axis parallel to the bottom of the map and let the distance between cities be determined by ρ_2. Suppose that the vehicle in which we are to travel is similar to a rocket whose engine is capable of being started only once prior to being refueled. If our fuel tank does not hold enough fuel to go $\frac{1}{2}$ units distance, then, relative to ρ_2, it is impossible to leave our present location since the distance, relative to the metric ρ_2, between any two distinct points is greater than or equal $\frac{1}{2}$ units. If our fuel tank holds enough fuel to go $\frac{1}{2}$ unit distance but not enough to go 1 unit distance, then we can proceed to any city on the map whose x-coordinate or y-coordinate is the same as our present x-coordinate or present y-coordinate, respectively, without stopping to refuel. Under these same conditions we can proceed to any point on the map by stopping once for fuel at the vertex of the right angle formed by the lines through the x-coordinate of our present position and the y-coordinate of our destination or by stopping for fuel at the vertex of the right angle formed by the lines through the y-coordinate of our present position and the x-coordinate of our destination. However if our fuel tank holds enough fuel to go 1 unit distance we can proceed directly to any point on the map without stopping for fuel.

It is also of interest to notice that distances as determined by the metric ρ_2 are invariant under translations of the coordinate system, but that distances as determined by the metric ρ_2 are not necessarily invariant under rotations.

A further interesting property of ρ_2 relates to the area of geometric figures in the plane. Consider the simple rectangle $ABCD$. A child in public school will assert confidently that the area of $ABCD$ is the product of the lengths of the sides of the figure; that is $AB \times AD$. With reference to ρ_2, the distance from A to B is $\frac{1}{2}$, and that from A to D is also $\frac{1}{2}$. Consequently, the area of $ABCD$ should be $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. Now, rotate the figure ever so slightly, as in the drawing below:

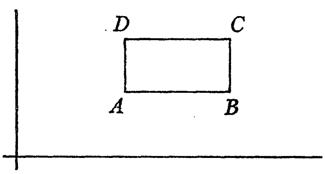


FIG. 1

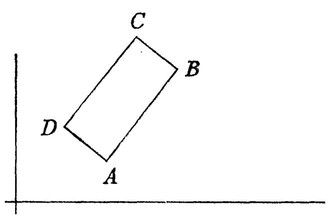


FIG. 2

The distance from A to B is then 1, as is that from A to D; consequently, the area of this figure is $1 \times 1 = 1$. We see that area is no longer an invariant under rotation, since the metric does not have this property.

Returning to the original rectangle, let us adjoin to it a second rectangle $A'B'C'D'$ by translation of $ABCD$ to the right, so that $A'D'$ coincides with BC, as indicated. The area of $A'B'C'D'$ is $\frac{1}{4}$ as before. Consider, however, the area of the combined rectangle $AB'C'D$; the distance from A to B' is $\frac{1}{2}$ and the distance from A to D is $\frac{1}{2}$, so that in accordance with our definition of area, the area of $AB'C'D$ must also be $\frac{1}{4}$. We find that in this metric the area of a whole is not the sum of the area of its contiguous parts, and we conclude that area is not an additive set function. This may seem remarkable, but in reality it is not. It is very well known [1] that proper subsets of the unit square in the plane, with reference to the familiar Euclidean metric, can be constructed such that their area, together with the area of the complement in the square totals to more than the area of the square. These sets are not so simply described as are the rectangles with reference to ρ_2, and are usually considered to be so pathological as to be excluded from consideration entirely as not possessing measure.

A noteworthy property of this metric which is not shared by the Euclidean metric is that the area of a proper subset of a given set may exceed that of the containing set, e.g., as illustrated below:

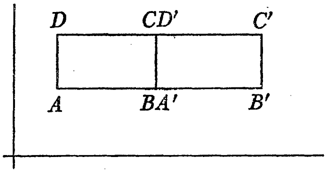


FIG. 3

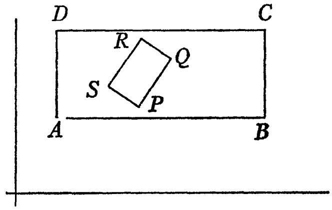


FIG. 4

Here $ABCD$ has area $\frac{1}{4}$ but $PQRS$ has area 1.

It is possible to create a class of measurable sets which preserve the additive property relative to ρ_2 in much the same way that such a class is defined relative to the Euclidean metric. There are indeed an infinite number of ways in which this can be done. We remark first that the area of the null set, and that of any discrete point or any finite collection of points is zero. The areas of rectangles with sides parallel to the x and y axes are $\frac{1}{4}$; the areas of any other rectangles are 1. If we consider then any finite collection of discrete points or rectangles which are not contiguous or overlapping this collection possesses an additive measure. The measure is defined as the sum of the areas of the individual rectangles or points making up the collection.

Reference

1. F. Hausdorff, Grundzüge der Mengenlehre, Leipzig, 1914, pp. 469–472.

ANSWERS

A440. If the 2, 3, and 10 inch doilies have centers at points C, A, and B, respectively, then these points form the vertices of a right triangle with sides 5, 12, and 13 inches. Complete the figure ABC into the rectangle with its fourth vertex denoted by letter O. From O, draw lines through B, C, and A, cutting the given circles in points P, Q, and R. Then $OP=OQ=OR=15$ inches. This is the required radius. For a general solution, see Scripta Mathematica, vol. 21 (1955), pages 46–47.

A441. We have

$$\begin{aligned}\ln(1+x) &= x - x^2/2 + \cdots \\ &= .0004 - (0004)^2/2 + \cdots \\ &= .00040016\end{aligned}$$

Also,

$$\begin{aligned}\ln(1.0004)^\pi &= \pi \ln(1.0004) \\ &= 3.141593(.000400) \\ &= .001257\end{aligned}$$

A442. Suppose Catherine did not share with Descartes a prejudice against negative radii. Then the three quadratics

$$(h-h_1)^2+(k-k_1)^2=(r+r_1)^2$$
$$(h-h_2)^2+(k-k_2)^2=(r+r_2)^2$$
$$(h-h_3)^2+(k-k_3)^2=(r-r_3)^2$$

reduce to three linear equations of the form

$$\begin{array}{l} 2(h_1-h_2)h+2(k_1-k_2)k+2(r_1-r_2)r=(h_1^2-h_2^2)+(k_1^2-k_2^2)-(r_1^2-r_2^2) \\ \vdots \end{array}$$

Now simply substitute all possible combinations of signs for r_1, r_2, and r_3 in

$$\begin{bmatrix} h \\ k \\ r \end{bmatrix}=\frac{1}{2}\begin{bmatrix} h_1-h_2 & k_1-k_2 & r_1-r_2 \\ h_2-h_3 & k_2-k_3 & r_2-r_3 \\ h_3-h_1 & k_3-k_1 & r_3-r_1 \end{bmatrix}^{-1}\begin{bmatrix} (h_1^2-h_2^2)+(k_1^2-k_2^2)-(r_1^2-r_2^2 \\ (h_2^2-h_3^2)+(k_2^2-k_3^2)-(r_2^2-r_3^2) \\ (h_3^2-h_1^2)+(k_3^2-k_1^2)-(r_3^2-r_1^2) \end{bmatrix}$$

to obtain the eight possible circles (h, k), r.

A443. For the base b write the number as

$$\sum_{i=0}^{n} a_i b^{n-i}$$

and denote the permutation of the a_i by a_{ip}. We have for the difference of the two numbers

$$\begin{aligned} \sum_{i=0}^{n} a_i b^{n-i}-\sum_{i=0}^{n} a_{ip} b^{n-i} &= \sum_{i=0}^{n}(a_i-a_{ip})b^{n-i} \\ &= \sum_{i=0}^{n}(a_i-a_{ip})(b^{n-i}-1)+\sum_{i=1}^{n}(a_i-a_{ip}). \end{aligned}$$

The last term is zero and, by induction, $b-1$ divides into the factor $b^{n-i}-1$.

(Quickies on page 295)

Correction for Spectral decomposition of matrices for high school students by Albert Wilansky, this MAGAZINE 41, 2 (1968) 51–59. It has been brought to my attention by Harry D. Ruderman and Paul Rosenbloom that the proof of Theorem 5 is not "immediate from Theorem 1 and the fact that the trace of I is 2." In fact it is conceivable that a resolution of the identity might have $p+2$ members; the trace of the sum would be $p+2$, i.e., 2. Here is a proof of Theorem 5: suppose $(A_1, A_2, A_3, \cdots A_k)$ is a resolution of the identity. Then $(A_1+A_2)^2=A_1^2+A_1A_2+A_2A_1+A_2^2=A_1+A_2$. Hence A_1+A_2 is idempotent. Thus $A_1+A_2=I$, by Theorem 1, since its trace is 2; and so $A_3+A_4+\cdots+A_k=0$. Multiply this equation by A_3, obtaining $A_3=0$. Similarly $A_4=\cdots=A_k=0$.

ON THE RANK OF A MATRIX

V. C. WILLIAMS and F. S. CATER, Portland State College

Let $A=(a_{ij})$ be an $n\times m$ matrix with entries in a field F. Each of the n rows of A can be regarded as a vector in coordinate m-space, F^m, and each of the m columns of A can be regarded as a vector in F^n. The *row space* of A is the subspace of F^m spanned by the rows of A and *column space* of A is the subspace of F^n spanned by the columns of A. The *row rank* of A is the dimension of the row space of A and the *column rank* of A is the dimension of the column space of A. It is proved in most undergraduate courses in linear algebra that the row rank and column rank are equal. This can be done by resorting to the determinants of square minors of A [**3**, pp. 22–24], or by computing the dimension of the solution space of a certain system of linear equations [**2**, pp. 47–51]. Products of linear transformations and the matrices associated with them can also be employed [**1**, pp. 234–235; **5**, pp. 35, 42, 48–50; and **6**, pp. 100–108]. In [**4**] Liebeck gives a short proof valid only for complex scalars. The purpose of this note is to present two very simple proofs that row rank and column rank are equal without resorting to any of these notions. Undergraduates should easily follow our arguments which the authors have not seen in any of the standard texts on linear algebra.

THEOREM. *Let A be an $n\times m$ matrix with entries in a field F, and let $p=$ column rank of A. Then $p=$ row rank of A.*

Proof. Let z_j denote the jth column of A. Fix a subscript k $(1\leqq k\leqq m)$ and choose a scalar d_j for each $j\neq k$. For any scalars $c_1,\cdots,c_n$, the equations

$$\sum_{i=1}^{n} c_i a_{ij} = 0 \qquad (1\leqq j\leqq m) \tag{1}$$

hold if and only if the equations

$$\begin{aligned} &\sum_{i=1}^{n} c_i a_{ij} = 0 \qquad (j\neq k),\\ &\sum_{i=1}^{n} c_i\Big(a_{ik}+\sum_{j\neq k} d_j a_{ij}\Big) = 0 \end{aligned} \tag{2}$$

hold. Equations (1) state that $c_1r_1+\cdots+c_nr_n=0$ where r_i denotes the ith row of A. Equations (2) make the corresponding statement for the matrix formed by replacing the column z_k with the vector $z_k+\sum_{j\neq k} d_jz_j$. It follows that a maximal collection of linearly independent rows in the matrix remains a maximal collection of linearly independent rows when $z_k+\sum_{j\neq k} d_jz_j$ replaces z_k. Hence the row rank is unchanged when we add to one column of A a linear combination of the other columns of A.

We can (and do) reduce A to an $n\times m$ matrix B in which all but p columns are the 0 vector in F^n by adding to certain columns of A linear combinations of other columns of A. From the preceding paragraph it is clear that

$$\text{row rank } A = \text{row rank } B.$$

Finally, the row space of B gives rise, in a natural manner, to a subspace of F^p of the same dimension (just ignore the 0 columns of B) and hence

$$\text{row rank } A = \text{row rank } B \leqq p = \text{column rank } A.$$

The reverse inequality is found by considering A transpose.

Alternate proof. Here we modify somewhat the definitions of the row and column spaces of A. Let $E=\{e_1, \cdots, e_m\}$ be a basis of F^m. The row space of A with respect to E is the subspace of F^m spanned by

$$\sum_{j=1}^{m} a_{1j}e_j, \cdots, \sum_{j=1}^{m} a_{nj}e_j.$$

Then, although two row spaces of A need not coincide, they must be isomorphic. Their common dimension is the row rank of A. Likewise, the column space of A with respect to the basis $\{f_1, \cdots, f_n\}$ of F^n is spanned by

$$\sum_{i=1}^{n} a_{i1}f_i, \cdots, \sum_{i=1}^{n} a_{im}f_i.$$

An elementary row operation leaves invariant the row space with respect to a given basis E, and thus also the row rank.

Suppose V is the row space of a matrix A with respect to a basis E. Let A be transformed into B by an elementary column operation. Then a basis G may be obtained by an appropriate change in E such that V is also the row space of B with respect to G.

1. If the ith column of A is multiplied by a scalar $c \neq 0$, replace e_i in E with $c^{-1}e_i$.

2. If the ith and jth columns of A are interchanged, then interchange e_i and e_j in E.

3. If the ith column of A is replaced by the sum of the ith and jth columns of A, then replace e_j in E with $e_j - e_i$.

In short, row rank is invariant under the elementary operations. An analogous argument shows that column rank is also invariant under elementary operations. By means of such operations A may be reduced to a canonical form in which all entries are 0 except perhaps for several 1's on the main diagonal [6, p. 106]. For such a matrix the row and the column rank both equal the number of such 1's. Since the elementary operations on A leave unchanged the row and column ranks, it is clear that the row and column ranks of A are equal.

References

1. Garrett Birkhoff and Saunders MacLane, A Survey of Modern Algebra, rev. ed., Macmillan, New York, 1953.

2. C. W. Curtis, Linear Algebra, Allyn & Bacon, Boston, 1963.

3. Nathan Jacobson, Lectures in Abstract Algebra, vol. II, Van Nostrand, Princeton, 1953.

4. Hans Liebeck, A proof of the equality of the column and row rank of a matrix, Amer. Math. Monthly, 23 (1966) 1114.

5. E. D. Nering, Linear Algebra and Matrix Theory, Wiley, New York, 1964.

6. M. F. Smiley, Algebra of Matrices, Allyn & Bacon, Boston, 1965.

ON GOLDBERG'S INEQUALITY ASSOCIATED WITH THE MALFATTI PROBLEM

HYMAN GABAI and ERIC LIBAN, York College (CUNY)

1. Introduction. The Malfatti problem, which is discussed in detail in [1], concerns the question of how one may inscribe three nonoverlapping circles in a triangle so that the sum of the areas of the three circles is maximized. Until 1930 (see [2]) it was assumed that the problem had been completely solved in the early part of the 19th century by its proposer, Malfatti. In [1], Goldberg describes a method of constructing the three circles so that the sum of the areas of the circles is greater than the sum of the areas of the circles constructed by Malfatti's method. It is the purpose of this paper to show how one may, with the aid of a computer, verify and strengthen the result asserted in [1].

Consider a given triangle with vertex angles A, B, and C such that $A \leqq B \leqq C$. We assume, without loss of generality, that the inscribed circle has unit radius. Let $u = \tan(A/4)$, $v = \tan(B/4)$,

$$M = \left(\frac{(1+u)(1+v)(1+u+v-uv)}{4(1-uv)}\right)^3 + \left(\frac{(1-uv)(1+u)}{(1+v)(1+u+v-uv)}\right)^2 + \left(\frac{(1-uv)(1+v)}{(1+u)(1+u+v-uv)}\right)^2,$$

$$G_1 = 1 + \left(\frac{1-u}{1+u}\right)^4 + \left(\frac{1-v}{1+v}\right)^4, \qquad G_2 = 1 + \left(\frac{1-u}{1+u}\right)^4 + \left(\frac{1-u}{1+u}\right)^8,$$

and $G = \max(G_1, G_2)$.

It is known (see [1]) that M is the sum of the squares of the radii of the circles constructed by Malfatti's method, and that G is the sum of the squares of the radii of the circles constructed by Goldberg's method. It is asserted in [1] that $M < G$, and it is remarked in [1] that "a rigorous demonstration of this fact would be desirable but has not yet been developed." With the aid of an IBM 1130 computer it has been rigorously demonstrated that not only is it the case that $M < G$ but, moreover, $M < G_1$.

2. The proof. Since $A \leqq B \leqq C$ and $A+B+C = 180°$, $A \leqq 60°$ and $u = \tan(A/4) \leqq \tan 15° = t_1$. Since $2B \leqq 180° - A$, $v = \tan(B/4) \leqq \tan(22.5° - (A/8)) < \tan 22.5° = t_2$. Moreover, since $A+B \leqq 120°$,

$$\frac{u+v}{1-uv} = \tan\left(\frac{A}{4} + \frac{B}{4}\right) \leqq \tan 30° = t_3.$$

One must therefore show that $M < G_1$ for all values of u and v satisfying

$$(1) \qquad 0 \leqq u \leqq v \leqq t_1$$

and for all values of u and v satisfying

$$(2) \qquad 0 \leqq u \leqq \frac{t_3 - v}{1 + t_3 v} \quad \text{and} \quad t_1 \leqq v \leqq t_2.$$

We note that if $0 \leqq a \leqq u \leqq b < 1$ and $0 \leqq c \leqq v \leqq d < 1$, then $a+c-ac \leqq u+v-uv \leqq b+d-bd$. For if $u=b-\beta$ and $v=d-\delta$, then $u+v-uv=(b+d-bd)-(\beta-d\beta)-(\delta-b\delta)-\beta\delta \leqq b+d-bd$. Also, if $u=a+\alpha$ and $v=c+\gamma$, then $u+v-uv=(a+c-ac)+(\alpha-\alpha c)+(\gamma-(a+\alpha)\gamma)>a+c-ac$.

It follows, therefore, that if $0 \leqq a \leqq u \leqq b < 1$ and $0 \leqq c \leqq v \leqq d < 1$, then

$$M \leqq M' = \left(\frac{(1+b)(1+d)(1+b+d-bd)}{4(1-bd)}\right)^2 + \left(\frac{(1-ac)(1+b)}{(1+c)(1+a+c-ac)}\right)^2 + \left(\frac{(1-ac)(1+d)}{(1+a)(1+a+c-ac)}\right)^2$$

and

$$G \geqq G_1 \geqq G' = 1 + \left(\frac{1-b}{1+b}\right)^4 + \left(\frac{1-d}{1+d}\right)^4.$$

By a finite number of suitable choices of a, b, c, and d, one may show that if u and v satisfy (1) or (2), then $M' < G'$. In general, when choosing a, b, c, and d, one must choose $b-a$ and $d-c$ positive but sufficiently small. It was found that choosing intervals such that $b-a=d-c=0.01$ is not, in general, small enough. However by choosing intervals of length 0.001 it was found that $M' < G'$ for all values of u and v satisfying (1) or (2). (It is not the case that $M < G_2$ for all values of u and v satisfying (1) or (2). For example, if $u=v=0.25$ then $M > G_2$.)

To illustrate the sequence of computations which the computer was programmed to perform to demonstrate that $M' < G'$ for all values of u and v satisfying (1) or (2), the first few choices of a, b, c, d, and the corresponding values for M' and G' are listed below.

c	d	a	b	M'	G'
0	0.001	0	0.001	2.0670036	2.9840637
0.001	0.002	0	0.001	2.0632556	2.9761591
		0.001	0.002	2.0595124	2.9682545
0.002	0.003	0	0.001	2.0595283	2.9683175
		0.001	0.002	2.05579	2.960413
		0.002	0.003	2.0520723	2.9525712

References

1. Michael Goldberg, On the original Malfatti problem, this MAGAZINE, 40 (1967) 241–247.

2. H. Lob and H. W. Richmond, On the solutions of Malfatti's problem for a triangle, Proc. London Math. Soc., 2, 30 (1930) 287–304.

ON KAPREKAR'S PERIODIC OSCILLATING SERIES

A. P. SHAH and A. M. VAIDYA, Gujarat University, India

D. R. Kaprekar [2] has defined periodic oscillating series as follows. Let h be a positive integer. Take any integer k_1 less than and prime to $10h-1$. The periodic oscillating series of index h (for short pos h) starting from k_1 is the sequence

$$k_1, k_2, k_3, \cdots \tag{1}$$

where for $n \geqq 2$, k_n is defined recursively by:

If $k_{n-1} = 10a_{n-1} + b_{n-1}$, $0 \leqq b_{n-1} \leqq 9$, then $k_n = hb_{n-1} + a_{n-1}$.

Thus the pos 4 starting from 38 is the sequence

$$38, 35, 23, 14, 17, 29, 38, 35, 23, 14, 17, 29, \cdots. \tag{2}$$

Since $k_1 < 10h-1$, it is clear that $a_1 \leqq h-1$ and $b_1 \leqq 9$. Hence $k_2 = hb_1 + a_1 \leqq 9h + h - 1 = 10h - 1$. Indeed we cannot have $a_1 = h-1$ and $b_1 = 9$ at the same time; therefore $k_2 < 10h-1$. The argument can be repeated to show that each $k_n < 10h-1$, so that the series (1) must be periodic.

For example, the first period of pos 7 starting from 26 is

$$\begin{matrix} 26, 44, 32, 17, 50, 05, 35, 38, 59, 68, 62, \\ 20, 02, 14, 29, 65, 41, 11, 08, 56, 47, 53. \end{matrix} \tag{3}$$

Kaprekar has conjectured that the smallest period of the pos h starting from k_1 is independent of k_1 and indeed that it is equal to the index t to which 10 belongs modulo $10h-1$. In this paper, we prove this conjecture and derive an interesting result.

THEOREM 1. *If $k_1, k_2, \cdots$ is a* pos h, *then for $n \geqq 2$, k_n is the remainder when hk_{n-1} is divided by* $10h-1$.

Proof. If $k_{n-1} = 10a_{n-1} + b_{n-1}$, then $k_n = hb_{n-1} + a_{n-1}$. Hence

$$hk_{n-1} - k_n = 10ha_{n-1} + hb_{n-1} - hb_{n-1} - a_{n-1} = a_{n-1}(10h - 1),$$

i.e., $hk_{n-1} = a_{n-1}(10h-1) + k_n$. Since $k_n < 10h-1$, the theorem is proved.

THEOREM 2. [1]. *If $k_1 < 10h-1$ and $(k_1, 10h-1) = 1$, then the decimal expansion of $(k_1/10h-1)$ is a pure recurring one with period t where t is the index to which* 10 *belongs* modulo $10h-1$.

THEOREM 3. *If t is the index to which* 10 *belongs* modulo $10h-1$, *then* $10^{t-1} \equiv h \pmod{10h-1}$.

Proof. If $10^{t-1} \equiv u \pmod{10h-1}$, then since $10^t \equiv 1$, we must have $10u \equiv 1 \pmod{10h-1}$. Now $u = h$ is obviously a solution of the above congruence and as $(10, 10h-1) = 1$, there is no other solution.

THEOREM 4. *If $k_1, k_2, \cdots$ is a* pos h *and* $(k_1/10h-1) = 0 \cdot a_1 a_2 \cdots a_{t-1}\dot{a}_t$, *then* $(k_2/10h-1) = 0 \cdot a_t a_1 a_2 \cdots a_{t-2}\dot{a}_{t-1}$.

Proof. By Theorems 3 and 1, $10^{t-1}k_1 \equiv hk_1 \equiv k_2 \pmod{10h-1}$. Hence

$$\frac{10^{t-1}k_1}{10h-1} \quad \text{and} \quad \frac{k_2}{10h-1}$$

have the same fractional parts. Obviously the fractional part of the former is $0\cdot a_t a_1 a_2 \cdots \dot{a}_{t-1}$, while the latter, being less than 1, has no integral part. This completes the proof of the theorem.

Iterating the result of Theorem 4, we get

THEOREM 5. *For* $2 \leqq r \leqq t$,

$$\frac{k_r}{10h-1} = 0\cdot a_{t-r+2}a_{t-r+3} \cdots a_t a_1 a_2 \cdots \dot{a}_{t-r+1}.$$

THEOREM 6. *A* pos h *has the period* t *where* t *is the index to which* 10 *belongs* modulo $10h-1$.

Proof. We have (Theorem 5),

$$\frac{k_t}{10h-1} = 0\cdot a_2 a_3 \cdots a_t \dot{a}_1.$$

Applying the process of Theorem 4 once more, we get

$$\frac{k_{t+1}}{10h-1} = 0\cdot a_1 a_2 \cdots \dot{a}_t,$$

proving that $k_1 = k_{t+1}$ and establishing the theorem.

Kaprekar has also noticed the following property of a pos h.

THEOREM 7. *Let* $k_1, k_2, \cdots$ *be a pos* h. *If the period* t *of the series is even, say* $t=2m$, *then for* $i=1, 2, 3, \cdots, m$, $(k_i+k_{m+i}, 10\,h-1)>1$.

We note that in the pos 4 (2), $t=6$ and $(k_i+k_{3+i}, 39)=13>1$, and in the pos 7 (3), $t=22$ and $(k_i+k_{11+i}, 69)=23>1$.

Proof. Let

$$\frac{k_i}{10h-1} = 0\cdot a_1 a_2 \cdots \dot{a}_{2m};$$

then

$$\frac{k_{m+i}}{10h-1} = 0\cdot a_{m+1}a_{m+2} \cdots a_{2m}a_1 a_2 \cdots \dot{a}_m.$$

Thus

$$10^m k_i \equiv k_{m+i} \pmod{10h-1}. \tag{4}$$

Now $10^{2m} \equiv 1 \pmod{10h-1}$. If $10^m \equiv 1 \pmod{p}$ for every prime divisor p of $10h-1$, then we would have $10^m \equiv 1 \pmod{10h-1}$, but this is not true. Therefore for some prime divisor p of $10h-1$, $10^m \not\equiv 1 \pmod{p}$, but $(10^m)^2 = 10^{2m} \equiv 1 \pmod{p}$, so that we must have $10^m \equiv -1 \pmod{p}$. Then (4) implies $-k_i \equiv k_{m+i} \pmod{p}$; or $(k_i + k_{m+i},\ 10h-1) \geqq p > 1$.

Finally, we mention the following points.

(1) It can be easily shown that if we start with $k_1 \leqq 10h-1$ such that $(k_1, 10h-1) = d$, then the pos h starting from k_1 will have period t where t is the index to which 10 belongs modulo $(10h-1/d)$.

(2) In all that precedes, the number 10 plays a completely passive role and it may be replaced by any integer $g > 1$. Thus the property of the periodic oscillating series is not confined to the decimal notation and remains true when numbers are written in the scale of any integer $g > 1$.

Note. Since this paper was accepted for publication, the second named author has discovered much simpler proofs of the results of this paper. An outline of these proofs is given below: it is clear from Theorem 1 that for $n \geqq 2$ we have $k_n \equiv hk_{n-1} \pmod{10h-1}$. It follows by iteration that for $n \geqq 1$, $k_{n+1} \equiv h^n k_1 \pmod{10h-1}$. Since each $k_i < 10h-1$, it is obvious that $k_{n+1} = k_1$ if and only if $h^n \equiv 1 \pmod{10h-1}$. Thus the period of the pos h is the index to which h belongs modulo $10h-1$. It is trivially proved that 10 and h have the same order mod $10h-1$. This proves Theorem 6. Theorem 7 is similarly proved.

References

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 4th ed., Oxford University Press, New York, 1960, p. 111.

2. D. R. Kaprekar, The Mathematics of the New Self Numbers (Part 5), Deolali, 1967. p. 15

NOTE ON SUMS OF SQUARES OF CONSECUTIVE ODD INTEGERS

WILLIAM SOLLFREY, The RAND Corporation

In the latest of a series of papers, [**1**, **2**, **3**] Brother U. Alfred has investigated the conditions under which the sum of the squares of N consecutive odd integers can be a square. He derives several theorems to eliminate many values of N, and finds solutions for other values. Eight numbers below 1000 (193, 564, 577, 601, 673, 724, 772, 913) remain unresolved. Also, in his table of solutions there are 17 values for which the solution exceeds 10 digits, and one contains 34 digits. In this note the eight cases will be resolved, and solutions will be presented for the 17 cases, only one of which contains as many as 7 digits. A systematic procedure for finding these solutions will be demonstrated.

First, $N = 601$ falls out immediately by Brother Alfred's Theorem 7, since it is of the form $12m+1$, and $6m+1 = 301$ contains 7, a number of the form $4k+3$. Next, it may be proved by repeated consideration of residues with respect to 4 and 5 that $N = 564$ has no solution. Solutions have been found for all the other unresolved cases.

To illustrate the procedure, consider $N = 996$, the number for which the for some prime divisor p of $10h-1$, $10^m \not\equiv 1 \pmod{p}$, but $(10^m)^2 = 10^{2m} \equiv 1 \pmod{p}$, so that we must have $10^m \equiv -1 \pmod{p}$. Then (4) implies $-k_i \equiv k_{m+i} \pmod{p}$; or $(k_i + k_{m+i},\ 10h-1) \geqq p > 1$.

Finally, we mention the following points.

(1) It can be easily shown that if we start with $k_1 \leqq 10h-1$ such that $(k_1,\ 10h-1) = d$, then the pos h starting from k_1 will have period t where t is the index to which 10 belongs modulo $(10h-1/d)$.

(2) In all that precedes, the number 10 plays a completely passive role and it may be replaced by any integer $g > 1$. Thus the property of the periodic oscillating series is not confined to the decimal notation and remains true when numbers are written in the scale of any integer $g > 1$.

Note. Since this paper was accepted for publication, the second named author has discovered much simpler proofs of the results of this paper. An outline of these proofs is given below: it is clear from Theorem 1 that for $n \geqq 2$ we have $k_n \equiv hk_{n-1} \pmod{10h-1}$. It follows by iteration that for $n \geqq 1$, $k_{n+1} \equiv h^n k_1 \pmod{10h-1}$. Since each $k_i < 10h-1$, it is obvious that $k_{n+1} = k_1$ if and only if $h^n \equiv 1 \pmod{10h-1}$. Thus the period of the pos h is the index to which h belongs modulo $10h-1$. It is trivially proved that 10 and h have the same order mod $10h-1$. This proves Theorem 6. Theorem 7 is similarly proved.

References

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 4th ed., Oxford University Press, New York, 1960, p. 111.
2. D. R. Kaprekar, The Mathematics of the New Self Numbers (Part 5), Deolali, 1967. p. 15

NOTE ON SUMS OF SQUARES OF CONSECUTIVE ODD INTEGERS

WILLIAM SOLLFREY, The RAND Corporation

In the latest of a series of papers, [1, 2, 3] Brother U. Alfred has investigated the conditions under which the sum of the squares of N consecutive odd integers can be a square. He derives several theorems to eliminate many values of N, and finds solutions for other values. Eight numbers below 1000 (193, 564, 577, 601, 673, 724, 772, 913) remain unresolved. Also, in his table of solutions there are 17 values for which the solution exceeds 10 digits, and one contains 34 digits. In this note the eight cases will be resolved, and solutions will be presented for the 17 cases, only one of which contains as many as 7 digits. A systematic procedure for finding these solutions will be demonstrated.

First, $N = 601$ falls out immediately by Brother Alfred's Theorem 7, since it is of the form $12m+1$, and $6m+1 = 301$ contains 7, a number of the form $4k+3$. Next, it may be proved by repeated consideration of residues with respect to 4 and 5 that $N = 564$ has no solution. Solutions have been found for all the other unresolved cases.

To illustrate the procedure, consider $N = 996$, the number for which the analysis has been presented in [3]. The equation to be investigated is

$$996x^2 + 2\cdot 996\cdot 995x + 2\cdot 332\cdot 995\cdot 1991 = z^2. \tag{1}$$

Writing $z = 166z_1$ and multiplying by 3 gives

$$9x^2 + 18\cdot 995x + 6\cdot 995\cdot 1991 = 249z_1^2. \tag{2}$$

Set $3(x+995) = x_1$, transpose the constant, and there results

$$x_1^2 = 249z_1^2 - 3\cdot 995\cdot 997 = 249z_1^2 - 3\cdot 5\cdot 199\cdot 997. \tag{3}$$

The scheme is now to reduce systematically the large constant term on the right, while retaining the general form of (3). First, using a desk calculator and Barlow's Tables of Squares, add 997 repeatedly to 249 until a square is reached. The first such square is $174724 = (418)^2 = 249 + 997\cdot 175$. Subtract both sides of (3) from $(418)^2 z_1^2$ and obtain

$$(418)^2 z_1^2 - x_1^2 = 997(175z_1^2 + 3\cdot 5\cdot 199) = (418z_1 - x_1)(418z_1 + x_1). \tag{4}$$

Hence 997 must be a factor of $418z_1 - x_1$ or $418z_1 + x_1$. In either case, the equation may be written in the form

$$\begin{aligned} z_2(836z_1 - 997z_2) &= 175z_1^2 + 3\cdot 5\cdot 199 \\ 175z_1^2 - 836z_2z_1 + 997z_2^2 + 3\cdot 5\cdot 199 &= 0. \end{aligned} \tag{5}$$

For this quadratic to have a solution in integers, its discriminant must be a perfect square. The solution may thus be written in the form

$$z_1 = \frac{1}{175}(418z_2 \pm x_2) \tag{6}$$

$$x_2^2 = 249z_2^2 - 3\cdot 5^3\cdot 7\cdot 199. \tag{7}$$

Equation (7) is of the same form as (3), but the large factor 997 has been replaced by smaller factors. To eliminate 199, multiply (7) by 4 and then subtract z_2^2 to obtain

$$4x_2^2 - z_2^2 = 199(5z_2^2 - 3\cdot 5^3\cdot 7\cdot 4) = (2x_2 - z_2)(2x_2 + z_2). \tag{8}$$

Again, 199 must be a factor of $2x_2 - z_2$ or $2x_2 + z_2$. This leads to the pair of quadratics

$$5z_2^2 \pm 2z_3z_2 - 199z_3^2 - 3\cdot 5^3\cdot 7\cdot 4 = 0. \tag{9}$$

Again, the discriminant must be a perfect square. Factoring out a 2 from this discriminant we obtain

$$z_2 = \tfrac{1}{5}(\pm z_3 \pm 2x_3) \tag{10}$$

$$x_3^2 = 249z_3^2 + 3\cdot 5^4\cdot 7. \tag{11}$$

This procedure may be continued until the constant term becomes a perfect square. However, in this particular case the procedure may be stopped at this point.

Now begin expanding $\sqrt{249}$ as a continued fraction. After the fifth step, the remainder 21 is encountered. It is thus quickly found that a solution of (11) is $z_3=5^2\cdot 50$, $x_3=5^2\cdot 789$. Thus we have found two solutions for (7): $z_2=5\cdot 1528$, $x_2=5\cdot 24{,}111$; and $z_2=5\cdot 1628$, $x_2=5\cdot 25{,}689$. When the first set is substituted into (6), neither sign possibility yields an integer solution. However, the second set yields the solution $z_1=18{,}709$. Inserting this into (3) gives $x_1=295{,}218$, whence $x=97{,}411$. For comparison, Brother Alfred's solution contains 18 digits.

The same procedure may be applied to the six unresolved values and the 17 very large solutions. The equation corresponding to (3) can always be reduced to the form

$$p^2 = Nq^2 + r^2 \tag{12}$$

where N denotes the square-free factor of the original number. The value of r will generally be composite. Most often, though not always, the solution $q=0$ will lead to a negative value of x when the reduction is retraced. Now expand $\sqrt{N}$ by continued fractions. Usually, after a few steps, a remainder will occur which is either a squared factor of r^2 or a factor of r. The first possibility gives a solution of (12) directly. The second possibility yields a solution as follows. Suppose p_1, q_1 is a solution of

$$p_1^2 = Nq_1^2 \pm c \tag{13}$$

where $2r/c$ is an integer. Then a solution of (12) is given by

$$p = 2rp_1^2/c \mp r \tag{14}$$

$$q = 2rp_1q_1/c. \tag{15}$$

Also, a remainder may occur which appears earlier in the reduction procedure, as for $N=996$.

The reduction procedure must be retraced and, hopefully, integers will result at each step. If not, or if the final value obtained for x is even, the continued fraction must be carried further. The longest case, $N=193$, required 9 steps. In contrast, the solution of $p^2=Nq^2+1$, as suggested in [3] requires going through at least a full cycle of the periodic continued fraction, whence the very large values.

The solutions we have obtained are listed in Table 1. The second column of this table gives the number of digits in x in the corresponding solution of [3], with a dash in this column denoting a previously unresolved case.

The very great reductions in the size of the solutions are evident from the table. In the most extreme case, $N=649$, x has been reduced by a factor of about $2\cdot 10^{29}$. The same technique can be used to reduce the solutions in the original problem [1] involving sums of squares of consecutive integers.

Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of the RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors.

TABLE 1

N	Digits in Previous x	x	z
73	15	67	1,241
177	13	103	3,953
193	—	4,353,047	60,477,129
249	17	3,289	55,859
292	10	110,311	1,889,970
337	23	1,027	25,275
388	11	271,727	5,360,026
393	18	1,211	32,095
409	26	54,863	1,117,797
537	20	12,955	312,713
577	—	2,263	68,663
628	17	1,281	48,670
649	34	5,327	152,515
673	—	46,671	1,228,225
708	13	134,387	3,594,634
724	—	324,545	8,752,074
753	27	18,717	534,379
772	—	343,803	9,573,958
852	14	105	31,382
913	—	339,107	10,273,989
964	19	13,355	444,886
976	13	251,725	7,894,620
996	18	97,411	3,105,694

References

1. Brother U. Alfred, Consecutive integers whose sum of squares is a perfect square, this MAGAZINE, 37 (1964) 19–32.

2. S. Philipp, Note on consecutive integers whose sum of squares is a perfect square, this MAGAZINE, 37 (1964) 218–220.

3. Brother U. Alfred, Sums of squares of consecutive odd integers, this MAGAZINE, 40 (1967) 194–199.

Editorial Note

J. A. H. Hunter has pointed out that from $13045^2-193\cdot 939^2=-2^7$ and $167^2-193\cdot 12^2=97$ we get the minimal solution for the case of $n=193$,

$$3791^2 - 193\cdot 273^2 = -97\cdot 2^7$$

whence $x=3599$ and $z=52{,}689$ ($X^2-193Y^2=-97\cdot 2^7$ where $X=x+192$ and $193Y=z$).

This result is a considerable improvement over the result of Sollfrey which gives $x=4{,}353{,}047$ and $z=60{,}477{,}129$.

TABLE 1

N	Digits in Previous x	x	z
73	15	67	1,241
177	13	103	3,953
193	—	4,353,047	60,477,129
249	17	3,289	55,859
292	10	110,311	1,889,970
337	23	1,027	25,275
388	11	271,727	5,360,026
393	18	1,211	32,095
409	26	54,863	1,117,797
537	20	12,955	312,713
577	—	2,263	68,663
628	17	1,281	48,670
649	34	5,327	152,515
673	—	46,671	1,228,225
708	13	134,387	3,594,634
724	—	324,545	8,752,074
753	27	18,717	534,379
772	—	343,803	9,573,958
852	14	105	31,382
913	—	339,107	10,273,989
964	19	13,355	444,886
976	13	251,725	7,894,620
996	18	97,411	3,105,694

References

1. Brother U. Alfred, Consecutive integers whose sum of squares is a perfect square, this MAGAZINE, 37 (1964) 19–32.
2. S. Philipp, Note on consecutive integers whose sum of squares is a perfect square, this MAGAZINE, 37 (1964) 218–220.
3. Brother U. Alfred, Sums of squares of consecutive odd integers, this MAGAZINE, 40 (1967) 194–199.

Editorial Note

J. A. H. Hunter has pointed out that from $13045^2-193\cdot 939^2=-2^7$ and $167^2-193\cdot 12^2=97$ we get the minimal solution for the case of $n=193$,

$$3791^2 - 193\cdot 273^2 = -97\cdot 2^7$$

whence $x=3599$ and $z=52{,}689$ ($X^2-193Y^2=-97\cdot 2^7$ where $X=x+192$ and $193Y=z$).

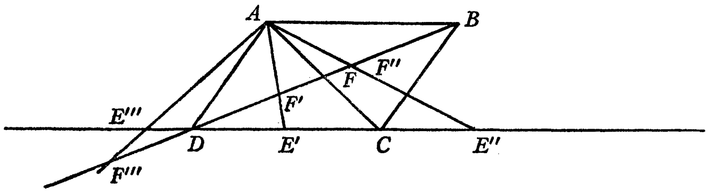
This result is a considerable improvement over the result of Sollfrey which gives $x=4{,}353{,}047$ and $z=60{,}477{,}129$.

A GEOMETRIC APPLICATION OF $f(n)=n/(n+1)$

FRANCINE ABELES, Newark State College

Introduction. The elementary geometric theorem which states that the diagonals of a parallelogram bisect each other is well known. It can, however, be placed in a much broader framework providing an interesting example of the limit of the real function $f(n)_{n\to\infty}=n/(n+1)$. The purpose of this paper is to develop the essentials utilizing the notion of points of division.

Preliminaries. In the figure below, the ratio in which C divides $\overline{DC}$ is undefined (∞), and F divides diagonal $\overline{DB}$ in the ratio 1/1. This is the familiar theorem on the diagonals of a parallelogram. It can easily be proved that if E' is the midpoint of $\overline{DC}$, i.e., E' divides $\overline{DC}$ in the ratio 1/1), then F' divides diagonal $\overline{DB}$ in the ratio 1/2 (i.e., F' is a point of trisection). Similarly, if E'' divides $\overline{DC}$ in the ratio $-3/1$, then F'' divides $\overline{DB}$ in the ratio 3/2. And if E''' divides $\overline{DC}$ in the ratio $-1/5$, then F''' divides $\overline{DB}$ (externally) in the ratio $-1/4$.



THEOREM. *In parallelogram $ABCD$, if a line through A intersects the base DC or its prolongation in a point E such that $(\overline{DE}/\overline{EC})=n$, then the distinct point F in which AE intersects diagonal DB effects the division, $(\overline{DF}/\overline{FB})=n/(n+1)$.*

Proof. Let E be any point on the base DC or its extension, $E\neq D$, such that AE meets diagonal DB or its extension in point F. (If $E=D=F$, there is nothing to prove.)

(1) $$\overrightarrow{DF} = x\overrightarrow{DB}, \quad x \neq 0, 1; \qquad \overrightarrow{DB} = \overrightarrow{DC} + \overrightarrow{CB},$$

(2) $$\overrightarrow{DE} = y\overrightarrow{DC}, \qquad y \neq 0, \quad -1$$

So,

(3) $$\overrightarrow{DF} = x/y\overrightarrow{DE} + x\overrightarrow{CB}.$$

$$\overrightarrow{DF} = \overrightarrow{DA} + \overrightarrow{AF}; \quad \overrightarrow{AF} = z\overrightarrow{AE},\ z \neq 0, 1; \quad \overrightarrow{AE} = \overrightarrow{DE} - \overrightarrow{DA},\ \overrightarrow{DA} = \overrightarrow{CB},$$

So,

(4) $$\overrightarrow{DF} = z\overrightarrow{DE} + (1 - z)\overrightarrow{CB}$$

Equating coefficients in (3) and (4), we obtain

(5) $$x/y = z \quad \text{and} \quad x = 1 - z$$

From (1),

$$\overrightarrow{DF} = x(\overrightarrow{DF} + \overrightarrow{FB}) \quad \text{and so} \quad \frac{\overline{DF}}{\overline{FB}} = \frac{x}{1-x}. \tag{6}$$

From (2),

$$\overrightarrow{DE} = y(\overrightarrow{DE} + \overrightarrow{EC}) \quad \text{and so} \quad \frac{\overline{DE}}{\overline{EC}} = \frac{y}{1-y}. \tag{7}$$

Using (5) to solve (7), the ratio of the base, we have $(y/1-y)=x/(1-2x)$. Letting $n=x/(1-2x)$, it is easily shown that (6), the ratio of the diagonal, is $n/(n+1)$.

Some Remarks. If E is strictly between D and C, we have the following specific cases: if the base is bisected, the diagonal is trisected; if the base is trisected, the diagonal is quadrisected, etc. Note that when E is a point of internal division, then $0<n<\infty$. If E is coincident with D or C, then $n=0$ or is undefined (∞), respectively. The latter situation, when the diagonal is bisected, illustrates $\lim_{n\to\infty}(n/n+1)=1$. If D is strictly between E and C, E and F points of external division, then $-1<n<0$. As E moves further to the left of D, n approaches -1. $f(n)$ is undefined at this point, corresponding to the statement that it is impossible to divide a segment in the ratio -1. When $n=-1/2$, $f(n)=-1$, i.e., $AE \| DB$.

If C is strictly between D and E, E a point of external division, then $-\infty<n<-1$. As E moves to the right of C, n approaches -1. Geometrically, $n=-1$ when $AE \| DC$.

The entire argument obviously applies to the consideration of diagonal CA with suitable adjustments. For example, the ratio of the base would be $(\overline{CE}/\overline{ED})$ and $(\overline{CF}/\overline{FA})$ for the diagonal. This indicates that the ratios are independent of any orientation assigned to the base or diagonal.

ON THE DIFFERENTIAL EQUATION $f'(x)=af(g(x))$

RAIMOND A. STRUBLE, North Carolina State University at Raleigh

In [1], W. R. Utz proposed the problem to "determine conditions for the existence of a real function $f(x)$, not identically zero, satisfying $f'(x)=af(g(x))$ wherein a is a given real constant and $g(x)$ is a given real function." The following would appear to be a rather complete resolution of this particular problem.

THEOREM. *Let $a\neq 0$, $b>0$, c and x_0 be given numbers and let $g(x)$ be continuous for $|x-x_0|\leqq b$. Then there exists a solution f of the initial value problem*

$$f'(x) = af(g(x)), \qquad f(x_0) = c \tag{1}$$

for $|x-x_0|\leqq\alpha$, where α is any positive number less than the minimum of b and $|a|^{-1}$.

From (1),

$$\overrightarrow{DF} = x(\overrightarrow{DF} + \overrightarrow{FB}) \quad \text{and so} \quad \frac{\overline{DF}}{\overline{FB}} = \frac{x}{1-x}. \tag{6}$$

From (2),

$$\overrightarrow{DE} = y(\overrightarrow{DE} + \overrightarrow{EC}) \quad \text{and so} \quad \frac{\overline{DE}}{\overline{EC}} = \frac{y}{1-y}. \tag{7}$$

Using (5) to solve (7), the ratio of the base, we have $(y/1-y)=x/(1-2x)$. Letting $n=x/(1-2x)$, it is easily shown that (6), the ratio of the diagonal, is $n/(n+1)$.

Some Remarks. If E is strictly between D and C, we have the following specific cases: if the base is bisected, the diagonal is trisected; if the base is trisected, the diagonal is quadrisected, etc. Note that when E is a point of internal division, then $0<n<\infty$. If E is coincident with D or C, then $n=0$ or is undefined (∞), respectively. The latter situation, when the diagonal is bisected, illustrates $\lim_{n\to\infty}(n/n+1)=1$. If D is strictly between E and C, E and F points of external division, then $-1<n<0$. As E moves further to the left of D, n approaches -1. $f(n)$ is undefined at this point, corresponding to the statement that it is impossible to divide a segment in the ratio -1. When $n=-1/2$, $f(n)=-1$, i.e., $AE \| DB$.

If C is strictly between D and E, E a point of external division, then $-\infty<n<-1$. As E moves to the right of C, n approaches -1. Geometrically, $n=-1$ when $AE \| DC$.

The entire argument obviously applies to the consideration of diagonal CA with suitable adjustments. For example, the ratio of the base would be $(\overline{CE}/\overline{ED})$ and $(\overline{CF}/\overline{FA})$ for the diagonal. This indicates that the ratios are independent of any orientation assigned to the base or diagonal.

ON THE DIFFERENTIAL EQUATION $f'(x)=af(g(x))$

RAIMOND A. STRUBLE, North Carolina State University at Raleigh

In [1], W. R. Utz proposed the problem to "determine conditions for the existence of a real function $f(x)$, not identically zero, satisfying $f'(x)=af(g(x))$ wherein a is a given real constant and $g(x)$ is a given real function." The following would appear to be a rather complete resolution of this particular problem.

THEOREM. *Let $a\neq 0$, $b>0$, c and x_0 be given numbers and let $g(x)$ be continuous for $|x-x_0|\leqq b$. Then there exists a solution f of the initial value problem*

$$f'(x) = af(g(x)), \qquad f(x_0) = c \tag{1}$$

for $|x-x_0|\leqq \alpha$, where α is any positive number less than the minimum of b and $|a|^{-1}$.

Proof. The proof is an easy application of the principle of contraction mappings [2] for complete metric spaces. Let α be any number satisfying $0<\alpha<\min\{b, |a|^{-1}\}$ and let $\beta=\max\{|x_0-g(x)| : |x-x_0|\leqq\alpha\}$ and $\gamma=\max\{\alpha, \beta\}$. Let S_γ denote the complete metric space consisting of all continuous functions f, defined for $|x-x_0|\leqq\gamma$ and satisfying $f(x_0)=c$, with the uniform metric

$$\rho(f_1, f_2) = \max\{|f_1(x)-f_2(x)| : |x-x_0| \leqq \gamma\}.$$

If $f\in S_\gamma$, let $F=T(f)$ be the function defined as follows:

$$\begin{aligned} F(x) &= c + a\int_{x_0}^{x} f(g(t))dt, &&\text{for } |x-x_0| \leqq \alpha \\ F(x) &= F(x_0+\alpha), &&\text{for } x_0+\alpha \leqq x \leqq x_0+\gamma \\ F(x) &= F(x_0-\alpha), &&\text{for } x_0-\gamma \leqq x \leqq x_0-\alpha. \end{aligned}$$

(Note that the composite function $f(g(t))$ is defined for $|t-x_0|\leqq\alpha$ since $|x_0-g(t)|\leqq\beta$ for $|t-x_0|\leqq\alpha$.) Then F is continuous for $|x-x_0|\leqq\gamma$ and satisfies $F(x_0)=c$. Hence T maps S_γ into S_γ. Moreover, if $f_1, f_2\in S_\gamma$, then with $F_1=T(f_1)$ and $F_2=T(f_2)$, we have

$$\begin{aligned} \rho(T(f_1), T(f_2)) &= \max_{|x-x_0|\leqq\gamma} |F_1(x)-F_2(x)| \leqq \max_{|x-x_0|\leqq\alpha} |F_1(x)-F_2(x)| \\ &= \max_{|x-x_0|\leqq\alpha} \left| a\int_{x_0}^{x} [f_1(g(t))-f_2(g(t))]dt \right| \\ &\leqq |a|\,\alpha \max_{|x-x_0|\leqq\gamma} |f_1(x)-f_2(x)| = |a|\,\alpha\rho(f_1, f_2). \end{aligned}$$

Since $|a|\alpha<1$, the mapping T is a contraction of S_γ and, therefore, possesses a unique fixed point f. Clearly f satisfies $f(x)=c+a\int_{x_0}^{x} f(g(t))dt$ for $|x-x_0|\leqq\alpha$ and is the desired solution of (1).

The solution function $f(x)$ constructed in the above proof is necessarily a constant for $x_0+\alpha\leqq x\leqq x_0+\gamma$ and $x_0-\gamma\leqq x\leqq x_0-\alpha$; thus it may be somewhat artificial. On the other hand, if $\alpha=\gamma$, it is the unique solution of (1) for $|x-x_0|\leqq\alpha$ and this would appear to be the most significant case of the theorem.

For example, if $x_0=0$ and $|g(x)|\leqq|x|$, then the unique solution of (1) is given by the infinite series

$$\begin{aligned} f(x) = c\Big[1 + ax + a^2\int_0^x g(x_1)dx_1 + a^3\int_0^x\int_0^{g(x_2)} g(x_1)dx_1dx_2 + \cdots \\ + a^n\int_0^x\int_0^{g(x_{n-1})}\cdots\int_0^{g(x_2)} g(x_1)dx_1dx_2\cdots dx_{n-1} + \cdots\Big], \end{aligned}$$

which is obtained by the Picard method of successive approximations. Actually the series is a solution in any interval on which $f(g(x))$ converges uniformly in x. In particular, with $g(x)=|x|$, the series converges for all x and the sum is ce^{ax} for $x\geqq 0$ and $c(2-e^{-ax})$ for $x<0$. With $g(x)=x^2$, the series becomes

$$f(x) = c\left[1 + ax + \frac{a^2x^3}{3} + \frac{a^3x^7}{(7)(3)} + \cdots + \frac{a^nx^{(2^n-1)}}{(2^n-1)(2^{n-1}-1)\cdots(3)} + \cdots\right]$$

and has the radius of convergence 1. With $g(x)=|x|^{1/2}$ we have $\alpha=\gamma$ only if $|a|<1$. The series solution is

$$f(x) = c\left[1 + ax + \frac{2}{3}a^2x|x|^{1/2} + \frac{2}{3}\frac{4}{7}a^3x|x|^{3/4} + \cdots + \frac{2}{3}\frac{4}{7}\cdots\frac{2^{n-1}}{2^n-1}a^nx|x|^{(2^{n-1})-1/2^{n-1}} + \cdots\right]$$

and converges for all x if $|a|<2$ and diverges for any $x\neq 0$ if $|a|\geqq 2$. It is interesting to note, however, that $f(x)=x|x|$ is a nontrivial solution of $f'(x)=2f(|x|^{1/2})$ for all x and satisfies $f(0)=0$.

It is not difficult to extend these considerations to differential equations of the form $f'(x)=F(x, f(x), f(g(x)))$, wherein $F(x, y, z)$ is Lipschitzian in y and z [3].

Results obtained in connection with research sponsored by the U. S. Army Research Office, Durham. The writer wishes to thank the referee for calling his attention to an error in the original manuscript.

References

1. W. R. Utz, The equation $f'(x)=af(g(x))$, Bull. Amer. Math. Soc., 71 (1965) 138.

2. S. C. Chu and J. B. Diaz, Remarks on a generalization of Banach's principle of contraction mappings, J. Math. Ana. Appl. 11 (1965) 440–446.

3. D. R. Anderson, An existence theorem for a solution of $f'(x)=F(x, f(g(x))$, SIAM Rev., 8 (1966) 359–362.

THE CONVERSE MALFATTI PROBLEM

MICHAEL GOLDBERG, Washington, D. C.

1. Introduction. Malfatti posed the problem of determining the sizes of three nonoverlapping circles of the greatest combined area which could be cut from a given triangle. He suggested that the solution consists of the three circles, now known as the Malfatti circles of the triangle, which are externally tangent to each other, and each of the circles is tangent to two sides of the triangle. This surmise is now known to be incorrect, as is shown in a paper by the author [1]. The solution is *never* the set of Malfatti circles.

The converse problem, which does not seem to have been considered previously, may be stated as follows:

Find the triangle of least area which can enclose three nonoverlapping circles of given radii.

This problem is more general than the original Malfatti problem. The given sets of radii can include sets which could not arise as the solutions of the original Malfatti problem. Furthermore, there are more types of configurations among the solutions of the converse problem. These, it will be seen, even include the Malfatti circles as solutions.

and has the radius of convergence 1. With $g(x)=|x|^{1/2}$ we have $\alpha=\gamma$ only if $|a|<1$. The series solution is

$$f(x) = c\left[1 + ax + \frac{2}{3}a^2x\,|x|^{1/2} + \frac{2}{3}\,\frac{4}{7}\,a^3x\,|x|^{3/4} + \cdots + \frac{2}{3}\,\frac{4}{7}\cdots\frac{2^{n-1}}{2^n-1}\,a^nx\,|x|^{(2^{n-1})-1/2^{n-1}} + \cdots\right]$$

and converges for all x if $|a|<2$ and diverges for any $x\neq 0$ if $|a|\geqq 2$. It is interesting to note, however, that $f(x)=x|x|$ is a nontrivial solution of $f'(x)=2f(|x|^{1/2})$ for all x and satisfies $f(0)=0$.

It is not difficult to extend these considerations to differential equations of the form $f'(x)=F(x, f(x), f(g(x)))$, wherein $F(x, y, z)$ is Lipschitzian in y and z [3].

Results obtained in connection with research sponsored by the U. S. Army Research Office, Durham. The writer wishes to thank the referee for calling his attention to an error in the original manuscript.

References

1. W. R. Utz, The equation $f'(x)=af(g(x))$, Bull. Amer. Math. Soc., 71 (1965) 138.

2. S. C. Chu and J. B. Diaz, Remarks on a generalization of Banach's principle of contraction mappings, J. Math. Ana. Appl. 11 (1965) 440–446.

3. D. R. Anderson, An existence theorem for a solution of $f'(x)=F(x, f(g(x))$, SIAM Rev., 8 (1966) 359–362.

THE CONVERSE MALFATTI PROBLEM

MICHAEL GOLDBERG, Washington, D. C.

1. Introduction. Malfatti posed the problem of determining the sizes of three nonoverlapping circles of the greatest combined area which could be cut from a given triangle. He suggested that the solution consists of the three circles, now known as the Malfatti circles of the triangle, which are externally tangent to each other, and each of the circles is tangent to two sides of the triangle. This surmise is now known to be incorrect, as is shown in a paper by the author [1]. The solution is *never* the set of Malfatti circles.

The converse problem, which does not seem to have been considered previously, may be stated as follows:

Find the triangle of least area which can enclose three nonoverlapping circles of given radii.

This problem is more general than the original Malfatti problem. The given sets of radii can include sets which could not arise as the solutions of the original Malfatti problem. Furthermore, there are more types of configurations among the solutions of the converse problem. These, it will be seen, even include the Malfatti circles as solutions.

2. The equilateral triangle solution. Let the given radii be a, b, c, where $a \geqq b \geqq c$. To simplify the problem, let the radius of the largest circle be taken as unity; that is, $a=1$. Then the triangle of smallest area which includes this circle is the circumscribed equilateral triangle. If the radii of the smaller circles do not exceed 1/3, they can be placed in two of the corners of the triangle without overlapping the large circle. Hence, the equilateral triangle is the solution for such a set of circles. (See Figure 1.)

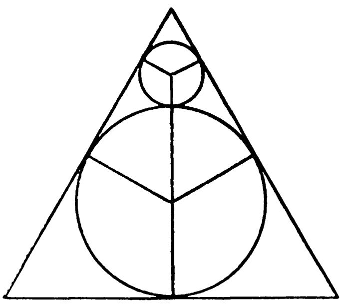


FIG. 1. Equilateral triangle

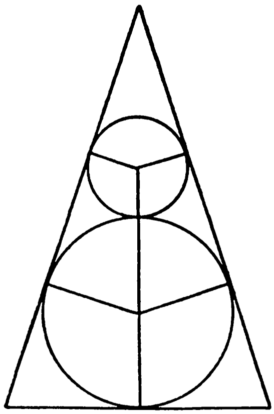


FIG. 2. Isosceles triangle

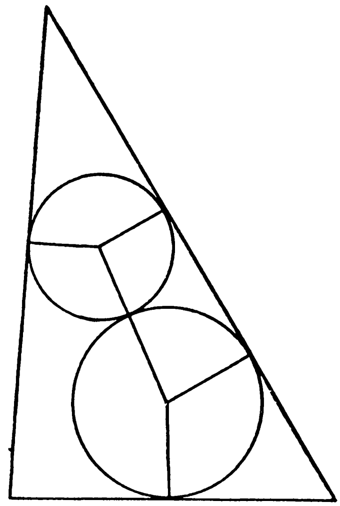


FIG. 3. Scalene triangle

3. The isosceles triangle solution. Let the second circle have a radius between 1/3 and $(1+\sqrt{17})/8$. Determine the smallest triangle which can include only the two largest circles. First, place the two circles in contact. Then, draw two sides of the sought triangle as two lines, each of which is tangent to both circles. The third side is then tangent to the other circle. If the area is to be minimized, the point of contact must be the midpoint of the third side. Hence the triangle is an isosceles triangle. If the third circle can now be placed in an unoccupied region, then this isosceles triangle is the sought solution. (See Figure 2.) The largest value that c can have is $c=b^2$.

4. The scalene triangle solution. Let the radius of the second circle be between $(1+\sqrt{17})/8$ and 1. Then, draw one side of the sought triangle as a line tangent to both of the two touching circles. Then, for the triangle of minimum area, each of the other two sides must touch a circle at the midpoint of the side. (See Figure 3.) If the third circle can be placed in an unoccupied region, then this scalene triangle is the sought solution. The maximum acceptable value of c is a function of b. It varies from $(1+\sqrt{17})^2/8^2=0.410$ to the value 0.719 approximately.

In Figure 4, the third circle can be placed in the smallest angle B, or in the third angle A. For some value of b, designated by k, the two maximum values of c are equal. By computation, k is equal to 0.85 approximately. For b less than k, the maximum value of c is obtained when c is placed in angle B. For b greater than k, the maximum value of c is obtained when c is placed in angle A.

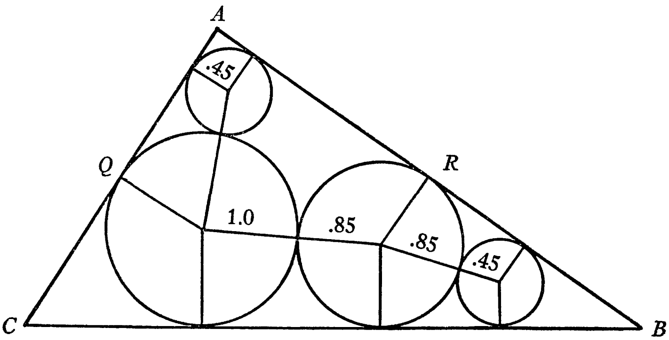


FIG. 4. Scalene triangle, transition

5. The transition case of two circles. The transition case from the isosceles solution to the scalene solution is the isosceles solution shown in Figure 5. If the first side is AB, and the third vertex is P, then the point of contact R is the midpoint of AB, and the point of contact Q is the midpoint of BP. If the second circle has radius B, and $\angle ABP=2E$, then $\sin E=(1-b)/(1+b)$, $\cos^2 E=1-\sin^2 E =4b/(1+b)^2$, $CB=2b\cot E$, $RB=CB\cos E=2b\cos^2 E/\sin E=b/\sin E+b+2$, or $b+(b+2)\sin E=2b\cos^2 E$, $b+(b+2)(1-b)/(1+b)=8b^2/(1+b)^2$, from which $4b^2-b-1=0$, so that $b=(1+\sqrt{17})/8=0.6404$ approximately.

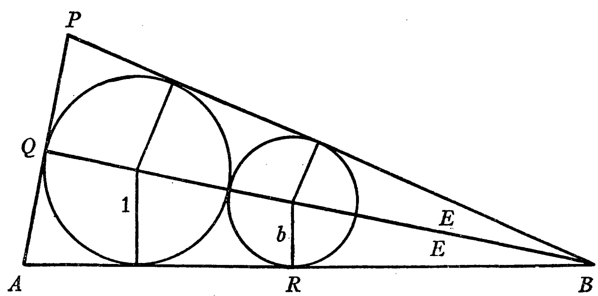


FIG. 5. Isosceles triangle, transition

6. Restrained-circle solutions. If the smallest circle cannot be placed in a vacant region of the smallest triangle enclosing the two larger circles, then other configurations of the three circles must be sought. The following solutions have been found by computation. For the smaller values of b, the arrangement is three circles tangent to the same side of the triangle with the large circle in the middle. The other two sides of the triangle are each tangent to two circles. For larger values of b and c, the largest circle no longer touches the three sides of the triangle and the two smaller circles do not touch each other. For still larger values of b and c, there is a region in which the Malfatti circles are the solutions.

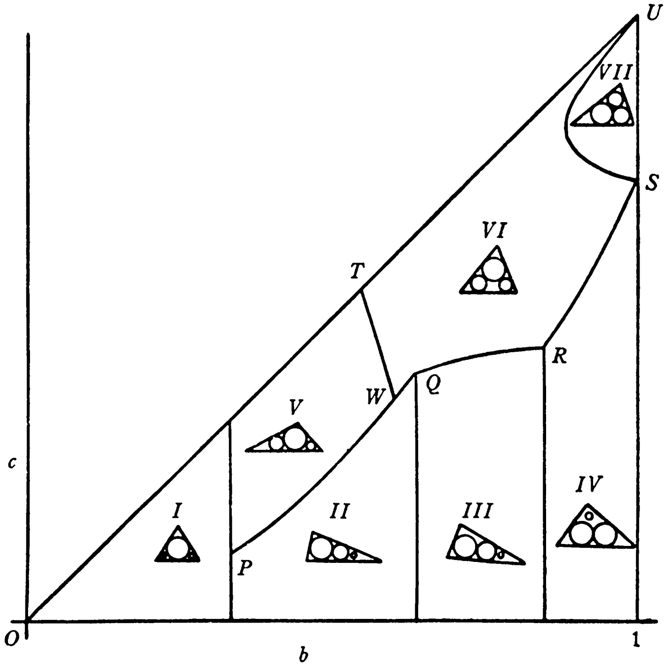


FIG. 6. Phase diagram

7. The phase diagram. The results are summarized in the phase diagram of Figure 6. Each of the possible sets of radii 1, b, c is shown by a point of the triangle of the phase diagram. This triangle is divided into seven regions. The nature of the solution in each region is shown by the geometric symbol in that region. In each of the lower regions (I, II, III, IV), at least one of the three circles is loose. In the upper regions (V, VI, VII), each of the circles is constrained. In region V, the three circles are tangent to the same side of the triangle. In region VI, each circle touches two sides of the triangle, but two of the circles do not touch each other. In region VII, the three circles are the Malfatti circles of the triangle; namely, each circle touches the other two circles and each circle touches two sides of the triangle.

The coordinates of the critical points of the diagram are as follows:

$P = 1/3,\ 1/9$

$Q = (1+\sqrt{17})/8, \quad (1+\sqrt{17})^2/64 = 0.640,\ 0.410$ approx.

$R = 0.85,\ 0.45$ approximately

$S = 1,\ 0.719$ approximately

$T = 0.55,\ 0.55$ approximately

$U = 1,\ 1$

$W = 0.60,\ 0.36$ approximately

The curve PQ is the parabola $c = b^2$. The other curves have not been derived.

8. The types of solutions with variation of the size of only one circle. Note that if the sizes of two of the circles are given, and the third circle is allowed to vary, the types of solutions will change with the choice of the size of the third circle. For example, if $a=1$, and $b=0.95$, then as c increases from zero to 0.95, the following sequence of solutions is obtained: IV, VI, VII, VI.

If the ratio of c to b is 0.625, then a line of slope 0.625 through the origin cuts through six regions, one of which is repeated, giving a sequence of seven solutions. This is illustrated by taking $c=10$, $b=16$, and letting a vary upward from 16. Then the following sequence of solutions is obtained

a	16–	17.2–	23.5–	25–	25.8–	26.7–	48–
Region	IV	VI	III	II	VI	V	I

If $c=10$, $b=12.5$, and a varies from 12.5 upward, the solutions are

a	12.5–	13.5–	21.7–	37.5–
Region	VII	VI	V	I

Reference

1. Michael Goldberg, On the original Malfatti problem, this MAGAZINE 40 (1967) 241–247.

DEFINITION VERSUS PROPERTY

WILLIAM R. RANSOM, Tufts University

Significance in naming. At first encounter, Adam may not have perceived the essential peculiarity of an animal, so that he could give it the most appropriate name: is "camel-leopard" the best definition of the giraffe? In mathematics there are numerous cases where a definition has been given, and a property has emerged worthy of ousting the original definition.

Parallels. For example, which should appear as the definition of parallels, nonintersection, or equidistance? The name, given by Greek-speaking Egyptians, "along side another," suggests that equidistance came first as a practical definition. We suppose that they were concerned with reestablishing bounds after the subsidence of the Nile's flood, and no thought of infinitely producing entered their heads. Yet Euclid took the other choice: "Parallel straight lines are such as are in the same plane, and being produced ever so far both ways, do not meet." His definition survives in most modern school texts. Yet denial is a bad attribute in a definition. The equidistant property makes a better definition: it avoids a negative, employs something visible, and easily leads to the nonmeeting as a property. The question of existence is not more difficult than in Euclid's way.

Straight and circular. The straight line is often defined as that which gives the shortest distance between two points. Euclid said it "lay evenly between two

8. The types of solutions with variation of the size of only one circle. Note that if the sizes of two of the circles are given, and the third circle is allowed to vary, the types of solutions will change with the choice of the size of the third circle. For example, if $a=1$, and $b=0.95$, then as c increases from zero to 0.95, the following sequence of solutions is obtained: IV, VI, VII, VI.

If the ratio of c to b is 0.625, then a line of slope 0.625 through the origin cuts through six regions, one of which is repeated, giving a sequence of seven solutions. This is illustrated by taking $c=10$, $b=16$, and letting a vary upward from 16. Then the following sequence of solutions is obtained

a	16–	17.2–	23.5–	25–	25.8–	26.7–	48–
Region	IV	VI	III	II	VI	V	I

If $c=10$, $b=12.5$, and a varies from 12.5 upward, the solutions are

a	12.5–	13.5–	21.7–	37.5–
Region	VII	VI	V	I

Reference

1. Michael Goldberg, On the original Malfatti problem, this MAGAZINE 40 (1967) 241–247.

DEFINITION VERSUS PROPERTY

WILLIAM R. RANSOM, Tufts University

Significance in naming. At first encounter, Adam may not have perceived the essential peculiarity of an animal, so that he could give it the most appropriate name: is "camel-leopard" the best definition of the giraffe? In mathematics there are numerous cases where a definition has been given, and a property has emerged worthy of ousting the original definition.

Parallels. For example, which should appear as the definition of parallels, nonintersection, or equidistance? The name, given by Greek-speaking Egyptians, "along side another," suggests that equidistance came first as a practical definition. We suppose that they were concerned with reestablishing bounds after the subsidence of the Nile's flood, and no thought of infinitely producing entered their heads. Yet Euclid took the other choice: "Parallel straight lines are such as are in the same plane, and being produced ever so far both ways, do not meet." His definition survives in most modern school texts. Yet denial is a bad attribute in a definition. The equidistant property makes a better definition: it avoids a negative, employs something visible, and easily leads to the nonmeeting as a property. The question of existence is not more difficult than in Euclid's way.

Straight and circular. The straight line is often defined as that which gives the shortest distance between two points. Euclid said it "lay evenly between two points," but did not obtain the geodesic property from his definition although moderns usually say (as Euclid does not) that the sum of two sides of a triangle is greater than the third side.

As for the circle, it is defined as the locus of a point at a given distance from its center, and its constant curvature becomes a property. It would be an interesting problem to reverse the status of these two aspects of the circle.

Plane figures. Polygon, which means many corners or angles, is defined as many *sides*. The equilateral triangle has three equal sides, but it also has three equal angles, and could be called equiangular just as well. The property and definition are quite interchangeable without inconvenience. Similarly "quadrilateral" and "quadrangle" are the same thing, although "complete quadrilateral" and "complete quadrangle" are quite different.

Logarithms and trigonometric functions. To Napier, the Briggs logarithms, so important in precomputer days, were defined by means of concordant arithmetic and geometric progressions. But as algebraic notation improved, the exponent property emerged and took its place as definition.

A similar interchange may be taking place in the definition of the sine as the ratio of side to hypotenuse. This is held by some less useful than the important periodic property associated with the numerical definition by means of an infinite series.

Conic Sections. Originally ellipse, parabola, and hyperbola were defined as plane sections of a cone, and the focal properties were derived. The definition of conics by means of focus and directrix, with ratio, sum or difference, prevail in current textbooks. But these features can be derived from the equations more readily than the equations can be derived from focal definitions. It may be that the focal properties, now accepted as definitions, should return to their status as properties.

Present proposal. It is the purpose of this article to set forth the advantages of a shift from property to definition in the case of the conic sections and the Naperian base, "*e*."

Making the conics useful. The greatest number of important and familiar functions are linears, squares, and reciprocals: for examples, $v=v_0+gt$, $s=gt^2/2$, $p=p_0v_0/v$, etc., etc. The simplest cases of these appear in analytic geometry as $y=x$, $y=x^2$, and $y=1/x$. Shifts and stretches (or shrinks) readily convert these to the more general cases:

$$y = ax + b,$$

$$y = b + (x - a)^2, \quad \text{or} \quad y = ax^2 + bx + c, \quad \text{and}$$

$$(x - a)(y - l) = 1.$$

Parabola. We may take the equation $y=x^2$ as definition of the "standard parabola," and show that $y=kx^2$—as $ky=(kx)^2$—has the same shape as $y=x^2$ but is $1/k$ times as large, and has the property that all its points are equidistant from a "focus" at (0, $1/4k$), and a "directrix" whose equation is $y=-1/4k$. The "reflection to focus" property can then be derived as customary directly from the equation $y=kx^2$.

Hyperbola. It is further proposed to take $y=1/x$ as definition of a "standard hyperbola," and show that $xy=a^2$ has the same shape but a different size. The asymptotes appear naturally and immediately from this definition.

The prevailing standard equation, $x^2/a^2-y^2/b^2=1$, has far fewer applications ("LORAN" is one) than the proposed $y=a^2/x$. It is readily derived by a 45° rotation, giving $x^2-y^2=2a^2$, from which stretches (or shrinks) give the familiar equation in a far less complicated manner than the methods which prevail in textbooks.

From the point of view of analytic geometry, there is a curious defect in the definitions of ellipse and hyperbola as $PF \pm PF' = 2a$. In the second squaring the distinction which the definition makes between sum and difference, + and −, disappears completely, and the thus lost distinction emerges in the appearance of the factor (a^2-c^2), which in no way accords with the definition the deduction started with!

Ellipse. The simplest and most useful definition is as a stretched circle. $x^2+y^2=a^2$ readily becomes $x^2/a^2+y^2/b^2=1$. This definition affords the draftsman a more convenient construction, and fits in better with perspective theory than the focal definition. The focal property is more easily derived from the equation than the equation from the focal property, and there is no question about the distinction from the hyperbolas.

The Napierian base. A more natural and easily applicable definition of $e=2.718\cdots$, arises after consideration that the graphs of $y=2^x$ and $y=3^x$ cross the y axis with slopes that are approximately 0.7 and 1.1; it is then natural to expect that there will be a number—call it e—such that $y=e^x$ crosses the y axis with a slope exactly one. Let this be taken as the definition of e: $\text{limit}_{x\to 0}(e^x-1/x)=1$.

This definition can be used directly in the differentiation of e^x, while the usual definition $\text{limit}_{x\to\infty}(1+1/x)^x=e$ entails formidable complications, and has a far less obvious motivation.

The value of e as $1+1+1/2!+1/3!+\text{etc.}$, comes out readily when the rule $d(e^x)=e^x\cdot dx$ is used in expanding by Maclaurin's Theorem, and this is much simpler than the work with the $(1+1/x)^x$ definition.

Experience. These proposals, to base the initial study of the conics on the equations $y=x^2$, $y=1/x$, and $x^2+y^2=a^2$, and the definition of e on a unit slope, have been tried out with ever so many classes by a good many different teachers, and no difficulties have been found to arise. It seems that they may well be widely adopted.

COMPLEX VARIABLES AND LINE-COORDINATES

R. KITTAPPA, University of Delaware, Newark

1. Setting up a correspondence between points (x, y) in a plane and complex numbers $x+iy$ has been very fruitful in the development of the theory of complex variables.

Is it possible to set up a similar correspondence between lines in a plane and complex numbers? Since the equation of a line in a plane contains two independent constants we can take those constants or the functions of those constants as the real and imaginary parts of a complex number and hence set up the required correspondence.

Now we have to choose a pair of independent constants, or functions of the constants, from the constants in the equation of a line.

The simple and advantageous way is to choose l and m in $lx+my=1$ as the real and imaginary part of the corresponding complex number.

The advantage of the choice lies in the fact that the point represented by a given complex number and the line represented by the same complex number will be pole and polar with respect to the circle $x^2+y^2=1$. So, by this choice, a complete duality is found to exist between the 'point complex plane' and the 'line complex plane' (if they can be called so), with respect to the base conic $x^2+y^2=1$.

This duality also brings along with it the one disadvantage of the point complex plane, viz., that there is only one point at infinity in the complex plane, and this is just a subterfuge to avoid the fact that the direction of the points at infinity in the complex plane are indeterminate.

Dually, in the line complex plane we have that the lines through the origin in the line complex plane are indeterminate as to their direction.

We can adopt a similar subterfuge here by agreeing that there is just one line through the origin so that one complex number, $\alpha+i\alpha$, will stand for all lines through the origin. However, since the origin is a finite point we may have to consider lines through the origin in many discussions and so this would be a drawback.

The following ruse can be adopted to advantage:

The equation of the line through the origin parallel to

$$lx + my = 1$$

is

$$lx + my = 0$$

or

$$lx + my = \epsilon, \quad \text{where} \quad \epsilon \to 0$$

or

$$\frac{l}{\epsilon}x + \frac{m}{\epsilon}y = 1 \quad \text{where} \quad \epsilon \to 0.$$

So, $Lt_{\epsilon\to 0}(l+im/\epsilon)$ will represent the line, $lx+my=0$. It can be verified that this representation of a line through the origin can be used without any disadvantage in all problems such as finding the intersection of two lines, the angle between two lines, etc.

It may be mentioned here that it is convenient to represent by L the complex number $l+im$ and the line represented by it just as we represent by z the complex number $x+iy$ and the point represented by it.

2. The new representation may not be very useful unless a simple polar representation giving scope for the use of DeMoivre's Theorem, etc., is found.

We have,

$$L = l + im = \sqrt{l^2+m^2}\left(\frac{l}{\sqrt{l^2+m^2}} + i\frac{m}{\sqrt{l^2+m^2}}\right).$$

So, the choice for the radius vector will be $\sqrt{l^2+m^2}$ or $|L|$ which is the reciprocal of the length of the perpendicular R drawn from the origin to the line L.

For the amplitude the choice will obviously be

$$\tan\phi = \frac{m}{l}$$

where ϕ is the angle made by the perpendicular from the origin to L with the initial position of the radius vector. So we have

$$L = \frac{1}{R}e^{i\phi}.$$

Incidentally, this suggests a good polar representation for use in line coordinate geometry, viz., (R, ϕ). This can be called line-polar as distinct from point-polar, (r, θ).

THEOREM 1. *If $f(L, \bar{L})=0$ is the equation in line complex of the envelope of a curve, then the first positive pedal with respect to the origin is given in point-polar equation by*

$$f\left(\frac{1}{r}e^{i\theta}, \frac{1}{r}e^{-i\theta}\right) = 0.$$

Proof. This follows from the fact that (R, ϕ) are the point-polar coordinates of the foot of the perpendicular from the origin.

3. Since a complex number can now be interpreted as a point and as a line in a plane, we will have many interpretations of a complex expression depending on the number of variables involved in the expression. We have,

THEOREM 2. *For an expression involving n complex variables the total number of interpretations will be 2^{n+1} which can be paired into 2^n duals.*

Proof. The proof of the theorem depends on the fact that not only can each complex variable have two interpretations but the expression taken as a whole can have two interpretations.

Let us consider an example involving one complex variable.

Variable	*Expression*	*Interpretation*
point W	point $\frac{W}{\lvert W\rvert}$	A point at unit distance from the origin and lying on the line joining the point W and the origin.
point W	line $\frac{W}{\lvert W\rvert}$	A line passing through the point W and at unit perpendicular distance from the origin.
line W	point $\frac{W}{\lvert W\rvert}$	A point at unit distance from the origin lying on the line drawn perpendicular to W from the origin.
line W	line $\frac{W}{\lvert W\rvert}$	A line parallel to W at unit distance from the origin.

4. After setting up the correspondence referred to in Sections 1 and 2, we give some results here in line complex geometry which can be easily verified by using the idea of duality or by coordinate geometry.

A. (1) A real number p represents the line $x=1/p$. A purely imaginary number iq represents the line $y=1/q$; $L=0$ represents the line at infinity.

(2) kL represents a line parallel to L for real k and a line perpendicular to L for imaginary k.

(3) $(mL_1+nL_2)/(m+n)$ represents a pencil through the intersection of L_1 and L_2, for real m and n.

(4) x-intercept of L is $2/(L+\bar{L})$.
y-intercept of L is $2i/(L-\bar{L})$.
Gradient of the line L is $-(i(L+\bar{L})/L-\bar{L})$.

(5) Condition of parallism of L_1 and L_2 is $\bar{L}_1L_2-L_1\bar{L}_2=0$.
Condition of perpendicularity of L_1 and L_2 is $\bar{L}_1L_2+L_1\bar{L}_2=0$.
Angle between L_1 and L_2 is given by

$$\tan\theta = (\bar{L}_1L_2 - L_1\bar{L}_2)/(i(\bar{L}_1L_2 + L_1\bar{L}_2)).$$

(6) $\bar{L}$ is the reflection of L in the x-axis and
$-L$ is the reflection of L in the origin.

(7) The sum of L_1+L_2 is constructed geometrically as follows:
Draw lines from the origin parallel to the lines L_1 and L_2. The diagonal not containing the origin of the parallelogram so formed is L_1+L_2.

(8) The product of L_1 and L_2 is constructed geometrically as follows:
Let the point complex coordinates of the foot of the perpendicular from the origin on L_1 and L_2 be z_1 and z_2 respectively. Find, geometrically, the product Z of z_1 and z_2. The line through Z perpendicular to OZ is the required line.

B. (1) THEOREM. *If $u(l, m)$ and $v(l, m)$ are the real and imaginary parts of an analytic function $f(L)$, then the common tangents of the families of en-*

velopes $u(l, m)=a$ and $v(l, m)=b$, drawn by taking one member from each family, will subtend a right angle at the origin, a and b being parameters.

To prove this statement we require the Cauchy-Riemann equations for the analytic function $f(L)$ and the following lemma which is the result corresponding to $\tan\psi=(dy/dx)$ for a point curve:

LEMMA. *For an envelope (of $lx+my=1$), $dl/dm=-\tan\theta$, where θ is the angle between the x-axis and the line joining the point o contact of $lx+my=1$ and the origin.*

The last result can be proved by considering the limiting case of two intersecting tangents when they coincide with one another.

(2) If $F(l, m)=c$, where c is a constant, is an envelope then ∇F is a line perpendicular to the join of the origin and the point of contact of $lx+my=1$.

(3) *Envelope integral.* Let $f(L)$ be continuous for all positions of the generator L of the envelope E. Divide E into n parts by choosing arbitrarily the generators $L_1, L_2, \cdots, L_{n-1}$ and call the generators at the ends L_0 and L_n. Choose a generator g_k between L_{k-1} and L_k. Form the sum

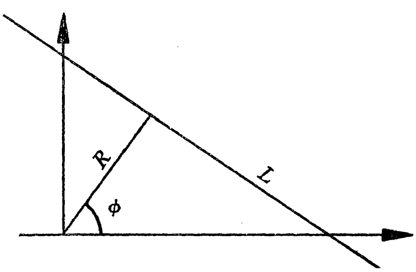
$$S_n=\sum_{k=1}^{k=n} f(g_k)(L_k-L_{k-1})=\sum_{k=1}^{k=n} f(g_k)\Delta L_k.$$

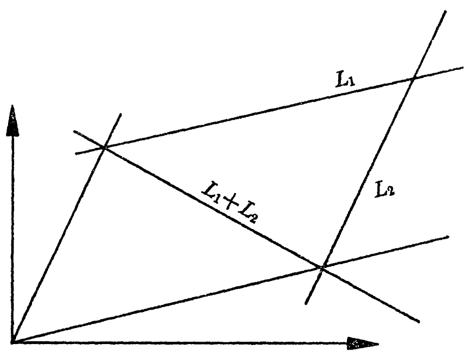
That the limit of the sum as the largest of $\|L_k\|$ approaches zero (if such a limit exists) can be denoted by $\int_E f(L)dL$ and may be called the envelope integral of $f(L)$ along the envelope E.

It may be noted that for the contour envelope integral, the curve enveloped need not be closed since it would be taken from one tangent to the same tangent along the envelope.

So, in writing out the integral theorems of line complex plane, we cannot expect them to be very similar to the theorems in point complex plane.

C. The results on transformations are similar to those of point complex plane.





CHARACTERIZING A FAMILY OF COMPLEX POLYTOPES WITHOUT USING COMPLEX COORDINATES

ERIC HALSEY, University of Washington

0. If α_n is the n-simplex, then $N_k(n)$, the number of k-faces of α_n, is

$$N_k(n) = \binom{n+1}{k+1}. \tag{1}$$

(See [1] for the notations and definitions used in this paper.) By letting the null set have dimension -1, we can have the numbers $N_k(n)$, for $-1 \leqq k \leqq n$ and $n = -1, 0, 1, 2, \cdots$, in one-to-one correspondence with the elements of Pascal's triangle, which are the very same binomial coefficients. It is interesting to inquire whether this type of combinatorial correspondence can be generalized to other families of regular complex polytopes; the paper gives an affirmative answer.

1. Shephard [1] shows that there are only two regular complex polytopes besides the regular (real) simplex that generalize to $n \geqq 5$ dimensions. These are (in his notation):

(i) $\gamma_n^m = m(2m^2)2(6)2(6) \cdots (6)2(6)2$, the generalized orthotope, whose m^n vertices have coordinates

$$(\epsilon^{k_1}\, \epsilon^{k_2}, \cdots, \epsilon^{k_n}) \quad \text{where} \quad \epsilon = e^{2\pi i/m} \qquad 1 \leqq k_1, k_2, \cdots, k_n \leqq m;\ m \geqq 2$$

in unitary n-space.

(ii) $\beta_n^m = 2(6)2(6) \cdots (6)2(6)2(2m^2)m$, the generalized cross polytope, whose mn vertices have coordinates

$$(\epsilon^k, 0, 0, \cdots, 0), (0, \epsilon^k, 0, \cdots, 0), \cdots, (0, \cdots, 0, \epsilon^k) \qquad 1 \leqq k \leqq m;\ m \geqq 2$$

in unitary n-space.

For $m=2$ these are just the real orthotope and the real cross polytope. Let the number of k-faces of γ_n^m and β_n^m be $N_k(\gamma_n^m)$ and $N_k(\beta_n^m)$, respectively. Since β_n^m is reciprocal to γ_n^m, then $N_k(\beta_n^m) = N_{n-k-1}(\gamma_n^m)$, and we will consider only the numbers $N_k(\gamma_n^m)$, which we will denote by $N_k^m(n)$. The value

$$N_k^m(n) = m^{n-k}\binom{n}{k} \qquad 0 \leqq k \leqq n \tag{2}$$

of $N_k^m(n)$ is easy to verify by observing that each k-face γ_k^m of γ_n^m is obtained by fixing $n-k$ of the coordinates of γ_n^m, or by the following induction on the dimension n.

The formula is obviously valid when $n=0$ (since γ_0^m is just a point), and in fact whenever $k=n$; likewise γ_1^m is the so-called "m-line" whose m vertices are the m mth roots of unity in the complex plane. Now let γ_{n-1}^m be centered at the origin in unitary $(n-1)$-space, with $N_k^m(n-1) = m^{n-k-1}\binom{n-1}{k}$. In n-space, m parallel translates of γ_{n-1}^m (in the new orthogonal "direction" or 1-space) are centered at the m mth roots of unity in that 1-space. All the vertices together of these m parallel replicas are the vertices of a γ_n^m, centered at the origin in n-

space. Evidently some of the k-faces γ_k^m of γ_n^m are exactly the k-faces of the m replica γ_{n-1}^m's, which are faces of γ_n^m, and each remaining one has as its vertices all the m translates of the vertices of a certain $(k-1)$-face γ_{k-1}^m of γ_{n-1}^m. Accordingly $N_k^m(n)=mN_k^m(n-1)+N_{k-1}^m(n-1)=m^{n-k}\binom{n}{k}$.

2. Pascal's triangle (notation: P_2) can be characterized as an array of numbers $p(i,j)$ $(i,j=0,1,2,\cdots)$ such that $p(0,0)=1$ and otherwise $p(i,j)=p(i-1,j)+p(i,j-1)$, provided that the latter terms are in P_2. If any are not, they are simply omitted from the sum. The nth *rank* of Pascal's triangle is $R_n(P_2)=\{p(i,j)\in P_2 \mid i+j=n\}$; the nth *stratum* is $S_n(P_2)=\{p(i,j)\in P_2 \mid j=n\}$. The fact that $p(i,j)=(i+j)!/i!j!$ is trivial if i or j is 0. Otherwise it follows from the following induction on rank:

$$p(i,j)=p(i-1,j)+p(i,j-1)=\frac{(i+j-1)!}{(i-1)!j!}+\frac{(i+j-1)!}{i!(j-1)!}$$

$$=\frac{i(i+j-1)!+j(i+j-1)!}{i!j!}=\frac{(i+j)!}{i!j!}.$$

3. We generalize Pascal's triangle to Pascal's n-simplex, P_n (for $n\geqq 2$), in the following way: P_n is an array of numbers $p(i_1,i_2,\cdots,i_n)$ $(i_1,i_2,\cdots,i_n=0,1,2,\cdots)$ such that $p(0,0,\cdots,0)=1$ and otherwise

$$p(i_1,i_2,\cdots,i_n)=\sum_{r=1}^{n}p(i_1,\cdots,i_r-1,\cdots,i_n)$$

provided that the latter terms are in P_n; if any is not, it is simply omitted from the sum. The kth rank of P_n is $R_k(P_n)=\{p(i_1,\cdots,i_n)\in P_n \mid \sum_{r=1}^{n}i_r=k\}$; the kth stratum is $S_k(P_n)=\{p(i_1,\cdots i_n)\in P_n \mid i_n=k\}$. We want to show that if $p(i_1,\cdots,i_n)$ is an element of P_n, then

$$p(i_1,\cdots,i_n)=\frac{(i_1+i_2+\cdots+i_n)!}{i_1!i_2!\cdots i_n!}=\frac{\left(\sum_{r=1}^{n}i_r\right)!}{\prod_{r=1}^{n}(i_r!)}. \tag{3}$$

If any $i_r=0$, then by the symmetry of P_n we can assume $r=n$, so that $i_n=0$ and $p(i_1,\cdots,i_n)\in S_0(P_n)$. But it is clear that $S_0(P_n)=P_{n-1}$, with $p(i_1,\cdots,i_{n-1},0)=p(i_1,\cdots,i_{n-1})$. We proceed by induction on n, starting with $n=2$, for which (3) was proved true in Section 2. Assume that (3) holds for P_{n-1}. Then

$$p(i_1,\cdots,i_{n-1},0)=p(i_1,\cdots,i_{n-1})=\frac{\left(\sum_{r=1}^{n-1}i_r\right)!}{\prod_{r=1}^{n-1}(i_r!)}$$

$$= \frac{\left(0 + \sum_{r=1}^{n-1} i_r\right)!}{0! \prod_{r=1}^{n-1} (i_r)!} = \frac{\left(\sum_{r=1}^{n} i_r\right)!}{\prod_{r=1}^{n} (i_r!)}.$$

If, on the other hand, no $i_r = 0$, then we proceed by induction on the rank of P_n:

$$p(i_1, \cdots, i_n) = \sum_{q=1}^{n} p(i_1, \cdots, i_q - 1, \cdots, i_n) = \sum_{q=1}^{n} \frac{\left(\left(\sum_{r=1}^{n} i_r\right) - 1\right)!}{(i_q - 1)! \prod_{r \neq q} (i_r!)}$$

$$= \sum_{q=1}^{n} \frac{i_q\left(\left(\sum_{r=1}^{n} i_r\right) - 1\right)!}{\prod_{r=1}^{n} (i_r!)} = \frac{\left(\left(\sum_{r=1}^{n} i_r\right) - 1\right)!}{\prod_{r=1}^{n} (i_r!)} \sum_{q=1}^{n} i_q = \frac{\left(\sum_{r=1}^{n} i_r\right)!}{\prod_{r=1}^{n} (i_r!)}.$$

This means that $p(i_1, \cdots, i_n) \in P_n$ is just the coefficient of $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ in the multinomial expansion $(x_1 + x_2 + \cdots + x_n)^k$, where $p(i_1, \cdots, i_n) \in R_k(P_n)$. Letting $x_1 = x_2 = \cdots = x_n = 1$, we derive as a corollary

$$\sum_{R_k(P_n)} p(i_1, \cdots, i_n) = n^k. \tag{4}$$

To conclude this section we will show that for $p(i_1, \cdots, i_n) \in P_n$ and its "projection" $p(i_1, \cdots, i_{n-1}) \in P_{n-1}$,

$$p(i_1, \cdots, i_n) = \left\{ \begin{matrix} \sum_{r=1}^{n} i_r \\ i_n \end{matrix} \right\} p(i_1, \cdots, i_{n-1}).$$

The proof is just a straightforward calculation:

$$\left\{ \begin{matrix} \sum_{r=1}^{n} i_r \\ i_n \end{matrix} \right\} p(i_1, \cdots, i_{n-1}) = \frac{\left(\sum_{r=1}^{n} i_r\right)!}{i_n! \left(\sum_{r=1}^{n-1} i_r\right)!} \frac{\left(\sum_{r=1}^{n-1} i_r\right)!}{\prod_{r=1}^{n-1} (i_r!)} = \frac{\left(\sum_{r=1}^{n} i_r\right)!}{\prod_{r=1}^{n} (i_r!)}$$

$$= p(i_1, \cdots, i_n).$$

What this means is that $S_k(P_n)$ is just a "copy" of P_{n-1} with every number in the jth rank of P_{n-1} multiplied by $\binom{j+k}{k}$.

4. We are now almost done, as we have enough equipment to find the combinatorial numbers we want. Look at the intersection $T_{ij}(n) = R_i(P_n) \cap S_j(P_n)$ of the ith rank with the jth stratum of P_n. We wish to calculate the numerical sum $A_{ij}(n)$ of all the $p(i_1, \cdots, i_n)$ in $T_{ij}(n)$. (Of course $T_{ij}(n)$ is empty and $A_{ij}(n) = 0$ whenever $i<j$.)

$$A_{ij}(n) = \sum{}^* p(i_1, \cdots, i_n) = \sum{}^* \begin{bmatrix} \sum_{r=1}^{n} i_r \\ i_n \end{bmatrix} p(i_1, \cdots, i_{n-1})$$

$$= \binom{i}{j} \sum{}^* p(i_1, \cdots, i_{n-1})$$

(where $\sum^*$ is understood to be taken over all the numbers $p(i_1, \cdots, i_n)$ in $T_{ij}(n)$)

$$= \binom{i}{j} \sum_{R_{i-j}(P_{n-1})} p(i_1, \cdots, i_{n-1}).$$

In light of (4) this is just

$$A_{ij}(n) = \binom{i}{j}(n-1)^{i-j}.$$

But then, for $n \geqq 3$, $A_{ij}(n)$ is exactly $N_j^{n-1}(i)$, the number of j-faces in γ_i^{n-1}. As desired, we have achieved a complete characterization of the numbers $N_k(\gamma_n^m)$ for the generalized complex orthotopes in terms of the coefficients of generalized Pascal's triangles.

References

1. H. S. M. Coxeter, Regular Polytopes, Macmillan, New York, 1963.
2. G. C. Shephard, Regular complex polytopes, Proc. London Math. Soc., 3, 2 (1952).

A GENERALIZED EULER ϕ-FUNCTION

ECKFORD COHEN, University of Florida

1. In [2] Klee discussed the function $\phi_k(n)$ defined to be the number of positive integers $m \leqq n$ such that (m, n) is divisible by no kth power $\neq 1$. Evidently, ϕ_k becomes the Euler ϕ-function in the case $k=1$. Later McCarthy [3] gave an independent discussion of the function ϕ_k from a point of view different from that of Klee. He also considered the average order of ϕ_k, proving

$$(1) \qquad \Phi_k(n) \overset{\text{def.}}{=} \sum_{m=1}^{n} \phi_k(m) = \alpha_k n^2 + O(n), \qquad k > 1,$$

where

$$(2) \qquad \alpha_k = 1/2\zeta(2k), \qquad \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad s > 1.$$

sum $A_{ij}(n)$ of all the $p(i_1, \cdots, i_n)$ in $T_{ij}(n)$. (Of course $T_{ij}(n)$ is empty and $A_{ij}(n) = 0$ whenever $i<j$.)

$$A_{ij}(n) = \sum{}^* p(i_1, \cdots, i_n) = \sum{}^* \left[\begin{matrix} \sum_{r=1}^{n} i_r \\ i_n \end{matrix} \right] p(i_1, \cdots, i_{n-1})$$

$$= \binom{i}{j} \sum{}^* p(i_1, \cdots, i_{n-1})$$

(where $\sum^*$ is understood to be taken over all the numbers $p(i_1, \cdots, i_n)$ in $T_{ij}(n)$)

$$= \binom{i}{j} \sum_{R_{i-j}(P_{n-1})} p(i_1, \cdots, i_{n-1}).$$

In light of (4) this is just

$$A_{ij}(n) = \binom{i}{j} (n-1)^{i-j}.$$

But then, for $n \geqq 3$, $A_{ij}(n)$ is exactly $N_j^{n-1}(i)$, the number of j-faces in γ_i^{n-1}. As desired, we have achieved a complete characterization of the numbers $N_k(\gamma_n^m)$ for the generalized complex orthotopes in terms of the coefficients of generalized Pascal's triangles.

References

1. H. S. M. Coxeter, Regular Polytopes, Macmillan, New York, 1963.
2. G. C. Shephard, Regular complex polytopes, Proc. London Math. Soc., 3, 2 (1952).

A GENERALIZED EULER ϕ-FUNCTION

ECKFORD COHEN, University of Florida

1. In [2] Klee discussed the function $\phi_k(n)$ defined to be the number of positive integers $m \leqq n$ such that (m, n) is divisible by no kth power $\neq 1$. Evidently, ϕ_k becomes the Euler ϕ-function in the case $k=1$. Later McCarthy [3] gave an independent discussion of the function ϕ_k from a point of view different from that of Klee. He also considered the average order of ϕ_k, proving

$$\Phi_k(n) \overset{\text{def.}}{=} \sum_{m=1}^{n} \phi_k(m) = \alpha_k n^2 + O(n), \qquad k > 1, \tag{1}$$

where

$$\alpha_k = 1/2\zeta(2k), \qquad \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad s > 1. \tag{2}$$

The corresponding result in the classical case $k=1$ is

$$\Phi(n) \overset{\text{def.}}{=} \sum_{m=1}^{n} \phi(m) = \frac{n^2}{2\zeta(2)} + O(n \log n), \qquad n \geqq 2. \tag{3}$$

McCarthy's proof of (1), like the classical proof of (3), is based on an explicit representation of ϕ_k and hence involves the use of inversion functions. In this note we prove (1) differently from McCarthy, making no appeal to the Moebius function or its generalizations. The method which we use dates back to Dirichlet who applied it to the ϕ-function. The point to be noted here is that, while Dirichlet's method fails to yield a result for $\Phi(n)$ comparable to that in (3), it easily suffices to prove McCarthy's formula (1). For an exposition of Dirichlet's method in the classical case as well as for further references, the reader may consult [1].

2. We depart from the formula

$$\sum_{d^k e = n} \phi_k(e) = n, \tag{4}$$

where the summation is over all ordered pairs of positive integers d, e such that $d^k e = n$. To prove (4) let $(a, b)_k$ denote the greatest common kth power divisor of a, b. Also, for each kth power divisor d^k of n let $S_k(d)$ denote the set of integers m, $0 < m \leqq n$, such that $(m, n)_k = d^k$. A number $m = d^k m'$ is contained in $S_k(d) \rightleftarrows 0 < m' \leqq n/d^k$, $(m', n/d^k)_k = 1$. There are therefore $\phi_k(n/d^k)$ numbers in $S_k(d)$. The formula (4) now results, if one observes that the sets $S_k(d)$ partition the set, $1 \leqq m \leqq n$, as d^k ranges over the kth power divisors of n.

In addition to (4) we require a simple estimate for the remainder of the series in (2) in the case $s=2$. We observe that for $n \geqq 1$

$$\sum_{m>n} \frac{1}{m^2} < \sum_{m=n+1}^{\infty} \frac{1}{m(m-1)} = \sum_{m=2}^{\infty} \frac{1}{m(m-1)} - S_n$$

where (with vacuous sums assumed 0)

$$S_n = \sum_{m=2}^{n} \frac{1}{m(m-1)} = \sum_{m=2}^{n} \left(\frac{1}{m-1} - \frac{1}{m} \right) = 1 - \frac{1}{n} \cdot$$

Hence the sum of the above series with partial sums S_n is 1 and

$$\sum_{m>n} \frac{1}{m^2} < \frac{1}{n}, \qquad n \geqq 1. \tag{5}$$

3. In the following we suppose that $k \geqq 2$. Placing

$$\Phi_k(n) = \alpha_k n^2 + R_k(n), \tag{6}$$

it is our aim to prove that $R_k(n) = O(n)$. For real $x \geqq 1$ we define $\Phi_k(x) = \Phi_k([x])$ and $R_k(x) = R_k([x])$, where $[x]$ denotes the largest integer $\leqq x$. Since

$$\sum_{m \leqq n} \sum_{d^k e = m} \phi_k(e) = \sum_{d^k e \leqq n} \phi_k(e) = \sum_{d \leqq \sqrt[k]{n}} \Phi_k\left(\frac{n}{d^k}\right),$$

it follows from (4) on replacing n by m and summing from 1 to n that

$$\sum_{m\leq\sqrt[k]{n}} \Phi_k\left(\frac{n}{m^k}\right) = \frac{n(n+1)}{2}. \tag{7}$$

By (6),

$$\sum_{m\leq\sqrt[k]{n}} \Phi_k\left(\frac{n}{m^k}\right) = \alpha_k \sum_{m\leq\sqrt[k]{n}} \left[\frac{n}{m^k}\right]^2 + \sum_{m\leq\sqrt[k]{n}} R_k\left(\frac{n}{m^k}\right). \tag{8}$$

Writing $[x]=x-\delta$, $\delta=\delta(x)$, it follows, on squaring and taking absolute values, that

$$\left|\sum_{m\leq\sqrt[k]{n}} \left[\frac{n}{m^k}\right]^2 - \sum_{m\leq\sqrt[k]{n}} \frac{n^2}{m^{2k}}\right| \leqq \sum_{m\leq\sqrt[k]{n}} \left(2\delta\frac{n}{m^k}+\delta^2\right) \leqq \sum_{m\leq\sqrt[k]{n}} \left(\frac{2n}{m^k}+1\right)$$
$$\leqq 3n \sum_{m\leq\sqrt[k]{n}} \frac{1}{m^k} \leqq (\text{const.})n,$$

since the last sum is the partial sum of a convergent series. Hence

$$\sum_{m\leq\sqrt[k]{n}} \left[\frac{n}{m^k}\right]^2 = n^2\left(\zeta(2k) - \sum_{m>\sqrt[k]{n}} \frac{1}{m^{2k}}\right) + O(n).$$

By (5),

$$\sum_{m>\sqrt[k]{n}} \frac{1}{m^{2k}} = \sum_{m^k>n} \left(\frac{1}{m^k}\right)^2 \leqq \sum_{h>n} \left(\frac{1}{h}\right)^2 < \frac{1}{n},$$

and therefore

$$\sum_{m\leq\sqrt[k]{n}} \left[\frac{n}{m^k}\right]^2 = n^2/2\alpha_k + O(n).$$

From (8) and the fact that the right side of (7) $=n^2/2+O(n)$, it follows that

$$\sum_{m\leq\sqrt[k]{n}} R_k\left(\frac{n}{m^k}\right) = G_k(n), \tag{9}$$

where

$$G_k(n) = O(n), \qquad n \geqq 1. \tag{10}$$

We shall prove as a consequence of (9) that

$$|R_k(n)| \leqq c_k n, \qquad n \geqq 1, \tag{11}$$

where c_k is a constant defined as follows. First, the case $n=1$ in (5) permits us to write

$$\sum_{m=2}^{\infty} \frac{1}{m^2} = \epsilon < 1. \tag{12}$$

Put $\epsilon' = 1 - \epsilon$ so that

$$\epsilon + \epsilon' = 1, \qquad 0 < \epsilon' < 1. \tag{13}$$

By (10), $G_k(n)/n$ is bounded, which assures the existence of a constant c_k such that

$$| G_k(n) | \leqq c_k\epsilon' n, \qquad n \geqq 1. \tag{14}$$

Taking c_k to be a number which satisfies (14), we prove (11) by induction on n. By (9)

$$R_k(n) = G_k(n), \qquad 1 \leqq n < 2^k,$$

which verifies (11) for all $n < 2^k$, by virtue of (13) and (14). Assume now that $n \geqq 2^k$ and that $|R_k(a)| \leqq c_k a$ for each positive integer $a < n$ (induction hypothesis). From (9) one obtains, on transposing to the right, the terms of the sum for which $m > 1$ and taking absolute values,

$$| R_k(n) | \leqq \sum_{1 < m \leqq \sqrt[k]{n}} \left| R_k\left(\frac{n}{m^k}\right) \right| + | G_k(n) |. \tag{15}$$

In the (nonvacuous) summation on the right, $n \geqq m^k > 1$ so that

$$1 \leqq [n/m^k] \leqq n/m^k < n,$$

which makes the induction hypothesis applicable to the terms of this sum. From this observation, (12), (13), and (14), it results that

$$| R_k(n) | \leqq c_k \sum_{1 < m \leqq \sqrt[k]{n}} [n/m^k] + c_k\epsilon' n \leqq c_k n \sum_{1 < m \leqq \sqrt[k]{n}} \frac{1}{m^k} + c_k\epsilon' n$$

$$< c_k n \sum_{m=2}^{\infty} \frac{1}{m^k} + c_k\epsilon' n \leqq c_k n \left(\sum_{m=2}^{\infty} \frac{1}{m^2} + \epsilon' \right) = c_k n.$$

It can be shown that the O-constant in (1) can be chosen independently of k, or in other words, the constant c_k can be chosen to have the same value for all $k > 1$. To see this, observe that $1 < \zeta(k) \leqq \zeta(2)$ if $k > 1$. The second half of this inequality is all that is needed for verifying that the O-constants arising from the computation leading to (10) are independent of k. The first inequality suffices to show that α_k is a bounded function of k, and hence that multiplication by α_k does not alter the situation concerning the final O-constant c_k.

References

1. Eckford Cohen, The average order of the Euler ϕ-function, submitted to the Amer. Math. Monthly.

2. V. L. Klee, A generalization of Euler's ϕ-function, Amer. Math. Monthly, 55 (1948) 358–359.

3. P. J. McCarthy, On a certain family of arithmetic functions, Amer. Math. Monthly, 65 (1958) 586–590.

EXTRA DIVIDENDS FROM A CALCULUS PROBLEM

C. STANLEY OGILVY, Hamilton College

If $P(x, y)$ is a point in the first quadrant, what line through P, together with the positive coordinate axes, forms the right triangle that minimizes (I) the sum of the base and the altitude; (II) the hypotenuse?

Although the student of elementary calculus may prefer to handle (I) and (II) as separate problems, the work is made more systematic if we note that to minimize the hypotenuse is also to minimize a^2+b^2, where a and b are the intercepts. Therefore we consider both questions as special cases of the more general problem of minimizing $f(a, b)=a^n+b^n$, $n>0$, subject to the constraint $x/a+y/b=1$. Using the Lagrange multiplier rule, we have with

$$\phi(a, b, \lambda) = a^n + b^n + \lambda(x/a + y/b - 1),$$

$$\frac{\partial\phi}{\partial a} = na^{n-1} - \frac{\lambda x}{a^2} = 0$$

$$\frac{\partial\phi}{\partial b} = nb^{n-1} - \frac{\lambda y}{b^2} = 0$$

from which $x/y=(a/b)^{n+1}$. If we now eliminate y between this and the constraining equation, we obtain

$$x(a^n + b^n) = a^{n+1} \quad \text{or} \quad x^{n/n+1}(a^n + b^n)^{n/n+1} = a^n.$$

Similarly,

$$y^{n/n+1}(a^n + b^n)^{n/n+1} = b^n.$$

Adding the two last equations gives us, finally,

$$a^n + b^n = (x^{n/n+1} + y^{n/n+1})^{n+1},$$

from which we read off the solutions to (I) and (II) by putting $n=1$ and $n=2$, respectively.

Setting $a^n+b^n=c^n$ gives the more familiar form

$$x^{n/n+1} + y^{n/n+1} = c^{n/n+1} \qquad (1)$$

and suggests what the "dividend" is going to be. Once the constant c has been established for the given (x, y), the question naturally arises: What other (x, y) have different individual values a and b but the same c? The answer is just those points on the curve of (1). Furthermore the line L through (x, y) determining c must be *tangent* to the curve at (x, y). For suppose it were not. Then there would be a point P' on the curve through which a line could be drawn parallel to L but nearer to the origin than L, with intercepts $a'<a$, $b'<b$. But $(a')^n+(b')^n<c^n$ is impossible if P' is on the curve.

The dividend, then, is the interpretation of the constant in (1) and the consequences that follow for special n. When $n=1$ we have the parabolic segment, shown as the unbroken part of the curve in the figure, whose equation is

$$x^{1/2} + y^{1/2} = c^{1/2}, \tag{2}$$

with $a+b=c$. This says that the tangents to such an arc cut off intercepts whose sum is constant. If we eliminate radicals from (2), the result is

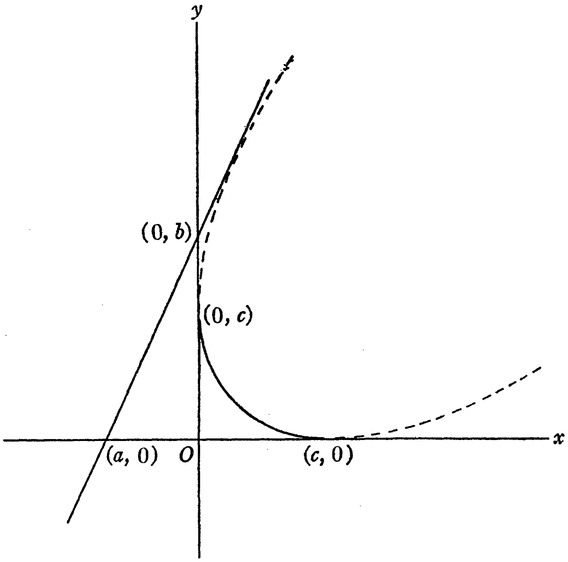
$$x^2 + 2xy + y^2 - 2cx - 2cy + c^2 = 0,$$

which has for its graph both the broken and unbroken parts of the curve. It is the *algebraic* sum of the intercepts cut off by a tangent to any point on the entire parabola that is equal to c.

If $n=2$, the curve is the hypocycloid of four cusps, with the well known property that the segment of the tangent line intercepted by the axes is of constant length c, because $a^2+b^2=c^2$. If $n=3$, equation (1) tells us that the curve $x^{3/4}+y^{3/4}=c^{3/4}$ has intercepts the sum of whose cubes is constant; and so on.

If $n<0$, the problem becomes that of maximizing the function $f(a, b)$ instead of minimizing it. All the calculations made for $n>0$ hold for $n<0$, except for the adjustments in wording required by the fact that the extremal is now a maximum instead of a minimum.

In summary, we have the simple result that the first quadrant tangents to the curve $x^m+y^m=c^m$ cut off intercepts a and b such that $a^n+b^n=c^n$, where $m=n/(n+1)$.



BOOK REVIEWS

EDITED BY DMITRI THORO, San Jose State College

Materials intended for review should be sent to: Dmitri Thoro, Department of Mathematics San Jose State College, San Jose, California 95114.

Introduction to Analysis and Abstract Algebra. By John E. Hafstrom. W. B. Saunders, Philadelphia, Pennsylvania, 1967. viii+339 pp. $6.50.

A student's first introduction to an axiomatic approach to mathematics requires a careful formulation of the notion of a proof. Before proceeding to a course in real variables for undergraduates, the author very nicely handles this problem by studying such topics as axioms for the ordered field of real numbers, inductive sets, completeness property for real numbers, denumerable and nondenumerable sets, equivalence relations and mappings as well as groups and fields. In the second half of the book the author takes up such topics as sequences, limits, continuity, integration and approximation of functions by polynomials. The topics in real analysis are, as the author points out, topics that every young mathematician should understand.

The author has very carefully considered the problem of how a student will begin to understand mathematics. Due to this careful thought, he carefully motivates theorems and definitions and quite often points out what type of proof will be used in a theorem, i.e., constructive, direct, indirect, by induction. This approach together with the numerous examples should be of great help to the student in understanding the material and working the numerous problems. The problems are often constructed to reinforce the intuitive notions that are formulated abstractly in definitions.

The book is primarily intended as a text for a one year course in real variable. I would like to use the first half of the book for a one semester course in an introduction to axiomatic mathematics. The second half of the book alone could easily be used as a text for a one semester course in real variable.

It would be helpful if there were more figures in the first half of the book. Also it would be helpful if there were a more liberal use of heavy type to set off axioms and theorems more noticeably.

GARY HAGGARD, San Jose State College

BRIEF MENTION

The Nature of Mathematics. By F. H. Young. Wiley, New York, 1968. xi+407 pp. $7.50.

"The general purpose of the text is to leave the student with a knowledge of at least some of the aims, techniques, and results of mathematics and with an appreciation of the role of mathematics in the world today." Included are chapters on the nature and results of mathematics in ancient times, properties of number systems (from the natural numbers to the reals), a typical development of mathematics from the concrete to the abstract and back to the concrete, residue classes and solutions of congruences, coordinate geometry, functions and relations, an introduction to calculus, matrices and determinants, digital computers. An appealing presentation which might be used for a wide variety of courses.

BOOK REVIEWS

EDITED BY DMITRI THORO, San Jose State College

Materials intended for review should be sent to: Dmitri Thoro, Department of Mathematics San Jose State College, San Jose, California 95114.

Introduction to Analysis and Abstract Algebra. By John E. Hafstrom. W. B. Saunders, Philadelphia, Pennsylvania, 1967. viii+339 pp. $6.50.

A student's first introduction to an axiomatic approach to mathematics requires a careful formulation of the notion of a proof. Before proceeding to a course in real variables for undergraduates, the author very nicely handles this problem by studying such topics as axioms for the ordered field of real numbers, inductive sets, completeness property for real numbers, denumerable and non-denumerable sets, equivalence relations and mappings as well as groups and fields. In the second half of the book the author takes up such topics as sequences, limits, continuity, integration and approximation of functions by polynomials. The topics in real analysis are, as the author points out, topics that every young mathematician should understand.

The author has very carefully considered the problem of how a student will begin to understand mathematics. Due to this careful thought, he carefully motivates theorems and definitions and quite often points out what type of proof will be used in a theorem, i.e., constructive, direct, indirect, by induction. This approach together with the numerous examples should be of great help to the student in understanding the material and working the numerous problems. The problems are often constructed to reinforce the intuitive notions that are formulated abstractly in definitions.

The book is primarily intended as a text for a one year course in real variable. I would like to use the first half of the book for a one semester course in an introduction to axiomatic mathematics. The second half of the book alone could easily be used as a text for a one semester course in real variable.

It would be helpful if there were more figures in the first half of the book. Also it would be helpful if there were a more liberal use of heavy type to set off axioms and theorems more noticeably.

GARY HAGGARD, San Jose State College

BRIEF MENTION

The Nature of Mathematics. By F. H. Young. Wiley, New York, 1968. xi+407 pp. $7.50.

"The general purpose of the text is to leave the student with a knowledge of at least some of the aims, techniques, and results of mathematics and with an appreciation of the role of mathematics in the world today." Included are chapters on the nature and results of mathematics in ancient times, properties of number systems (from the natural numbers to the reals), a typical development of mathematics from the concrete to the abstract and back to the concrete, residue classes and solutions of congruences, coordinate geometry, functions and relations, an introduction to calculus, matrices and determinants, digital computers. An appealing presentation which might be used for a wide variety of courses.

First Steps in Probability. By Meyer Dwass. McGraw-Hill, New York, 1967. viii+282 pp. $8.50.

". . . intended for a first course in mathematical probability theory. The prerequisite is a mastery of elementary algebra. . . . The main purpose of the book is to build, as quickly as possible, the concepts of sample space and probability measure, random variables, expected value, and frequency distribution for one or several variables. The theoretical high point is the law of large numbers." Also treated are regression and correlation for two random variables, special distributions, the central-limit theorem (without proofs), simple random walks and countable sample spaces. Contains especially attractive illustrations.

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles Valley College

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. 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.

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 February 1, 1969.

PROPOSALS

705. *Proposed by Max Rumney, London, England.*

Devise a method and a proof for the method of placing the natural numbers 1 to n^2 in a square array n by n so that: a) the column sums $\equiv 0 \pmod{n^2+1}$, b) the differences of the products of the extreme numbers of the diagonals of every square in the array $\equiv 0 \pmod{n^2+1}$, and c) the difference of the products of any two rows $\equiv 0 \pmod{n^2+1}$.

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

In Problem 594 [this MAGAZINE, March, 1966] it was shown that

$$AD + BE + CF \leqq 2R + 5r$$

where R is the circumradius, r the inradius, and AD, BE and CF the altitudes of the triangle ABC. Strengthen this inequality by showing $AD+BE+CF \leqq 2R+4r+2r^2/R$.

707. *Proposed by Joseph Malkewitch, University of Wisconsin.*

For what values of k is there a convex polygon with k sides which can be dissected into squares and equilateral triangles which have the same length of side?

First Steps in Probability. By Meyer Dwass. McGraw-Hill, New York, 1967. viii+282 pp. \$8.50.

". . . intended for a first course in mathematical probability theory. The prerequisite is a mastery of elementary algebra. . . . The main purpose of the book is to build, as quickly as possible, the concepts of sample space and probability measure, random variables, expected value, and frequency distribution for one or several variables. The theoretical high point is the law of large numbers." Also treated are regression and correlation for two random variables, special distributions, the central-limit theorem (without proofs), simple random walks and countable sample spaces. Contains especially attractive illustrations.

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles Valley College

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. 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.

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 February 1, 1969.

PROPOSALS

705. *Proposed by Max Rumney, London, England.*

Devise a method and a proof for the method of placing the natural numbers 1 to n^2 in a square array n by n so that: a) the column sums $\equiv 0 \pmod{n^2+1}$, b) the differences of the products of the extreme numbers of the diagonals of every square in the array $\equiv 0 \pmod{n^2+1}$, and c) the difference of the products of any two rows $\equiv 0 \pmod{n^2+1}$.

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

In Problem 594 [this MAGAZINE, March, 1966] it was shown that

$$AD + BE + CF \leqq 2R + 5r$$

where R is the circumradius, r the inradius, and AD, BE and CF the altitudes of the triangle ABC. Strengthen this inequality by showing $AD+BE+CF \leqq 2R+4r+2r^2/R$.

707. *Proposed by Joseph Malkewitch, University of Wisconsin.*

For what values of k is there a convex polygon with k sides which can be dissected into squares and equilateral triangles which have the same length of side?

708. *Proposed by Norman Schaumberger, Bronx Community College, New York.*

Prove that

$$\sum_{k=1}^{n-1} (n-k) \cos 2k\pi/n = -n/2.$$

709. *Proposed by H. S. Hahn, West Georgia College.*

Consider a convex body every point of whose surface is at a distance r from the surface of a regular tetrahedron with edges of length one. Find its surface area and its volume.

710. *Proposed by J. A. H. Hunter, Toronto, Canada.*

Solve the "no given digits" puzzle noting the decimal points and the repeating decimal in the quotient.

$$\begin{array}{l}
x\ x\ x\ x\ )\ x \cdot x\ x\ x\ x\ (\ \cdot \dot{x}\ x\ x\ x\ x\ x\ \dot{x} \\
\qquad\qquad\ \underline{x \cdot x\ x\ x\ x} \\
\qquad\qquad\qquad x\ x\ x\ x\ x \\
\qquad\qquad\qquad \underline{x\ x\ x\ x\ x} \\
\qquad\qquad\qquad\qquad x\ x\ x\ x \\
\qquad\qquad\qquad\qquad \underline{x\ x\ x\ x} \\
\qquad\qquad\qquad\qquad\qquad\quad x
\end{array}$$

711. *Proposed by Thomas Shewczyk, University of Wisconsin at Waukesha.*

If the numbers $a_1, a_2, \cdots a_n$ are positive, then show that

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} \frac{1}{a_i}\right) \geqq n^2$$

SOLUTIONS

Late Solutions

Eric Langford, Naval Postgraduate School, Monterey, California: **679** *(Two solutions).*

Erratum. In the solution to Problem 672 [May, 1968, page 162] the name of the proposer, R. S. Luthar, was omitted from the list of solvers.

C-Numbers

684. [*March,* 1968] *Proposed by J. David Martin, Washington State University.*

Define a C-number as a number having a common factor with all nonprime numbers smaller than itself. List all C-numbers, and prove that the list is exhaustive.

Solution by Michael Stolnicki, Oakland Community College, Michigan.

If n is a number such that $p_j^2 < n \leqq p_{j+1}^2$ where p_j is the jth prime, then for n to be a C-number it must have a common factor with the numbers 2^2, 3^2, 5^2, $\cdots$, p_j^2. The smallest such n is $2 \cdot 3 \cdot 5 \cdot \cdots \cdot p_j$. But if $p_j \geqq 7$ this contradicts

$n \leqq p_{j+1}^2$. (By Bertrand's postulate, $p_{j+1} < 2p_j$. Hence $p_{j+1}^2 < 4p_j^2 < 8p_jp_{j-1} < 2\cdot 3\cdot 5\cdot \cdots \cdot p_{j-1}\cdot p_j$ for $p_j \geqq 11$. Since $p_{j+1}^2 < 2\cdot 3\cdot 5\cdot \cdots \cdot p_j$ is also true for $p_j = 7$, we have it for all $p_j \geqq 7$.) Thus for n to be a C-number, $n \leqq 49$. The only such numbers are 2, 3, 4, 6, 8, 12, 18, 24, and 30. The integer 1 is a C-number trivially.

Also solved by Robert J. Bridgman, Mansfield State College, Pennsylvania; Mannis Charosh, New Utrecht High School, Brooklyn, New York; James M. Howard, Ferris State College, Michigan; Richard A. Jacobson, Houghton College, New York; Erwin Just, Bronx Community College, New York; J. D. E. Konhauser, Macalester College, Minnesota; Norbert J. Kuenzi, Iowa City, Iowa; Eugene McGovern, Ossining, New York; John J. Moore, Niagara Falls, New York; Prasert Na Nagara, Kasetsart University, Bangkok, Thailand; E. P. Starke, Plainfield, New Jersey; Robert P. Urbanski, McGill University; and the proposer.

Several solvers found the problem in "Elementary Number Theory," by Uspenski and Heaslet, Pages 89–90, and in "The Enjoyment of Mathematics," by Rademacher and Toeplitz, Pages 187–192.

Fermat's Principle

685. [*March*, 1968] *Proposed by Jack M. Elkin, Polytechnic Institute of Brooklyn.*

Prove Fermat's Principle for a circular mirror. That is, given two points, A and B, inside a circle, locate P such that $AP+PB$ is an extremum.

I. *Solution by Murray S. Klamkin, Ford Scientific Laboratory, Dearborn, Michigan.*

The problem is not formulated correctly. For the case of a constant refractive medium, Fermat's principle states that the path of light is such that $AP+PB$ is an extremum. Thus we will not be proving Fermats' principle by locating P such that $AP+PB$ is an extremum. Perhaps the proposer wishes to establish the law of reflection from Fermat's principle or conversely. In either case, this is a well-known problem and a simple solution follows by the use of level lines. Also, for greater generality, we can just as easily use an arbitrary convex closed curve with continuous curvature.

Consider the family of curves $AP+PB=k$ (constant). These are ellipses having A and B as foci. Clearly, the minimum occurs (possibly at more than one point) for the smallest ellipse of the family which is tangent to the given curve. The point (or points of tangency) will correspond to the minimizing point P. Then since the focal radii make equal angles with any tangent line, we obtain the law of reflection. Similarly, the maximum occurs for the largest ellipse of the family which is tangent to the given curve.

The converse theorem follows just as easily. If $\overline{AP}$ and $\overline{PB}$ make equal angles with the given curve, then the ellipse with foci A and B and which passes through P must be tangent to the given curve at P. Either the ellipse will be locally inside the given curve at P or outside of it. In the former case, $AP+BP$ will be a local minimum and in the latter case a local maximum.

II. *Comments by Leon Bankoff, Los Angeles, California.*

This is essentially the Billiard Problem of Alhazen (965–1039 A.D.), which appears as Problem 41 in Dörrie's "100 Great Problems of Elementary Mathematics" (Dover Reprint, N. Y., 1965). In its optical application, the problem is associated with Fermat's principle that "nature always acts by the shortest path." A solution to an analogous problem is given on Page 73 of the 1869 issue of the "Lady's and Gentleman's Diary," a source of reference somewhat less accessible than the work by Dörrie. The European reader may prefer to consult the French counterpart of Dörrie's book, "Célèbres Problèmes Mathématiques," by Edouard Callandreau, Éditions Albin Michel, Paris, 1949, Page 305, Problem 71, or Dörrie's German text, "Triumph der Mathematik," Physica-Verlag, Würzburg, 1958.

Scholarly enthusiasts who are not allergic to the dust of obscure library shelves may enjoy delving into Volume I of Leybourn's "Diary Questions," Pages 167–9, which gives three solutions originally published in the "Ladies' Diary" for 1727–1728.

Most of the published solutions involve one of the four intersections of the given circle with the equilateral hyperbola whose diameter is AB and whose ordinate axis is parallel to the line connecting the inverses of A and B with respect to the given circle.

One of the solutions in the 1869 "Diary" locates the point P as the point of tangency of the given circle with one of the family of confocal ellipses whose foci are A and B. In the proposed problem, both A and B lie within the circle. Hence the required ellipse lies entirely within the circle and touches the circle at the point P on the circumference for which $AP+BP$ is a minimum.

The location of P by means of conic sections precludes the possibility of a construction with Euclidean tools, except in the trivial case where A and B lie on a circle concentric with the given circle. In that case, P lies on the perpendicular bisector of the line joining A and B, and is easily found by ruler and compass or by a Mascheroni construction with compass alone.

An interesting sidelight mentioned in the 1869 "Diary" is that "this question occurs in the construction of steam boilers. The brace in the form of $A'P$, $B'P$, OP (where A' and B' are the inverses of A and B with respect to the given circle whose center is O) is stronger when the angle $A'PB'$ is bisected by OP."

Also solved by Michael Goldberg, Washington, D.C.; Lew Kowarski, Morgan State College, Maryland; and the proposer.

Little Fermat Theorem

686. [March, 1968] *Proposed by Stanley E. Payne, Miami University, Ohio.*

For each positive integer $m>1$ define the polynomial $f_m(x)$ by $f_m(x) = \sum\phi(d)x^{m/d}$, where ϕ is Euler's phi function and the sum is extended over all positive divisors of m. For any integer n, prove that

$$f_m(n) \equiv 0 \quad (\text{modulo } m).$$

Solution by E. Rosenthall, McGill University.

If a canonical form of m is $p_1^{e_1}p_2^{e_2} \cdots p_r^{e_r}$ then it suffices to prove

$$f_m(n) = 0 \ (\text{mod } p^a)$$

where p^a is any one of the $p_i^{e_i}$. We can write $m = p^a D$ where $(p, D) = 1$.

Then

$$\sum_{d/m} \phi(d) n^{m/d} = \sum_{d/D} \sum_{k=0}^{a} \phi(p^k d)(n^{D/d})^{p^{a-k}} = \sum_{d/D} \phi(d) A(d),$$

where, putting $\delta = D/d$,

$$\begin{aligned} A(d) &= \sum_{k=0}^{a} (p^k - p^{k-1})(n^\delta)^{p^{a-k}} \\ &= \sum_{k=0}^{a-1} p^k[(n^\delta)^{p^{a-k}} - (n^\delta)^{p^{a-k-1}}] + p^a n^\delta \\ &\equiv 0 \pmod{p^a}, \qquad \text{since } b^{p^t} \equiv b^{p^{t-1}} \pmod{p^t} \quad \text{for all integers } b. \end{aligned}$$

Also solved by Douglas Lind, University of Virginia; Peter A. Lindstrom, Union College, New York; Thomas E. Moore, University of Notre Dame; E. P. Starke, Plainfield, New Jersey; and the proposer.

Lind gave a reference to P. A. MacMahon ("Proceedings of the London Mathematical Society," 23, Pages 305–313).

Central Symmetry

687. [March, 1968] *Proposed by Sidney H. L. Kung, Jacksonville University, Florida.*

Prove that if the perimeter of a quadrilateral $ABCD$ is cut into two portions of equal length by all straight lines passing through a fixed point O in it, the quadrilateral is a parallelogram.

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

Through O, draw a line EF which does not pass through a vertex of the quadrilateral. Then take points at equal distances from E and draw lines through them and O. They meet the side containing F at equal distances from F since they are bisectors of the perimeter. But since the three equally spaced points at E are in perspective with the three equally spaced points at F, the two sides containing E and F are parallel. Similarly, the other two sides are parallel and the quadrilateral is a parallelogram with O as its center.

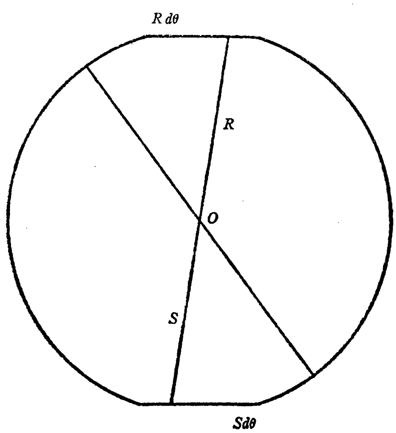
II. *Solution by Murray S. Klamkin, Ford Scientific Laboratory, Dearborn, Michigan.*

We consider a more general problem where we have a closed curve which is starlike with respect to the fixed point O and has the same perimeter property as the quadrilateral.

By the perimeter property,

$$Rd\theta = Sd\theta$$

or the curve must be centro-symmetric with respect to O. If C is a quadrilateral, it follows that it must then be a parallelogram.



Also solved by Prasert Na Nagara, Kasetsart University, Bangkok, Thailand; E. P. Starke, Plainfield, New Jersey; and the proposer (two solutions).

A Carlitz Congruence

688. [March, 1968] *Proposed by L. Carlitz, Duke University.*

Let n be an odd positive integer, a an arbitrary integer prime to n. Show that

$$\sum_{s=1}^{(n-1)/2} [1/2 + as/n] \equiv 0 \pmod 2$$

provided a is odd or $n \equiv \pm 1 \pmod 8$.

I. *Solution by the proposer.*

Put

$$as \equiv r_s \pmod n \qquad (|r_s| < \tfrac{1}{2}n),$$

so that

(*) $$as = r_s + k_s n,$$

where k is an integer. It follows that

$$-\frac{1}{2} + \frac{as}{n} < k_s < \frac{1}{2} + \frac{as}{n}$$

and therefore

$$k_s = \left[\frac{1}{2} + \frac{as}{n}\right].$$

Then from (*)

$$\sum_{s=1}^{\frac{1}{2}(n-1)} \left[\frac{1}{2} + \frac{as}{n}\right] = \sum_{s=1}^{\frac{1}{2}(n-1)} k_s = \frac{1}{n} \sum_{s=1}^{\frac{1}{2}(n-1)} (as - r_s)$$

$$\equiv (a+1) \sum_{s=1}^{\frac{1}{2}(n-1)} s \pmod 2,$$

since, except for order the numbers $|r_s|$ are identical with the numbers 1, 2, $\cdots$, $\frac{1}{2}(n-1)$. Thus

$$\sum_{s=1}^{\frac{1}{2}(n-1)} \left[\frac{1}{2} + \frac{as}{n}\right] \equiv \frac{1}{8}(n^2-1)(a+1) \pmod 2.$$

The right member is evidently congruent to 0 when a is odd or when $n \equiv \pm 1 \pmod 8$.

II. *Comment by E. P. Starke, Plainfield, New Jersey.*

The following points might be worth noting in connection with any solution. We write

$$S(a, n) = \sum_{s=1}^{(n-1)/2} \left[\frac{1}{2} + \frac{as}{n}\right].$$

Then

$$S(a, n) + S(n-a, n) = (n^2-1)/8.$$
$$S(a+n, n) - S(a, n) = (n^2-1)/8.$$
$$S(a, n) = \left[\frac{a}{2}\right]\left(\frac{n-1}{2}\right) - \sum_{r=1}^{[a/2]} \left[\frac{(2r-1)n}{2a}\right]$$

for $1 < a < n$.

The following is a simple restatement of Euler's criterion for the quadratic character of $a \pmod n$ if n is prime. However, if the symbol (a/n) is interpreted as the Jacobi symbol, then the statement is true for all odd n:

$$S(a, n) - \sum_{s=1}^{(n-1)/2} [as/n] \equiv \left(\frac{a}{n}\right) \pmod 2.$$

The proofs of these relations are simple and straightforward.

One incorrect solution was received.

Another Triangle Inequality

689. [March, 1968] *Proposed by Alexandru Lupas, Institutul De Calcul, Cluj, Romania.*

Let the lengths of the sides of a triangle be $a_1 \geqq a_2 \geqq a_3$. Also let r and R, respectively, denote the radii of the inscribed and circumscribed circle of an arbitrary triangle with angles B_i, $(0, \pi/2)$, $i = 1, 2, 3$. Then the following inequality holds

$$\sum_{i=1}^{3} (a_{i+1}a_{i+2})^{\cos B_i} \leqq 2 + 1/2 \sum_{i=1}^{3} a_i^2 - r/R$$

equality holding only when $a_1 = a_2 = a_3 = 1$ and $B_1 = B_2 = B_3$.

Solution by the proposer.

It is well known that the following inequality

$$x^y \leqq xy - y + 1, \qquad y \in (0, 1),$$

holds for x positive. Now, for $x = a_{i+1}a_{i+2}$, and $y = \cos B_i$, $B_i \in (0, \pi/2)$, it is easy to obtain

$$\text{(1)} \qquad \sum_{i=1}^{3} (a_{i+1}a_{i+2})^{\cos B_i} \leqq 3 + \sum_{i=1}^{3} (a_{i+1}a_{i+2})^{\cos B_i} - \sum_{i=1}^{3} \cos B_i.$$

On the other hand, taking into account that

$$\text{(2)} \qquad \sum_{i=1}^{3} \cos B_i = 1 + \frac{r}{R}$$

and

$$\text{(3)} \qquad \sum_{i=1}^{3} a_{i+1}a_{i+2} \cos B_i \leqq \frac{1}{2} \sum_{i=1}^{3} a_i^2, \quad \text{(P. Szász)},$$

we obtain from (1) — (3) the desired inequality.

Because in (1) equality holds only if $a_1 = a_2 = a_3 = 1$, and in (3) equality occurs for $B_i = A_i$, $i = 1, 2, 3$, that is B_i are the opposite angles of the sides a_1, a_2, a_3, we obtain the equality in this case.

A Recurrence Relation

690. [March, 1968] *Proposed by J. M. Gandhi, University of Alberta, Canada.*

Prove that

$$\sum_{j=0}^{\gamma-1} \sum_{i=0}^{j} j!/(j-i)!$$

$$= \begin{vmatrix} 0! & -\binom{1}{1} & 0 & 0 & \cdots & 0 \\ 1! & +\binom{2}{1} & -\binom{2}{2} & 0 & \cdots & 0 \\ 2! & -\binom{3}{1} & +\binom{3}{2} & -\binom{3}{3} & \cdots & 0 \\ — & — & — & — & \cdots & — \\ (\gamma-2)! & — & — & — & \cdots & -\binom{\gamma-1}{\gamma-1} \\ (\gamma-1)! & -1^{\gamma}\binom{\gamma}{1} & +(-1)^{\gamma-1}\binom{\gamma}{2} & — & \cdots & \binom{\gamma}{\gamma-1} \end{vmatrix}$$

Note. A number of typographical errors in the statement of the problem were undiscovered prior to printing. Thus, no solutions were received. However, the corrected proposal appears above and solutions are again solicited from readers.

A Long Probate

91. [March, 1951] *Proposed by Cyril Tobin, St. Francis Xavier University, Antigonish, Nova Scotia.*

An old man had three married daughters, Mrs. Jones, Mrs. Smith and Mrs. White. Each daughter had one and only one child. When the old man died, following the terms of his will, his fortune was distributed among his daughters, sons-in-law, and grandchildren in the following manner.

Each of the nine persons received a number of envelopes (the number of envelopes being different for each person). In each envelope was found a number of dollars equal to the number of envelopes which the person received. The share of money for each woman exceeded that of her husband by the amount that her husband's share exceeded that of their child. Although no two families received the same amount of money, all three families did receive the same number of envelopes. The total number of envelopes received by Mrs. Jones and Mrs. Smith was equal to the total number of envelopes received by Mrs. White and Mr. Jones.

Before the will was read all nine persons were penniless. After the administration, all were wealthy but not one of them was a millionaire. What were the nine sums of money?

I. *Solution by Leon Bankoff, Los Angeles, California.*

Let $(A>B>C)$, $(D>E>F)$, $(X>Y>Z)$ denote the number of envelopes received by the nine heirs grouped according to family, with specific family names to be assigned later.

The conditions may be represented symbolically as follows:

(1) $$0<(A\neq B\neq C\neq D\neq E\neq F\neq X\neq Y\neq Z)<1{,}000.$$

(2) $$A^2+C^2=2B^2;\qquad D^2+F^2=2E^2;\qquad X^2+Z^2=2Y^2.$$

(3) $$A+B+C=D+E+F=X+Y+Z.$$

(4) No. envelopes (Mrs. J. + Mrs. S.) = No. envelopes (Mrs. W. + Mr. J).

We note that the condition involving the inequality of the amounts of money received by each family is redundant since $A^2+B^2+C^2=3B^2$; $D^2+E^2+F^2=3E^2$; $X^2+Y^2+Z^2=3Y^2$ and, by (1), $3B^2\neq 3E^2\neq 3Y^2$.

The approach employed here is to obtain all solutions satisfying (1), (2), and (3) and to discard those that do not comply with (4). Although a fairly complete discussion and solution of the equation $a^2+c^2=2b^2$, (with $a>b>c$), may be found in Dickson's "History of the Theory of Numbers," Chelsea, N. Y., 1952, Vol. II, Page 435, et seq., a simple derivation of the parametric solution is given here.

Writing the equation in the form

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

and solving as primitive Pythagorean triplets, we have the two solutions

$$(a+c)/2 = u^2 - v^2; \qquad (a-c)/2 = 2uv; \qquad b = u^2 + v^2,$$

or

$$(a+c)/2 = 2uv; \qquad (a-c)/2 = u^2 - v^2; \qquad b = u^2 + v^2,$$

where $u > v$, with u, v relatively prime positive integers of different parity. Note that $a^2+c^2=2b^2$ implies that a and c are of like parity while $(a+c)/2$ and $(a-c)/2$ need not be.

Solving these two sets of parametric equations for a, b and c, we find that all prime solutions of $a^2+c^2=2b^2$ are given by $a=2uv+u^2-v$; $b=u^2+v^2$; $c=\pm(u^2-v^2-2uv)$. From this general solution we must select the sign that renders c positive. For use in a later step of this solution, we observe that we have two different parametric representations for $a+b+c$, namely, $4uv+u^2+v^2$ and $3u^2-v^2$, depending on which sign is taken for c.

Now let $(A, B, C)=j(a, b, c)$; $(D, E, F)=k(d, e, f)$; and $(X, Y, Z)=h(x, y, z)$, where j, k and h are integers and where the lower case letters in parentheses represent distinct primitive solutions obtained by using different pairs of parameters corresponding to u and v above, say (r, s), (m, n), and (p, q), respectively.

For trial purposes, assign the minimum values to (r, s) and (m, n), namely, $r=2$, $s=1$, $m=3$, $n=2$, to obtain the two smallest primitive solutions, $a=7$, $b=5$, $c=1$, and $d=17$, $e=13$, $f=7$. In passing we note that any other pairs of acceptable parameters yield greater values for a, b, c and d, e, f.

We now have

$$\begin{aligned} A + B + C &= j(a+b+c) = 13j, \\ D + E + F &= k(d+e+f) = 37k, \\ X + Y + Z &= h(x+y+z) = 13j = 37k. \end{aligned}$$

Thus $X+Y+Z$ is divisible by 13 and by 37 and hence by 481. If the quotient is w, we have

$$X + Y + Z = 481w = h(4pq + p^2 + q^2) \quad \text{or} \quad h(3p^2 - q^2),$$

the latter two members differing according to the sign chosen for z in terms of p and q.

Solutions of $481w=h(4pq+p^2+q^2)$ within the specified bounds are found by letting $h=w$, thus obtaining $p=20$, $q=1$, and $p=15$. $q=4$. Neither solution is acceptable here because the expression $4pq+p^2+q^2$ was obtained by using a negative value for z, contrary to the requirements of (1).

The only integral solution of $481w=h(3p^2-q^2)$ consistent with the stated restrictions is attained by letting $w=q=3$ and $h=1$, yielding $m=22$ and $n=3$

as parameters for the generation of the X, Y, Z triad related to the other two. So $X+Y+Z=x+y+z=1443=13j=37k$, from which we find $j=111$ and $k=39$. We now have a complete solution

$$A = 777 \qquad D = 663 \qquad X = 607$$
$$B = 555 \qquad E = 507 \qquad Y = 493$$
$$C = 111 \qquad F = 273 \qquad Z = 343$$

Condition (4) is satisfied if we assign A, B, C to the Whites, D, E, F to the Smiths, and X, Y, Z to the Joneses.

The proof that this solution is unique involves laborious exhaustion of all other possibilities within the prescribed limits and the observation that none of them meets condition (4). For the convenience of the interested problemist, a complete tabulation of these partial solutions is given here, along with the parameters used to generate the primitive triads.

A	847	924	693	231	462	987	497	994	533	769
B	605	660	495	165	330	705	353	706	377	565
C	121	132	99	33	66	141	47	94	13	217
D	803	876	657	219	438	897	483	966	527	759
E	583	636	477	159	318	663	345	690	373	561
F	187	204	153	51	102	273	69	138	23	231
X	637	764	573	191	382	713	357	714	497	617
Y	533	596	447	149	298	617	303	606	355	533
Z	403	356	267	89	178	503	237	474	71	401
	(2, 1)		(2, 1)			(2, 1)	(17, 8)		(5, 2)	(23, 6)
	(7, 2)		(7, 2)			(4, 1)	(2, 1)		(18, 7)	(4, 1)
	(5, 4)		(17, 10)			(19, 16)	(10, 1)		(2, 1)	(22, 7)

II. *Solution by F. L. Miksa, Aurora, Illinois.*

Let W, H, and C be the number of envelopes and W^2, H^2, and C^2 be the number of dollars received by the wife, husband, and child, respectively, of one family. Then

$$W^2 - H^2 = H^2 - C^2 \quad \text{or} \quad W^2 + C^2 = 2H^2.$$

Place $W=x+y$ and $C=x-y$, whereupon $x^2+y^2=H^2$. The primitive solutions of this last equation are given by

$$x = m^2 - n^2, \qquad y = 2mn, \qquad H = m^2 + n^2,$$

where m and n are of different parity, $m > n$, and $(m, n)=1$. It follows that $W=m^2+2mn+n^2$, $C=m^2-2mn-n^2$, and the number of envelopes received by each family is $N=W+C+H=3m^2-n^2$. Since none of the beneficiaries became a millionaire, none of the values of C, H, W can exceed 999, so $N \leqq 2997$.

From a table of Pythagorean triangles, the possible values of W, C, and H can be written immediately. For example, the triplet $(x, y, H)=(3, 4, 5)$ gives

$(W, C, H) = (7, 1, 5)$. Thus we write the 128 primitive solutions for which $N < 2997$ and C, H, W do not exceed 999. Since there are no duplicate values of N in this set, if a solution exists at least two of the values of N must be derived from primitive solutions. The last phase of the work consists in testing for a group of three solutions such that the number of envelopes received together by Mrs. Jones and Mrs. Smith is equal to the number of envelopes received together by Mrs. White and Mr. Jones. A solution follows:

	Jones	*Smith*	*White*		*Jones*	*Smith*	*White*
W	607	663	777	W^2	368449	439569	603729
H	493	507	555	H^2	243049	257049	308025
C	343	273	111	C^2	117649	74529	12321
	1443	1443	1443		729147	771147	924075

Also solved by C. W. Trigg, San Diego, California; and the proposer.

Comment on Problem 671

671. [November, 1967, and May, 1968] *Proposed by A. Wilansky, Lehigh University.*

Let $t_1, t_2, \cdots, t_r$ be real numbers with $t_1+t_2+\cdots+t_r=0$ and let $\{x_n\}$ be a bounded sequence of real numbers. Show that if the $\lim_{n\to\infty}(t_1x_{n-r+1}+t_2x_{n-r+2}+\cdots+t_rx_n)$ exists, it must be zero.

Comment by Jerry L. Pietenpol, Cleveland, Ohio

In Solution I, the derived inequalities are valid only if $t_i \geqq 0$, $(i=1, 2, \cdots r)$ whereas the condition $\sum_i t_i = 0$ requires at least one t_i to be negative (except for the trivial case $t_1 = \cdots = t_r = 0$). Indeed, the stated inequalities would imply $S_n = 0$ for all $n > N$ which is clearly not true in general.

Comment on Q397

Q397. Determine

$$\lim_{n\to\infty} \frac{1}{n} \sqrt[n]{((n+1)(n+2)\cdots(n+n))}.$$

[*Submitted by Murray S. Klamkin*]

Comment by Eckford Cohen, Manhattan, Kansas

This limit can also be evaluated by applying a weak form of Stirling's formula. We may write

$$c_n \overset{\text{def.}}{=} \frac{1}{n}((n+1)(n+2)\cdots(n+n))^{1/n} = 4\left(\frac{a_{2n}^2}{a_n}\right)$$

where $a_n = \sqrt[n]{n!}/n$. It follows that

$$\lim_{n\to\infty} c_n = \frac{4}{e}$$

from the well-known result, $\lim_{n\to\infty} a_n = 1/e$. The latter result can be proved in a number of ways. For a simple proof based on the exponential function, we refer to S. Saks and A. Zygmund, "Analytic Functions," Chapter 7, Section 5.

Comment on Q426

Q426. Without using calculus, determine the least value of the function $f(x) = (x+a+b)(x+a-b)(x-a+b)(x-a-b)$, where a and b are real constants.

[*Submitted by Roger B. Eggleton*]

Comment by M. S. Klamkin, Ford Scientific Laboratory, Dearborn, Michigan.

A more direct solution can be obtained by noting that

$$\begin{aligned} f(x) &= (a + b + x)(a + b - x)(a - b + x)(a - b - x) \\ &= ((a + b)^2 - x^2)((a - b)^2 - x^2) \\ &= (a^2 + b^2 - x^2)^2 - 4a^2b^2 \end{aligned}$$

Thus the minimum is $-4a^2b^2$ for $x^2 = a^2 + b^2$.

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.

Q440. A lady made three circular doilies with radii of 2, 3, and 10 inches, respectively. She placed them on a circular table so that each doily touched the other two. If each doily also touched the edge of the table, what was its radius?

[*Submitted by Walter W. Howard*]

Q441. Find $\ln(1.0004)^{\pi}$ to six decimal places.

[*Submitted by Wray G. Brady*]

Q442. E. T. Bell ("Men of Mathematics") relates an amusing story of Descartes assigning his pupil, Catherine the Great, the famous Appollonian problem of constructing a circle tangent to three given circles. To vent his hidden contempt for her scholarly pretensions, he neglected to warn the poor girl that synthetic geometry should be used. She used his new analytic geometry and, supposedly, was led into a trap, the required solution of three simultaneous quadratics. On the contrary, show how an easy triumph was possible.

[*Submitted by Charles E. Maley*]

Q443. Take the digits of a number expressed in any given base and permute them in any order. Prove that the difference between the two numbers is divisible by a number one less than the base.

[*Submitted by Michael Garrick and Mrs. Jack Lochhead*]

(Answers on pages 247-248)

A noteworthy property of this metric which is not shared by the Euclidean metric is that the area of a proper subset of a given set may exceed that of the containing set, e.g., as illustrated below:

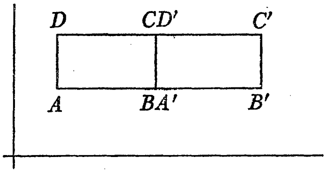


FIG. 3

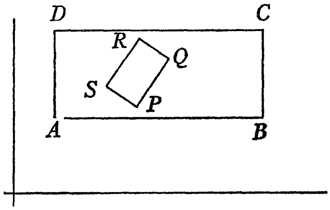


FIG. 4

Here $ABCD$ has area $\frac{1}{4}$ but $PQRS$ has area 1.

It is possible to create a class of measurable sets which preserve the additive property relative to ρ_2 in much the same way that such a class is defined relative to the Euclidean metric. There are indeed an infinite number of ways in which this can be done. We remark first that the area of the null set, and that of any discrete point or any finite collection of points is zero. The areas of rectangles with sides parallel to the x and y axes are $\frac{1}{4}$; the areas of any other rectangles are 1. If we consider then any finite collection of discrete points or rectangles which are not contiguous or overlapping this collection possesses an additive measure. The measure is defined as the sum of the areas of the individual rectangles or points making up the collection.

Reference

1. F. Hausdorff, Grundzüge der Mengenlehre, Leipzig, 1914, pp. 469–472.

ANSWERS

A440. If the 2, 3, and 10 inch doilies have centers at points C, A, and B, respectively, then these points form the vertices of a right triangle with sides 5, 12, and 13 inches. Complete the figure ABC into the rectangle with its fourth vertex denoted by letter O. From O, draw lines through B, C, and A, cutting the given circles in points P, Q, and R. Then $OP = OQ = OR = 15$ inches. This is the required radius. For a general solution, see Scripta Mathematica, vol. 21 (1955), pages 46–47.

A441. We have

$$\begin{aligned} \ln(1+x) &= x - x^2/2 + \cdots \\ &= .0004 - (0004)^2/2 + \cdots \\ &= .00040016 \end{aligned}$$

Also,

$$\begin{aligned} \ln(1.0004)^\pi &= \pi \ln(1.0004) \\ &= 3.141593(.000400) \\ &= .001257 \end{aligned}$$

A442. Suppose Catherine did not share with Descartes a prejudice against negative radii. Then the three quadratics

$$(h-h_1)^2+(k-k_1)^2=(r+r_1)^2$$
$$(h-h_2)^2+(k-k_2)^2=(r+r_2)^2$$
$$(h-h_3)^2+(k-k_3)^2=(r-r_3)^2$$

reduce to three linear equations of the form

$$2(h_1-h_2)h+2(k_1-k_2)k+2(r_1-r_2)r=(h_1^2-h_2^2)+(k_1^2-k_2^2)-(r_1^2-r_2^2)$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

Now simply substitute all possible combinations of signs for r_1, r_2, and r_3 in

$$\begin{bmatrix} h \\ k \\ r \end{bmatrix}=\frac{1}{2}\begin{bmatrix} h_1-h_2 & k_1-k_2 & r_1-r_2 \\ h_2-h_3 & k_2-k_3 & r_2-r_3 \\ h_3-h_1 & k_3-k_1 & r_3-r_1 \end{bmatrix}^{-1}\begin{bmatrix} (h_1^2-h_2^2)+(k_1^2-k_2^2)-(r_1^2-r_2^2) \\ (h_2^2-h_3^2)+(k_2^2-k_3^2)-(r_2^2-r_3^2) \\ (h_3^2-h_1^2)+(k_3^2-k_1^2)-(r_3^2-r_1^2) \end{bmatrix}$$

to obtain the eight possible circles (h, k), r.

A443. For the base b write the number as

$$\sum_{i=0}^{n} a_i b^{n-i}$$

and denote the permutation of the a_i by a_{ip}. We have for the difference of the two numbers

$$\sum_{i=0}^{n} a_i b^{n-i}-\sum_{i=0}^{n} a_{ip} b^{n-i}=\sum_{i=0}^{n}(a_i-a_{ip})b^{n-i}$$
$$=\sum_{i=0}^{n}(a_i-a_{ip})(b^{n-i}-1)+\sum_{i=1}^{n}(a_i-a_{ip}).$$

The last term is zero and, by induction, $b-1$ divides into the factor $b^{n-i}-1$.

(Quickies on page 295)

Correction for Spectral decomposition of matrices for high school students by Albert Wilansky, this MAGAZINE 41, 2 (1968) 51–59. It has been brought to my attention by Harry D. Ruderman and Paul Rosenbloom that the proof of Theorem 5 is not "immediate from Theorem 1 and the fact that the trace of I is 2." In fact it is conceivable that a resolution of the identity might have $p+2$ members; the trace of the sum would be $p+2$, i.e., 2. Here is a proof of Theorem 5: suppose $(A_1, A_2, A_3, \cdots A_k)$ is a resolution of the identity. Then $(A_1+A_2)^2=A_1^2+A_1A_2+A_2A_1+A_2^2=A_1+A_2$. Hence A_1+A_2 is idempotent. Thus $A_1+A_2=I$, by Theorem 1, since its trace is 2; and so $A_3+A_4+\cdots+A_k=0$. Multiply this equation by A_3, obtaining $A_3=0$. Similarly $A_4=\cdots=A_k=0$.