MATHEMATICS MAGAZINE

CONTENTS

CODEN: MAMGA8

VOLUME 46 • MAY 1973 • NUMBER 3

MATHEMATICS MAGAZINE

GERHARD N. WOLLAN, *EDITOR*

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

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

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

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 \$7.00. Members of the Mathematical Association of America and of Mu Alpha Theta may subscribe at the special rate of \$5.00. Single issues of the current volume may be purchased for \$1.40. Back issues may be purchased, when in print, for \$1.50.

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

POLYGONS IN ARRANGEMENTS GENERATED BY n POINTS

BRANKO GRÜNBAUM, University of Washington

1. Introduction. In 1826 Jacob Steiner [**11**] initiated the study of *arrangements of lines* in the plane, that is, of the various ways in which a finite set of lines may partition the plane. In its many ramifications the subject has received a large number of contributions since Steiner's days, and the interest in it was rekindled in recent years due to the discovery of its relevance to other topics in pure mathematics (such as zonohedra; compare Coxeter [**3**]) and in applied mathematics (such as pattern recognition, switching functions, etc.; see, for example, Winder [**12**], or Harding [**6**]). Recent surveys [**4**] and [**5**] provide summaries of the known results, present many open problems and conjectures, and give extensive bibliographies.

The aim of the present note is to establish a rather elementary result on arrangements which shows, among others, that unexpected facts can still be found even in very accessible fields.

Before the formulation of our result we shall introduce some definitions and notation. For simplicity of exposition we shall restrict the attention to the *real projective plane* P^2; the translation to the setting of the Euclidean plane presents no difficulties, but the formulation is somewhat awkward.

Let $\{L_1, L_2, \cdots, L_n\}$ be a family of straight lines in P^2, not all passing through one point. The *arrangement* $\mathscr{A}$ determined by the $n(\mathscr{A}) = n$ lines is the partition of P^2 into *vertices* (that is, intersection points of the L_j's), *edges* (segments of the L_j's between successive vertices), and *faces* (polygonal regions which are the connected components of the complement in P^2 of $\bigcup_{j=1}^{n} L_j$). Thus an arrangement of lines in P^2 is, essentially, a special type of cell complex decomposition of P^2, and therefore notions of isomorphic arrangements, etc., may be defined in the usual way. If F is a face of an arrangement $\mathscr{A}$, we shall denote by $p(F)$ the number of edges of $\mathscr{A}$ that form the boundary of F. It is easy to verify that always $p(F) \leqq n(\mathscr{A})$; equality in this relation is possible for each value of $n(\mathscr{A})$ (an example with $n(\mathscr{A}) = 7$ is shown in Figure 1), but if $n(\mathscr{A}) \geqq 5$ then $\mathscr{A}$ contains at most one face F such that $p(F) = n(\mathscr{A})$.

If a set V of $m = m(V)$ points is given in P^2, such that no line contains V, the arrangement $\mathscr{A}(V)$ *spanned* by V is the arrangement generated by all the lines determined by points of V. Thus $n(\mathscr{A}(V)) \leqq \binom{m(V)}{2}$, with equality if and only if no three points of V are collinear; it may easily be shown (compare Laisant [**7**]) that the total number of faces of $\mathscr{A}(V)$, where $m(V) = m$, is at most $(m^4 - 6m^3 + 23m^2 - 22m + 8)/8$. We are interested in the question: What is the largest possible number of sides for a face of an arrangement $\mathscr{A}(V)$ spanned by a set of $m = m(V)$ points? The answer is given by the following result.

THEOREM. *If F is a face of an arrangement $\mathscr{A}(V)$ spanned by a set $V = \{v_1, \cdots, v_m\}$ of $m = m(V)$ points, then*

(i) *$p(F) \leqq m$ if m is odd;*

(ii) $p(F) \leqq m-1$ *if* m *is even*;
both inequalities are best possible.

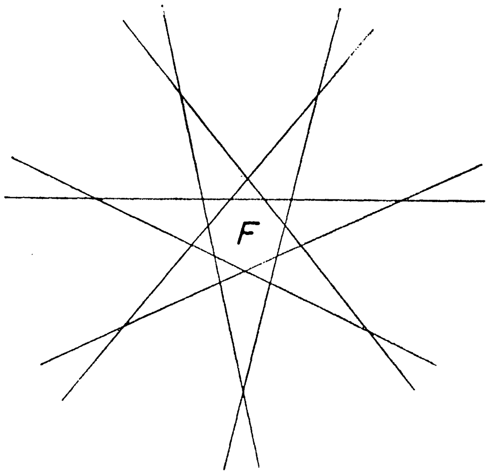


FIG. 1.

2. Proof of the theorem. We shall interpret P^2 as the extended real affine plane, taking as line at infinity a line that misses F and contains none of the points v_j. Choosing for V the vertices of a regular m-gon if m is odd, or all but one of the vertices of a regular $(m+1)$-gon if m is even, we obtain examples which show equality to be possible in the estimates of the theorem (the cases $n = 6$ and $n = 7$ are illustrated in Figure 2).

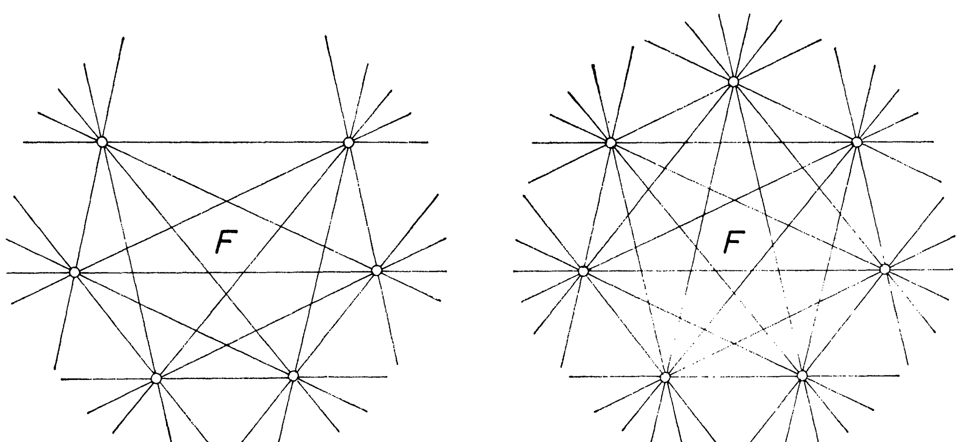


FIG. 2.

In order to prove the other assertions of the theorem we first observe that at most two lines through each v_j participate in the formation of F by containing one of the edges of F. Since each such line is counted at least from two of the v_j's it follows that $p(F) \leqq 2m(V)/2 = m(V)$, and that $p(F) = m(V)$ is possible only if precisely two lines through each point v_j are used, and each line passes through precisely two of the v_j's. This completes the proof of assertion (i), and we have only to prove that if $m = m(V)$ is even then $p(F) = m$ is impossible. To that end we shall investigate a few properties of arrangements $\mathscr{A}(V)$ for which there is a face F such that $p(F) = m$.

Let O be an interior point of F and let T_j denote the line determined by O and v_j; clearly $T_j \neq T_k$ for $j \neq k$. We claim:

(*) If $v_j, v_k \in V$ are on a line determined by an edge E of F then T_j and T_k are neighboring lines in the pencil $\{T_i\}$ of lines through O. Moreover, E does not belong

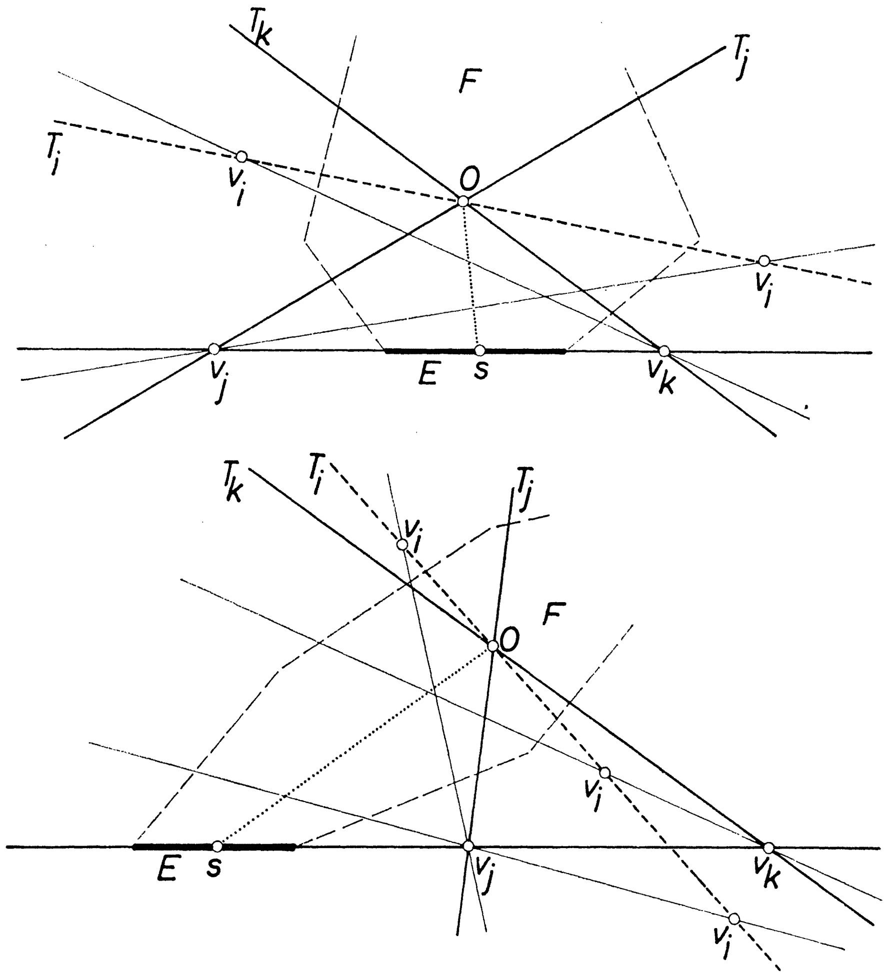


FIG. 3.

Illustration of the proof of assertion (*). The two cases differ in the mutual relation of v_j, v_k and E. In each case, a few typical positions of v_i are indicated.

to that pair of opposite angles determined at O by T_j and T_k which is free from other lines T_i.

The proof of this assertion can be read off at once from Figure 3; in it s is a relatively interior point of the edge E, and thus the segment Os is in F. If a line T_i would pass through the "forbidden" angle a contradiction would result, since the segment Os would be intersected either by the line v_iv_j or by the line v_iv_k.

Using (*) we may arrange the notation so that the lines $T_1, T_2, \cdots, T_m$ follow each other in the clockwise sense around O. We relabel, if needed, the points v_j so that $v_j \in T_j$, and we denote by E_j the edge of F which is collinear with v_j and v_{j+1}, by H_j the line containing E_j, v_j and v_{j+1}, and by s_j a relatively interior point of E_j. We observe:

(**) In the underlying affine plane v_{j-1} is (on H_{j-1}) between v_j and E_{j-1} if and only if v_{j+1} is (on H_j) between v_j and E_j.

Indeed, if v_{j-1} were between v_j and E_{j-1} but v_{j+1} not between v_j and E_j then, as indicated in Figure 4, the line $v_{j-1}v_{j+1}$ of $\mathscr{A}(V)$ would intersect F since it would intersect the segment Os_j. Similarly in case the roles of v_{j-1} and v_{j+1} are interchanged.

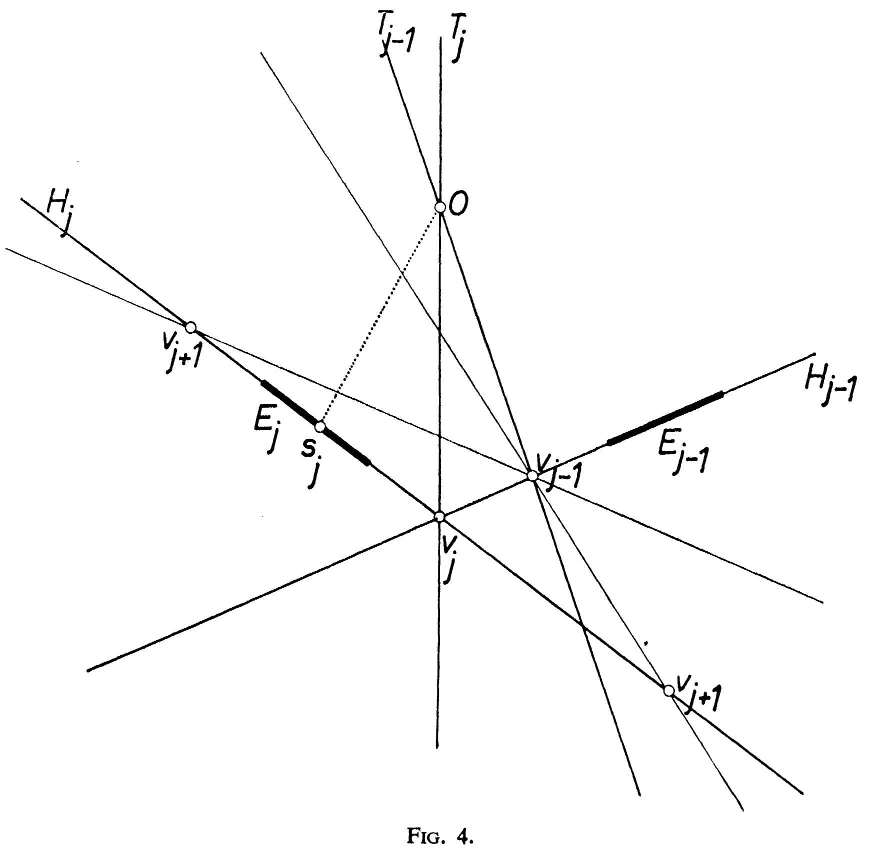


FIG. 4.

Illustration of the proof of assertion (**). Two typical positions of v_{j+1} are indicated.

We shall now designate, for each T_j, one of the halflines determined by O as T_j^+, the other as T_j^-. In order to determine the assignment of these labels on T_j we consider v_{j-1}, E_{j-1}, and v_j; if v_{j-1} does not separate E_{j-1} from v_j (on H_{j-1}) then the halfline of T_j containing v_j is designated T_j^+, the other T_j^-. On the other hand, if v_{j-1} separates E_{j-1} from v_j on H_{j-1} then the halfline of T_j containing v_j becomes T_j^-, and the other halfline T_j^+ (compare Figure 5).

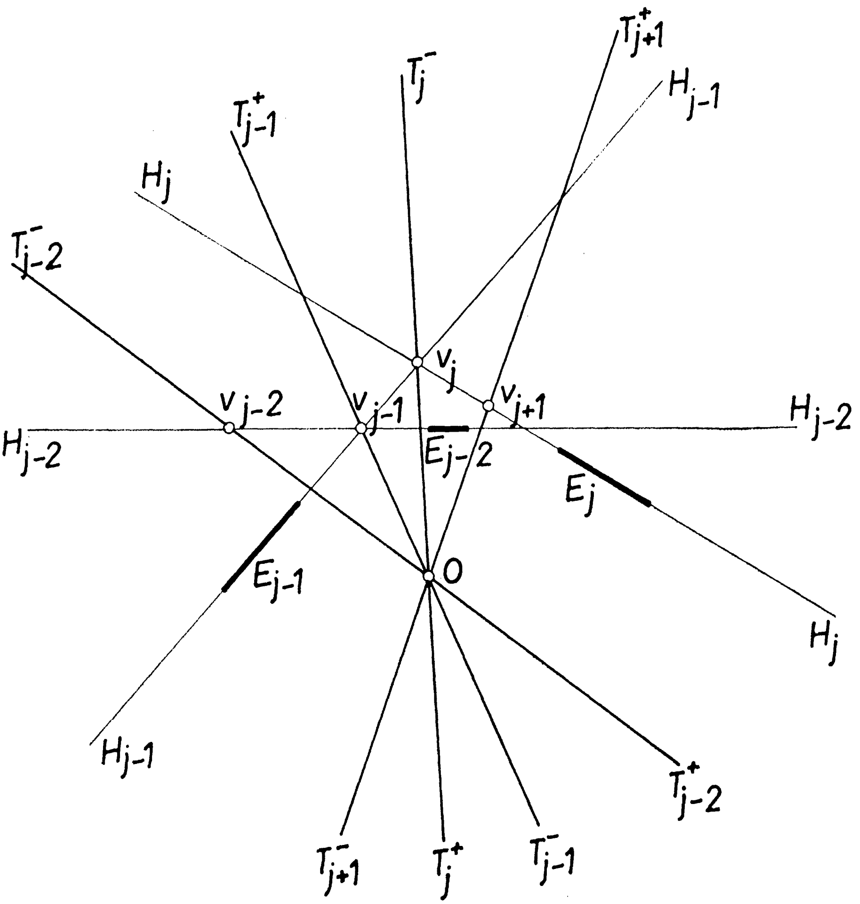


FIG. 5.

With this notation it follows at once from (**) that the two rays neighboring each T_j^+ are T_{j-1}^- and T_{j+1}^-. Therefore the total number of lines T_j is necessarily odd, and $p(F) = m(V)$ is impossible if $m(V)$ is even. This completes the proof of the theorem.

3. Remarks and problems.

(a) Our use of the drawings in the above proof could clearly be avoided,—but there seems to be no reason to do so. Quite the opposite—the illustrations make it quite obvious that the essential tools used in the proof are the polygonal case of the Jordan curve theorem and the order properties of the affine plane. Therefore it follows at once that the theorem remains valid (and the proof unchanged) if P^2 is replaced by the projective plane over any ordered field.

(b) By an *arrangement of pseudolines* $L_1, L_2, \cdots, L_n$ we mean the decomposition of P^2 into "vertices", curvilinear "edges", and "polygonal regions" effected by the simple closed curves L_j, $1 \leqq j \leqq n$, each two of which cross each other at precisely one point. (For more details and for references see Chapter 3 of [5].) An arrangement $\mathscr{A}(V)$ of pseudolines is *spanned* by a set V of points provided for each pair of points of V there is a pseudoline of the arrangement containing them, and each pseudoline contains at least two points of V. It is easily seen that our theorem may be generalized to arrangements $\mathscr{A}(V)$ of pseudolines spanned by a set V of points. The proof given above applies in this case as well; the only additional tool needed is Levi's "enlargement lemma" (see [8] or [5, page 47]):

Given an arrangement of pseudolines $L_1, \cdots, L_n$ and points v_1, v_2 which do not both lie on any one of the pseudolines L_j, there exists an arrangement of pseudolines $L, L_1, \cdots, L_n$ such that v_1 and v_2 are in L.

(c) Our theorem is a strengthening of the result formulated as Problem 20 in [10], which actually provided the motivation for the present note: In a convex polygon all diagonals are drawn, decomposing the polygon into smaller ones. What is the greatest number of sides possible for one of the smaller polygons, given the number of sides of the original polygon? As shown by remark (b) above, the answer to this problem would remain unchanged even if curvilinear "diagonals" were allowed, as long as no two diagonals have more than one point in common.

(d) If $\mathscr{A}$ is an arrangement of n lines and if $F_1, \cdots, F_r$ are distinct faces of $\mathscr{A}$, then a result of Canham [1] asserts that $\sum_{j=1}^{r} p(F_j) \leqq n + 2r(r-1)$, with equality possible for all r and all $n \geqq 2r(r-1)$. In particular, for $r = 2$ we have $p(F_1) + p(F_2) \leqq n + 4$, a result due to N. G. Gunderson (see [2]). It would be of interest to find analogues of Canham's result for arrangements $\mathscr{A}(V)$ spanned by sets of $m = m(V)$ points. In the special case $r = 2$ we conjecture that $p(F_1) + p(F_2) \leqq [3(m+1)/2]$ for $m \geqq 5$ is the best possible estimate.

(e) A family of simple closed curves $C_1, \cdots, C_n$ in the Euclidean plane is an *arrangement of curves* provided each two C_j's intersect and cross each other precisely twice (compare [5, Chapter 3]). Arrangements of curves are very similar to arrangements of lines or pseudolines, although significant differences exist; some of the differences are related to the possibility of *digonal faces*, the boundary of which consists of only two "edges". A recent result of Meyer [9] establishes that if F is a face of an arrangement of n curves then $p(F) \leqq 2n-2$, and if the arrangement contains no digonal faces then even $p(F) \leqq 2n-4$. An arrangement of curves $C_1, \cdots, C_n$ is *spanned* by a set V of $m = m(V)$ points provided each three points of V belong to some curve C_j, and each C_j contains at least three points from V. It would be of interest to find, in analogy to Meyer's result, estimates for $p(F)$ in terms of m for arrangements of curves spanned by sets of m points, or for those among them that contain no digonal faces.

This research supported in part by the Office of Naval Research under Grant N00014-67-A-0103-0003.

References

1. R. J. Canham, A theorem on arrangements of lines in the plane, Israel J. Math., 7 (1969) 393–397.

2. W. B. Carver, The polygonal regions into which a plane is divided by n straight lines, Amer. Math. Monthly, 48 (1941) 667–675.

3. H. S. M. Coxeter, The classification of zonohedra by means of projective diagrams, J. Math. Pures Appl., (9) 41 (1962) 137–156.

4. B. Grünbaum, Arrangements of hyperplanes, Proc. Second Louisiana Conference on Combinatorics, Graph Theory and Computing, Baton Rouge, 1971, pp. 41–106.

5. ———, Arrangements and Spreads, CBMS Regional Conference Series in Mathematics, Number 10., Amer. Math. Soc., 1972.

6. E. F. Harding, The number of partitions of a set of N points in k dimensions induced by hyperplanes, Proc. Edinburgh Math. Soc., 15 (1967) 285–289.

7. C.-A. Laisant, Remarques sur la théorie des régions et des aspects, Bull. Soc. Math. France, 10 (1881–82) 52–55.

8. F. Levi, Die Teilung der projektiven Ebene durch Gerade oder Pseudogerade, Ber. Math. -Phys. Kl. Sächs. Akad. Wiss. Leipzig, 78 (1926) 256–267.

9. W. Meyer, Maximum face size in an arrangement of curves. (To appear.)

10. D. O. Šklyarskiĭ, N. N. Čencov and I. M. Yaglom, Geometric Inequalities and Problems on Maxima and Minima, (in Russian), Nauka, Moscow, 1970.

11. J. Steiner, Einige Gesetze über die Theilung der Ebene und des Raumes, J. Reine Angew. Math., 1 (1826) 349–364.

12. R. O. Winder, Partitions of n-space by hyperplanes, SIAM J. Appl. Math., 14 (1966) 811–818.

SETTING THE HANDICAP IN BILLIARDS. A NUMERICAL INVESTIGATION

MARCEL F. NEUTS, Purdue University

Abstract. This paper illustrates a frequently occurring phenomenon in the numerical solution of mathematical problems. An analytic solution, even where explicitly available, may be too cumbersome for effective numerical computation. On the other hand a careful examination of the structure of the problem, keeping the capabilities of modern computers in mind, may reveal recurrence relations which may be conveniently programmed to yield the required numerical answers.

In the specific problem considered here we show as an illustration, how the handicap may be set in a billiards game between two players of unequal skills to make the game approximately even.

Introduction. The literature on applied mathematics and probability abounds with so-called explicit solutions to a large number of specific problems. The solution is explicit, only in the sense that in principle a formula can be given to express the required quantity in terms of the given data. As the problems become more compli-

This research supported in part by the Office of Naval Research under Grant N00014-67-A-0103-0003.

References

1. R. J. Canham, A theorem on arrangements of lines in the plane, Israel J. Math., 7 (1969) 393–397.

2. W. B. Carver, The polygonal regions into which a plane is divided by *n* straight lines, Amer. Math. Monthly, 48 (1941) 667–675.

3. H. S. M. Coxeter, The classification of zonohedra by means of projective diagrams, J. Math. Pures Appl., (9) 41 (1962) 137–156.

4. B. Grünbaum, Arrangements of hyperplanes, Proc. Second Louisiana Conference on Combinatorics, Graph Theory and Computing, Baton Rouge, 1971, pp. 41–106.

5. ———, Arrangements and Spreads, CBMS Regional Conference Series in Mathematics, Number 10., Amer. Math. Soc., 1972.

6. E. F. Harding, The number of partitions of a set of *N* points in *k* dimensions induced by hyperplanes, Proc. Edinburgh Math. Soc., 15 (1967) 285–289.

7. C.-A. Laisant, Remarques sur la théorie des régions et des aspects, Bull. Soc. Math. France, 10 (1881–82) 52–55.

8. F. Levi, Die Teilung der projektiven Ebene durch Gerade oder Pseudogerade, Ber. Math. -Phys. Kl. Sächs. Akad. Wiss. Leipzig, 78 (1926) 256–267.

9. W. Meyer, Maximum face size in an arrangement of curves. (To appear.)

10. D. O. Šklyarskiĭ, N. N. Čencov and I. M. Yaglom, Geometric Inequalities and Problems on Maxima and Minima, (in Russian), Nauka, Moscow, 1970.

11. J. Steiner, Einige Gesetze über die Theilung der Ebene und des Raumes, J. Reine Angew. Math., 1 (1826) 349–364.

12. R. O. Winder, Partitions of *n*-space by hyperplanes, SIAM J. Appl. Math., 14 (1966) 811–818.

SETTING THE HANDICAP IN BILLIARDS. A NUMERICAL INVESTIGATION

MARCEL F. NEUTS, Purdue University

Abstract. This paper illustrates a frequently occurring phenomenon in the numerical solution of mathematical problems. An analytic solution, even where explicitly available, may be too cumbersome for effective numerical computation. On the other hand a careful examination of the structure of the problem, keeping the capabilities of modern computers in mind, may reveal recurrence relations which may be conveniently programmed to yield the required numerical answers.

In the specific problem considered here we show as an illustration, how the handicap may be set in a billiards game between two players of unequal skills to make the game approximately even.

Introduction. The literature on applied mathematics and probability abounds with so-called explicit solutions to a large number of specific problems. The solution is explicit, only in the sense that in principle a formula can be given to express the required quantity in terms of the given data. As the problems become more complicated or advanced the explicit solution commonly reduces to an explicit *asymptotic* one only.

Many a student of mathematics is therefore left with an utterly truncated notion of what constitutes a *solution* to a specific problem. Even if he has remained fascinated by the almost universal applicability of mathematics, he tends to consider a problem solved if an explicit formula can be found and if some qualitative information is drawn from it through limit and approximation theorems.

The numerical analysis of the problem is often ignored or is left to the initiative of people "interested in numbers". Even assuming that the problem is of sufficient interest to attract the attention of a qualified numerical analyst, the latter is faced with a delicate task. If he does not have the time or the background to acquaint himself fully with the analytic derivation, he will often do a considerable amount of work to program an analytic expression that is utterly ineffective for numerical work. He may also use an inordinate amount of computer time to guard against such computer hazards as rounding error, even when they are unlikely to cause any problems.

Very often, if the person solving the problem analytically is also well versed in the use of the computer, he will recognize that *the structure* of the problem, rather than the formula-solution may be used to obtain a numerical solution.

It is the purpose of this paper to present an illustrative example pertaining to the foregoing discussion. Apart from its interest in the analysis of a popular game, our discussion provides an example of a problem whose analytic solution is beyond the level of the classical probability texts, yet whose numerical solution may well be presented to and stimulate the interest of an undergraduate audience.

We consider the classical game of billiards. Players successively get turns at trying to score points. A *turn* lasts as long as a player scores consecutive successes. If he fails to make a certain shot, the next player gets a turn and so on in a fixed order of the players. The winner is that player who scores a certain number of points before his opponents do. We shall limit our attention to a game with *two players* only. The extension to more players is routine.

We consider a game with players I and II. We assume that the successive attempts to score points form a sequence of Bernoulli trials and that player I has probability p_1, $0 < p_1 < 1$ of scoring a point at each of his trials; p_2 is the corresponding probability for player II, $0 < p_2 < 1$.

In order to win the game, player I must accumulate n_1 points *before* player II accumulates n_2. In the alternate case player II wins.

It clearly matters which player goes first. We shall therefore denote by $P_1(n_1, n_2)$ the probability that player I wins, given that he gets the first turn. Similarly $P_2(n_1, n_2)$ will denote the probability that player II wins, given that he gets the first turn.

In some cases a toss up is performed to decide who goes first. In particular, in the sequel we shall denote by $P(n_1, n_2)$ the probability that player I wins, if he starts the game with probability 1/2 and his opponent starts with probability 1/2.

Clearly we have that

$$P(n_1, n_2) = \tfrac{1}{2}[P_1(n_1, n_2) + 1 - P_2(n_1, n_2)] \tag{1}$$

for all n_1 and n_2.

The problem of setting *the handicap* is the following. For given values of p_1 and p_2, we wish to determine the values of n_1 and n_2 for which $P_1(n_1, n_2)$ or $P_2(n_1, n_2)$ or $P(n_1, n_2)$ is approximately 0.5.

Without loss of generality we may assume that $p_1 \geqq p_2$, i.e., that the better player is designated as player I. In reality there may be further restrictions. Often the number n_1 is fixed by the rules of the game and the question is then to determine how many points the better player should "give" to the weaker one to make the game about even.

The most detailed discussion of the probability aspects of the game of billiards was given in two papers in Dutch by O. Bottema and S. C. Van Veen (*Kansbere keningen bij het biljartspel I, II*, Nieuw Archief voor Wiskunde, 22, 1943, 16–33 and 123–158).

They obtained involved exact and also approximate expressions for the probabilities $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ in terms of certain hypergeometric series and their limiting formulas.

They also discuss the interesting variant in which both players always get *the same number of turns*. If the player who goes first reaches n_1 or n_2, as the case may be, before his opponent does, the latter gets one more turn. If he also completes his allotted number of points the game is a draw. The numerical investigation of the handicap problem is somewhat more involved for this case and we will not consider it here.

A number of interesting random variables, such as the duration of play, the number of turns per match and the longest run of successful shots for each player, were also investigated by these authors. We shall restrict our attention however to the problem of selecting the proper handicap.

In order to set up the handicap we must determine those values of n_1 and n_2 for which one of the probabilities $P_1(n_1, n_2)$, $P_2(n_1, n_2)$ or $P(n_1, n_2)$ is close to the value 0.5. Which of the three probabilities is to be used in a specific game depends on the selection rule for the first turn.

The quantities $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ were expressed explicitly in terms of p_1, p_2, n_1 and n_2 by Bottema and Van Veen. The resulting expressions are complicated series, which are of little help or no help in determining the handicap. They would require delicate computer programming; moreover a large number of such expressions would have to be evaluated, which would lead to the use of a considerable amount of computer time.

The authors referred to also obtained approximate formulas for the probabilities of interest. In spite of the delicate derivation of these approximations, little of additional use in numerical work is gained, since we are generally interested in fairly small values for n_1 and n_2 where the approximations are still rather crude.

The correct numerical values of n_1 and n_2 may be obtained by elementary means.

We shall reexamine the structure of the problem and obtain recurrence relations from which the desired n_1 and n_2 may be found by a simple algorithm.

We note the following geometric representation of the problem. If X_k and Y_k denote the numbers of points accumulated by I and II respectively in the first k attempts, then the point (X_k, Y_k) performs a simple type of random walk on the lattice points of the nonnegative orthant in the plane (Figure 1). The walk starts

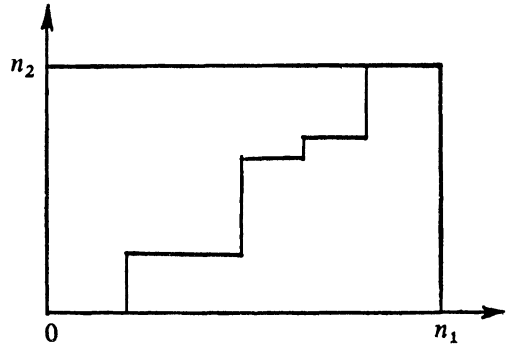


Fig. 1. $n_1=20$, $n_2=13$. A possible path of (X_k, Y_k).

at (0,0)— we set $X_0 = Y_0 = 0$—and proceeds horizontally or vertically by unit steps whenever a point is scored and depending on whether player I or II is playing. Whenever a failure occurs the walk remains in its previous position.

Player I wins if the set of points (n_1, m_2), $0 \leqq m_2 < n_2$ is reached eventually; player II wins if the set of points (m_1, n_2), $0 \leqq m_1 < n_1$ is reached eventually.

The recurrence relations. If player I starts, then we find by considering all possibilities after the first trial, that

$$P_1(n_1, n_2) = p_1 P_1(n_1 - 1, n_2) + q_1[1 - P_2(n_1, n_2)], \tag{2}$$

for $n_1 \geqq 1$, $n_2 \geqq 1$, and $q_1 = 1 - p_1$.

Similarly, if player II starts, then we find by the same argument that

$$P_2(n_1, n_2) = p_2 P_2(n_1, n_2 - 1) + q_2[1 - P_1(n_1, n_2)], \tag{3}$$

for $n_1 \geqq 1$, $n_2 \geqq 1$, and $q_2 = 1 - p_2$.

It is further obvious that we may set

$$P_1(0, n_2) = P_2(n_1, 0) = 1 \text{ and } P_1(n_1, 0) = P_2(0, n_2) = 0, \tag{4}$$

for $n_1 \geqq 1$ and $n_2 \geqq 1$.

We solve the equations (2) and (3) for $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ and obtain

$$\begin{aligned} P_1(n_1, n_2) &= p_1(1 - q_1 q_2)^{-1} P_1(n_1 - 1, n_2) \\ &\quad + q_1 p_2 (1 - q_1 q_2)^{-1}[1 - P_2(n_1, n_2 - 1)], \\ P_2(n_1, n_2) &= p_2(1 - q_1 q_2)^{-1} P_2(n_1, n_2 - 1) \\ &\quad + q_2 p_1 (1 - q_1 q_2)^{-1}[1 - P_1(n_1 - 1, n_2)]. \end{aligned} \tag{5}$$

The recurrence relations (5) are ideally suited for numerical computation. We see that the quantities $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ for which $n_1 + n_2 = k + 1$ are simply related to the corresponding string of quantities for which $n_1 + n_2 = k$. We so obtain an iterative scheme similar to *Pascal's triangle* for the binomial coefficients.

The numerical solution to the handicap problem. Depending on whether n_1 and n_2 are set before or after the player to go first is selected, we shall wish to find those pairs n_1 and n_2 for which $P_1(n_1, n_2)$, $P_2(n_1, n_2)$ or $P(n_1, n_2)$ are *close to* 0.5. Except for very special values of p_1 and p_2, we may not be able to find pairs (n_1, n_2) within the range of interest, for which these probabilities are *exactly* 0.5.

We therefore specify an interval (α, β), say (.495, .505), about 0.5 and write the computer program to print out the values of $n_1, n_2, P_1(n_1, n_2), P_2(n_1, n_2), P(n_1, n_2)$ *for which at least one of the latter three quantities belongs to* (α, β).

In order to make efficient use of computer memory storage, we compute all the quantities $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ for which $n_1 + n_2 = k + 1$ recursively from those for which $n_1 + n_2 = k$, by use of the recurrence relations (5), starting with $k = 1$, $P_1(1,0) = P_2(0,1) = 1 - P_1(0,1) = 1 - P_2(1,0) = 0$.

Only those $n_1, n_2, P_1(n_1, n_2)$, $P_2(n_1, n_2)$ and $P(n_1, n_2)$ for which one of the latter three quantities lies in (α, β) are stored in arrays for later printout.

In the examples computed by the author the bounds $\alpha = .495$ and $\beta = .505$ were chosen and k from 1 to 600. The latter bound is larger than required for most practical purposes. Even so, the central processing time on the CDC 6500 at Purdue was only 32 seconds (approx.) per set of values p_1 and p_2. With k up to 200, the central processing time was 7.5 seconds (approx.).

As far as the computer program is concerned, there is of course nothing particular about the values α and β chosen. If one wished to concentrate attention on any other subinterval of $[0, 1]$, it would suffice to choose the values of α and β accordingly. In particular for $\alpha = 0$ and $\beta = 1$ one would obtain a complete table of the three probabilities of interest.

Generalizations.

a. The claim is commonly made by billiard players that if a good player misses a shot, he is likely to leave a favorable configuration for his opponent at the next trial, whereas when a mediocre player misses a shot, the configuration that results is usually not particularly favorable to his opponent, or may even be markedly unfavorable.

To add an extra touch of realism to our model, we may assume that the probability of success for a player is p_i, $i = 1, 2$ at each trial, except for those trials that follow a missed shot for which the probability of success will be denoted by p_i', $i = 1, 2$. The first shot of the game is considered an "ordinary" shot for either player, i.e., the probability of a success at the first shot is p_i, $i = 1, 2$.

As before we set $1 - p_i = q_i$ and $1 - p_i' = q_i'$, with $0 < p_i < 1$, $0 < p_i' < 1$ for $i = 1, 2$. $P_1(n_1, n_2)$, $P_2(n_1, n_2)$ and $P(n_1, n_2)$ have the same interpretations as before.

The recurrence relations for $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ are now obtained as follows.

If player I goes first, he either makes the first point or else II and I will try alternatingly until one of them makes a point

The conditional probability that I makes the first *successful* shot, given that he starts is given by

$$(6)\qquad p_1 + q_1q_2'p_1' + q_1q_2'q_1'q_2'p_1' + q_1q_2'(q_1'q_2')^2p_1' + \cdots + q_1q_2'(q_1'q_2')^\nu p_1' + \cdots$$
$$= p_1 + q_1q_2'p_1'(1 - q_1'q_2')^{-1} = \theta_1 .$$

The conditional probability that player II makes the first *successful* shot is given by

$$(7)\qquad 1 - p_1 - q_1q_2'p_1'(1 - q_1'q_2')^{-1} = q_1[1 - q_2'p_1'(1 - q_1'q_2')^{-1}]$$
$$= q_1p_2'(1 - q_1'q_2')^{-1} = 1 - \theta_1 .$$

If player I makes the first successful shot the remaining game is similar to the original one, except that now $(n_1 - 1, n_2)$ points remain to be played. Similarly if player II makes the first successful shot.

The recurrence relations now become

$$(8)\qquad P_1(n_1,n_2) = \theta_1P_1(n_1 - 1,n_2) + (1 - \theta_1)[1 - P_2(n_1,n_2 - 1)]$$
$$P_2(n_1,n_2) = \theta_2P_2(n_1,n_2 - 1) + (1 - \theta_2)[1 - P_1(n_1 - 1,n_2)]$$

where

$$(9)\qquad \theta_2 = p_2 + q_2q_1'p_2'(1 - q_1'q_2')^{-1} .$$

The recurrence relations (8) are similar to those in formula (5). With a minor modification in the coefficients, the same computer program may be used to find the handicap.

We note that for $p_1' = p_1$ and $p_2' = p_2$ the formulae (8) reduce to those given in (5).

b. In an alternate generalization of the original model, we may assume that the probability of a success depends only on the number of consecutive successes scored during that turn. Specifically let $p_1(\nu)$ be the probability that player I has a success if he has already made exactly $\nu - 1$ successes during his current turn. Similarly we define $p_2(\nu)$ for the second player. We assume that $0 < p_1(\nu) < 1$ and $0 < p_2(\nu) < 1$ for all $\nu \geqq 1$.

The probabilities $P_1(n_1,n_2)$, $P_2(n_1,n_2)$ and $P(n_1,n_2)$ are defined as above. They no longer satisfy simple first order recurrence relations. Instead we obtain

$$(10)\qquad P_1(n_1,n_2) = q_1(1)[1 - P_2(n_1,n_2)]$$
$$+ \sum_{\nu=1}^{n_1} \prod_{r=1}^{\nu} p_1(r)q_1(\nu + 1)[1 - P_2(n_1 - \nu,n_2)],$$
$$P_2(n_1,n_2) = q_2(1)[1 - P_1(n_1,n_2)]$$
$$+ \sum_{\nu=1}^{n_2} \prod_{r=1}^{\nu} p_2(r)\, q_2(\nu + 1)[1 - P_1(n_1,n_2 - \nu)],$$

by conditioning on the number of points scored by the player who goes first, during his first turn. By solving for $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ in (10), we see that the quantities $P_1(n_1, n_2)$ and $P_2(n_1, n_2)$ for which $n_1 + n_2 = k + 1$ may be computed recursively in terms of those for which $n_1 + n_2 \leqq k$. Although the computer program to solve the handicap problem for this case is easy to write, substantially more memory storage and computation time are required to obtain the numerical answers for this model. In the absence of concrete data, the author has not performed any such computations. A listing of the FORTRAN IV program for the handicap problem and tables of the numerical values corresponding to the parameter values

$$p_1 = 0.1\ (0.2)\ 0.9 \qquad \alpha = 0.495$$

$$p_2 = 0.1\ (0.2)\ p_1 \qquad \beta = 0.505$$

$$n_1 + n_2 \leqq 600$$

may be obtained from the author upon request, by writing to Department of Statistics, Purdue University, West Lafayette, Indiana 47907.

Table 1. This table illustrates the numerical results for the handicap problem corresponding to:

$$p_1 = .85 \qquad p_2 = .75.$$

The numerical values of $n_1, n_2, P_1(n_1, n_2)$, $P_2(n_1, n_2)$ and $P(n_1, n_2)$ are given for all cases where at least one of the latter three probabilities belongs to the interval (α, β), where:

$$\alpha = .498 \qquad \beta = .502$$

and where $n_1 + n_2 \leqq 600$.

TABLE I

n_1	n_2	P_1	P_2	P	n_1	n_2	P_1	P_2	P	n_1	n_2	P_1	P_2	P
6	5	.740	.501	.619	8	6	.708	.501	.604	13	7	.589	.589	.500
16	7	.500	.663	.418	27	16	.616	.498	.559	29	17	.610	.501	.555
30	16	.556	.556	.500	31	15	.500	.612	.444	44	25	.592	.498	.547
45	24	.547	.544	.502	46	23	.501	.589	.456	46	26	.588	.500	.544
47	25	.544	.545	.500	48	24	.500	.589	.455	48	27	.585	.502	.541
49	26	.542	.545	.498	61	34	.578	.498	.540	62	33	.540	.537	.501
63	32	.501	.576	.463	63	35	.575	.500	.538	64	34	.538	.538	.500
65	33	.500	.576	.462	65	36	.573	.501	.536	66	35	.536	.539	.498
67	34	.498	.576	.461	78	43	.569	.498	.535	79	42	.535	.533	.501
80	41	.501	.567	.467	80	44	.567	.500	.534	81	43	.533	.534	.500
82	42	.499	.568	.466	82	45	.565	.501	.532	83	44	.532	.535	.499
84	43	.498	.568	.465	95	49	.502	.560	.471	95	52	.563	.499	.532
96	51	.532	.530	.501	97	50	.501	.561	.470	97	53	.561	.500	.531
98	52	.530	.531	.500	99	51	.499	.561	.469	99	54	.559	.501	.529
100	53	.529	.532	.499	101	52	.498	.562	.468	112	58	.502	.556	.473

TABLE I (Cont'd)

n_1	n_2	P_1	P_2	P	n_1	n_2	P_1	P_2	P	n_1	n_2	P_1	P_2	P
112	61	.558	.499	.530	113	60	.529	.527	.501	114	59	.501	.556	.472
114	62	.556	.500	.528	115	61	.528	.528	.500	116	60	.499	.557	.471
116	63	.555	.501	.527	117	62	.527	.529	.499	118	61	.498	.557	.471
118	64	.553	.502	.526	128	68	.529	.525	.502	129	67	.502	.552	.475
129	70	.554	.499	.528	130	69	.527	.526	.501	131	68	.501	.552	.474
131	71	.553	.500	.526	132	70	.526	.526	.500	133	69	.499	.553	.473
133	72	.551	.501	.525	134	71	.525	.527	.499	135	70	.498	.554	.472
135	73	.550	.502	.524	145	77	.527	.523	.502	146	76	.501	.549	.476
146	79	.551	.499	.526	147	78	.526	.524	.501	148	77	.500	.549	.476
148	80	.549	.500	.525	149	79	.525	.525	.500	150	78	.500	.550	.475
150	81	.548	.501	.524	151	80	.523	.526	.499	152	79	.499	.550	.474
152	82	.547	.502	.523	162	86	.525	.522	.502	163	85	.501	.546	.478
163	88	.548	.499	.525	164	87	.524	.523	.501	165	86	.500	.547	.477
165	89	.547	.500	.524	166	88	.523	.523	.500	167	87	.500	.547	.476
167	90	.546	.501	.523	168	89	.522	.524	.499	169	88	.499	.548	.475
169	91	.545	.502	.521	170	90	.521	.525	.498	178	96	.547	.498	.524
179	95	.524	.521	.502	180	94	.501	.544	.479	180	97	.546	.499	.523
181	96	.523	.522	.501	182	95	.500	.544	.478	182	98	.545	.500	.522
183	97	.522	.522	.500	184	96	.500	.545	.477	184	99	.544	.501	.521
185	98	.521	.523	.499	186	97	.499	.546	.477	186	100	.542	.501	.520
187	99	.520	.524	.498	195	105	.545	.498	.523	196	104	.523	.520	.502
197	103	.501	.542	.480	197	106	.544	.499	.522	198	105	.522	.521	.501
199	104	.500	.542	.479	199	107	.543	.500	.521	200	106	.521	.521	.500
201	105	.500	.543	.478	.201	108	.542	.501	.521	202	107	.520	.522	.499
203	106	.499	.544	.478	203	109	.541	.501	.520	204	108	.519	.523	.489
212	111	.502	.539	.481	212	114	.543	.498	.522	213	113	.522	.519	.502
214	112	.501	.540	.481	214	115	.542	.499	.521	215	114	.521	.520	.501
216	113	.500	.541	.480	216	116	.541	.500	.521	217	115	.520	.521	.500
218	114	.500	.541	.479	218	117	.540	.501	.520	219	116	.519	.521	.499
220	115	.499	.542	.478	220	118	.539	.501	.519	221	117	.519	.522	.498
229	120	.502	.538	.482	229	123	.541	.498	.522	230	122	.521	.518	.501
231	121	.501	.539	.481	231	124	.540	.499	.521	232	123	.520	.519	.501
233	122	.500	.539	.481	233	125	.539	.500	.520	234	124	.520	.520	.500
235	123	.500	.540	.480	235	126	.539	.501	.519	236	125	.519	.520	.499
237	124	.499	.540	.479	237	127	.538	.501	.518	238	126	.518	.521	.498
239	125	.498	.541	.479	246	129	.502	.537	.483	246	132	.540	.498	.521
247	131	.521	.518	.501	248	130	.501	.537	.482	248	133	.539	.499	.520
249	132	.520	.518	.501	250	131	.500	.538	.481	250	134	.538	.500	.519
251	133	.519	.519	.500	252	132	.500	.538	.481	252	135	.537	.501	.518
253	134	.518	.520	.499	254	133	.499	.539	.480	254	136	.536	.501	.518
255	135	.517	.520	.498	256	134	.498	.540	.479	256	137	.535	.502	.517
263	138	.502	.535	.483	263	141	.539	.498	.520	264	140	.520	.517	.501
265	139	.501	.536	.483	265	142	.538	.499	.519	266	141	.519	.518	.501
267	140	.500	.537	.482	267	143	.537	.500	.519	268	142	.518	.518	.500
269	141	.500	.537	.481	269	144	.536	.501	.518	270	143	.517	.519	.499
271	142	.499	.538	.481	271	145	.535	.501	.517	272	144	.517	.520	.498
273	143	.498	.538	.480	273	146	.534	.502	.516	280	147	.502	.534	.484

TABLE I (Cont'd)

n_1	n_2	P_1	P_2	P	n_1	n_2	P_1	P_2	P	n_1	n_2	P_1	P_2	P
280	150	.537	.498	.520	281	149	.519	.517	.501	282	148	.501	.535	.483
282	151	.537	.499	.519	283	150	.518	.517	.501	284	149	.500	.535	.482
284	152	.536	.500	.518	285	151	.518	.518	.500	286	150	.500	.536	.482
286	153	.535	.500	.517	287	152	.517	.519	.499	288	151	.499	.537	.481
288	154	.534	.501	.516	289	153	.516	.519	.499	290	152	.498	.537	.481
290	155	.533	.502	.516	296	157	.519	.516	.502	297	156	.502	.533	.484
297	159	.536	.498	.519	298	158	.519	.516	.501	299	157	.501	.534	.484
299	160	.536	.499	.518	300	159	.518	.517	.501	301	158	.500	.534	.483
301	161	.535	.500	.517	302	160	.517	.517	.500	303	159	.500	.535	.482
303	162	.534	.500	.517	304	161	.516	.518	.499	305	160	.499	.536	.482
305	163	.533	.501	.516	306	162	.516	.519	.499	307	161	.498	.536	.481
307	164	.532	.502	.515	313	166	.519	.515	.502	314	165	.502	.532	.485
314	168	.535	.498	.518	315	167	.518	.516	.501	316	166	.501	.533	.484
316	169	.535	.499	.518	317	168	.517	.516	.501	318	167	.500	.533	.483
318	170	.534	.500	.517	319	169	.517	.517	.500	320	168	.500	.534	.483
320	171	.533	.500	.516	321	170	.516	.518	.499	322	169	.499	.535	.482
322	172	.532	.501	.516	323	171	.515	.518	.499	324	170	.498	.535	.482
324	173	.532	.502	.515	330	175	.518	.515	.502	331	174	.502	.531	.485
331	177	.534	.499	.518	332	176	.518	.515	.501	333	175	.501	.532	.484
333	178	.534	.499	.517	334	177	.517	.516	.501	335	176	.500	.533	.484
335	179	.533	.500	.517	336	178	.516	.516	.500	337	177	.500	.533	.483
337	180	.532	.500	.516	338	179	.516	.517	.499	339	178	.499	.534	.483
339	181	.531	.501	.515	340	180	.515	.518	.499	341	179	.498	.534	.482
341	182	.531	.502	.515	342	181	.514	.518	.498	347	184	.518	.514	.501
348	183	.502	.531	.485	348	186	.534	.499	.518	349	185	.517	.515	.501
350	184	.501	.531	.485	350	187	.533	.499	.517	351	186	.517	.515	.501
352	185	.500	.532	.484	352	188	.532	.500	.516	353	187	.516	.516	.500
354	186	.500	.532	.484	354	189	.531	.500	.515	355	188	.515	.517	.499
356	187	.499	.533	.483	356	190	.531	.501	.515	357	189	.515	.517	.499
358	188	.498	.533	.482	358	191	.530	.502	.514	359	190	.514	.518	.498
364	193	.518	.514	.502	365	192	.501	.530	.486	365	195	.533	.499	.517
366	194	.517	.515	.501	367	193	.501	.531	.485	367	196	.532	.499	.516
368	195	.516	.515	.501	369	194	.500	.531	.485	369	197	.531	.500	.516
370	196	.516	.516	.500	371	195	.500	.532	.484	371	198	.531	.500	.515
372	197	.515	.516	.499	373	196	.499	.532	.483	373	199	.530	.501	.514
374	198	.514	.517	.499	375	197	.498	.533	.483	375	200	.529	.502	.514
376	199	.514	.517	.498	380	203	.533	.498	.517	381	202	.517	.514	.502
382	201	.501	.529	.486	382	204	.532	.499	.517	383	203	.516	.514	.501
384	202	.501	.530	.486	384	205	.531	.499	.516	385	204	.516	.515	.501
386	203	.500	.530	.485	386	206	.531	.500	.515	387	205	.515	.515	.500
388	204	.500	.531	.484	388	207	.530	.500	.515	389	206	.515	.516	.499
390	205	.499	.531	.484	390	208	.529	.501	.514	391	207	.514	.516	.499
392	206	.498	.532	.483	392	209	.529	.502	.514	393	208	.513	.517	.498

THINK-A-DOT REVISITED

JOHN A. BEIDLER, University of Scranton

1. Introduction. The purpose of this paper is to answer a question raised in [1] about the game THINK-A-DOT. In answering that question, we introduce in section two some elementary concepts from Automata Theory. In subsequent sections we describe the game THINK-A-DOT, discuss the mathematical theory of the game, solve the problem in which we are interested, and suggest some generalizations.

2. Automata theory. Let Σ be a finite nonempty collection of elements. We will call Σ an alphabet. Σ^* is defined to be the collection of all finite strings of elements from Σ including the empty string. We indicate the empty string with the symbol ε. If Σ were the English alphabet, then any string of alphabetic characters would be an element in Σ^*. This would include every word that appears in an English language dictionary as well as any meaningless collection of letters. For example, "*bstfh*", "*qdjkb*", and " ", that is, the empty string ε, would be in Σ^*.

If $x \in \Sigma^*$, the *length* of x, $|x|$, is the number of appearances of symbols from Σ in the word x. If $x = \varepsilon$, $|x| = 0$.

An automaton A is a triple $A = (Q, \Sigma, \delta)$ where

(i) Σ is called the set of *inputs*;

(ii) Q is a finite collection of elements and is called the *set of states*;

(iii) δ is a function, $\delta\colon Q \times \Sigma^* \to Q$, such that for all $x, y \in \Sigma^*$ and for all $p \in Q$,

$$\delta(p, xy) = \delta(\delta(p, x), y).$$

As an example, consider the automaton $A = (Q, \Sigma, \delta)$ where $Q = \{p, q, r\}$, $\Sigma = \{a, b\}$ and δ is defined by

$$\delta(r, a) = p, \quad \delta(p, a) = \delta(r, b) = q. \quad \delta(p, b) = \delta(q, a) = \delta(q, b) = r.$$

A way to represent an automaton pictorially is with a labeled directed graph. This is done by labeling the nodes as states and the arcs with the letters of the input alphabet. Figure 1 is the graph associated with this automaton.

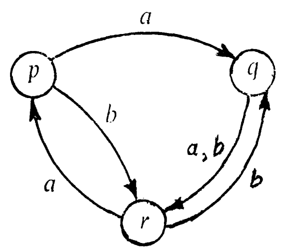


Fig. 1.

An automaton $A = (Q, \Sigma, \delta)$ is said to be *strongly connected* if for all $p, q \in Q$

there is an $x \in \Sigma^*$ such that $\delta(p,x) = q$. The automaton in Figure 1 is strongly connected while that of Figure 2 is not. Observe in Figure 2, there is no $x \in \Sigma^*$ such that $\delta(q,x) = p$.

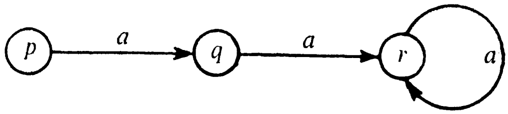


FIG. 2.

An automaton $A = (Q,\Sigma,\delta)$ is said to be *abelian* if for all $p \in Q$ and for all $x, y \in \Sigma^*$, $\delta(p,xy) = \delta(p,yx)$. A state $p \in Q$ is said to be abelian if for all $x, y \in \Sigma^*$, $\delta(p,xy) = \delta(p,yx)$. The automaton of Figure 1 is not abelian,

$$p = \delta(p,ba) \neq \delta(p,ab) = r,$$

while the automaton of Figure 2 is trivially abelian.

Much time has gone into the study of strongly connected and abelian automata (Fleck [**2**], Bayer [**3**], Bavel [**4**], and others). The lemma which follows will indicate how easy it is to determine if an automaton is abelian given that it is strongly connected.

LEMMA 1. *Let $A = (Q,\Sigma,\delta)$ be strongly connected; $p \in Q$ is an abelian state if and only if A is abelian.*

Proof. Let $p \in Q$ be abelian and $q \in Q$ be any state in A. Since A is strongly connected, there is a $z \in \Sigma^*$ such that $\delta(p,z) = q$. Then for any $x, y \in \Sigma^*$,

$$\begin{aligned} \delta(q,xy) &= \delta(p,z(xy)) = \delta(p,(zx)y) \\ &= \delta(p,y(zx)) = \delta(p,(yz)x) \\ &= \delta(p,(zy)x) = \delta(p,z(yx)) = \delta(q,yx). \end{aligned}$$

The converse is obvious.

COROLLARY. *Let $A = (Q,\Sigma,\delta)$ be strongly connected. A is abelian if and only if there is some state $p \in Q$ which is abelian.*

A strongly connected abelian automaton is referred to as a *perfect* automaton. For an automaton $A = (Q,\Sigma,\delta)$ and for $x, y \in \Sigma^*$, we say that x and y are *equivalent*, $x \equiv y$, if for all $q \in Q$, $\delta(q,x) = \delta(q,y)$. An automaton A is said to be *state independent* whenever there is a state $q \in Q$ such that for some $x, y \in \Sigma^*$, $\delta(q,x) = \delta(q,y)$, then $x \equiv y$. Observe that the automaton of Figure 1 is not state independent, $\delta(q,a) = \delta(q,b)$ while $\delta(p,a) \neq \delta(p,b)$.

LEMMA 2. *A perfect automaton is state independent.*

Proof. Let $A = (Q,\Sigma,\delta)$ be perfect and let $p \in Q$ be such that for $x, y \in \Sigma^*$ we have $\delta(p,x) = \delta(p,y)$. Since A is strongly connected, for any $q \in Q$ there is a $z \in \Sigma^*$ such that $\delta(p,z) = q$. Using this fact along with the fact A is abelian, we have

$$\delta(q,x) = \delta(p,zx) = \delta(p,xz) = \delta(\delta(p,x),z)$$
$$= \delta(\delta(p,y),z) = \delta(p,yz) = \delta(p,zy) = \delta(q,y),$$

and hence we have $x \equiv y$.

Actually, much stronger results than those obtained above have been established ([3] p. 7 and [4] p. 14). The way these lemmas are used is quite simple. Consider the automaton of Figure 3, first observe that it is strongly connected. To prove it is abelian, simply observe that for the state p, $\delta(p,ab) = \delta(p,ba)$, then using the corollary of Lemma 1 we have that it is abelian. Since it is strongly connected and abelian, the automaton is perfect and hence by Lemma 2 it is state independent.

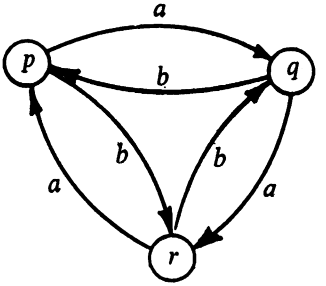


FIG. 3.

3. THINK-A-DOT. The game THINK-A-DOT consists of eight flip-flop gates and connecting channels. There are three holes at the top of the game through which marbles are dropped. As a marble passes through one of the gates it is directed toward the lower left or the lower right. As a marble passes through a gate, the sense of the gate is changed so that the next marble through that gate will be sent in the opposite direction. Figure 4 is a picture of the game and Figure 5 is a schematic

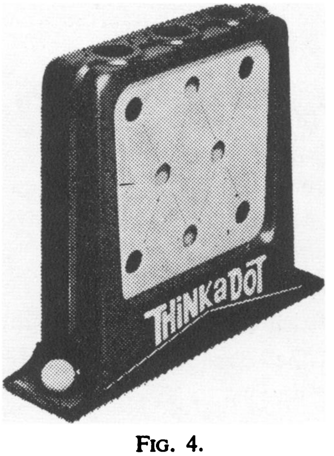


FIG. 4.

description of the connections between the eight flip-flop gates. In Figure 5, the eight circles labeled A through H represent the eight gates and the arcs represent the connecting channels. The holes at the top of the game (see Figure 4) cause a marble dropped into one of them to go directly into gate A, B, or C. We use the letters L, M, and R (for left, middle, and right) to indicate the action of dropping a marble directly into gates A, B, and C respectively. The gates form various patterns when the various combinations of their orientations are considered. By the orientation, we mean the direction in which the next marble through the gate would go.

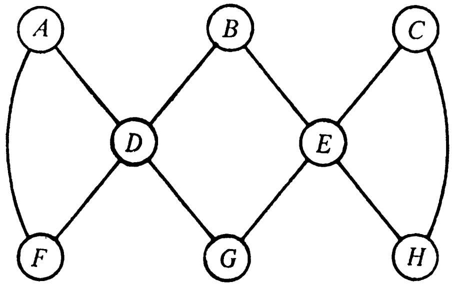


FIG. 5.

The object of the game is to get from one pattern of gates to another (if possible) and do so in the least number of moves. We give a necessary and sufficient condition for a given sequence of marble drops to be a minimum sequence of drops. Furthermore, for a particular sequence of drops which is not minimal, we show how to find an equivalent minimal sequence of drops.

To some extent the solution to this problem is given in [**5**], p. 10. However, the object of this paper is not just to present the result, but to illustrate a use of automata theory.

4. The mathematical theory of THINK-A-DOT. We say that a gate is of *sense*-0 if its orientation is to the left, otherwise we say the gate is of *sense*-1. The set of gates A, B, C, F, G, and H (see Figure 5) is referred to as the *test set*. There are 256 arrangements (patterns) of the eight gates (see [**1**], p. 193). We refer to these patterns by number (0 through 255) according to the following scheme: Assign the values 1, 2, 4, 8, 16, 32, 64, 128 to the gates A through H respectively. Then the number assigned to a pattern is the sum of the products of each gate's value with its sense number. For example, if gates A, B, G, and H have sense-1 and the rest are of sense-0, the pattern number is

$$1 \cdot 1 + 1 \cdot 2 + 0 \cdot (4 + 8 + 16 + 32) + 1 \cdot 64 + 1 \cdot 128 = 195.$$

A gate pattern is called *even* if the number of gates of sense-0 in the test set is even, otherwise the pattern is called *odd*. Note that pattern number 2 is an odd pattern while pattern number 5 is even. This might appear to be confusing. However, the pattern numbers are used to distinguish each pattern from every other pattern while the classification into the even and odd categories is for a completely different purpose. Hence no confusion should arise.

From [**1**] we state without proof

THEOREM 1. *Through appropriate marble drops, an even (odd) pattern can reach any other even (odd) pattern and only even (odd) patterns.*

That is, the patterns are divided into two classes, one of even patterns and the other consisting entirely of odd patterns. You can get from any one pattern in one class to any other pattern in the same class. It is impossible to go from one pattern in one class to a pattern in the other class. The reason for this is when a marble is dropped the sense of two gates in the test set is always changed. This means that if you have an even pattern and drop a marble, the resulting pattern must be even. Similarly, an odd pattern always yields an odd pattern. For the details one should read [1].

5. THINK-A-DOT as an automaton. Observe that THINK-A-DOT is an automaton $A = (Q, \Sigma, \delta)$ where:

(i) Q is the set of gate patterns;

(ii) $\Sigma = \{L, M, R\}$ (the left, middle, and right marble drops);

(iii) δ is the function $\delta: Q \times \Sigma^* \to Q$ where Σ is any "string" of marble drops.

It suffices for our purposes to look at the automaton $A_e = (Q_e, \Sigma, \delta_e)$ where Σ is the same as above and Q_e is the set of even gate patterns and δ_e is δ restricted to Q_e. First, from Theorem 1 it is obvious that A_e is strongly connected. Since A_e is strongly connected, our first major step is to prove

THEOREM 2. *A_e is abelian.*

Proof. First observe for gate pattern 0, which is an even pattern:

(i) $\delta(0, LM) = \delta(33, M) = 11 = \delta(42, L) = \delta(0, ML)$;

(ii) $\delta(0, LR) = \cdots = \delta(0, RL)$;

(iii) $\delta(0, MR) = \cdots = \delta(0, RM)$.

This same fact can be established for every other even gate pattern. From here, one may use an inductive argument to establish that A_e is abelian. The inductive hypothesis is that for all $q \in Q_e$ and for all $x_1, x_2 \in \Sigma^*$ such that $|x_1 x_2| < k$ we have

$$\delta(q, x_1 x_2) = \delta(q, x_2 x_1).$$

One must then show that if y_1, $y_2 \in \Sigma^*$ and $|y_1 y_2| = k$, then for all $q \in Q_e$, $\delta(q, y_1 y_2) = \delta(q, y_2 y_1)$. For the sake of brevity, this part of the proof is omitted.

COROLLARY. *A_e is a perfect automaton.*

COROLLARY. *A_e is state independent.*

All of the statements made about the automaton A_e are true for the automaton $A_0 = (Q_0, \Sigma, \delta_0)$ where Q_0 is the set of odd gate patterns and δ_0 is δ restricted to Q_0.

6. The minimum move problem. Since we know A_e is state independent, it is sufficient for our purposes to observe the relationship of the effect on the automaton by one input string to the effect by another input string by observing the results of using the input string on just one state and generalizing from that result. We can do this because of state independence. That is, if for some particular state $p \in Q$ and for some $x, y \in \Sigma^*$, $\delta(p,x) = \delta(p,y)$ then $x \equiv y$.

LEMMA 3. $L^8 \equiv M^8 \equiv R^8 \equiv \varepsilon$.

Proof. For pattern $0 \in Q_e$ we have

$$\delta(0,L^8) = \delta(0,M^8) = \delta(0,R^8) = \delta(0,\varepsilon) = 0.$$

Since A_e is state independent, the result follows.

The automaton A_e is abelian. This along with Lemma 3 implies that we can get from any one gate pattern in Q_e to any other in Q_e and not have to drop more than 7 marbles in any one hole. This means that at most 21 marble drops would be needed. However, this bound can be reduced.

LEMMA 4 ([1]). *For $p, q \in Q_e$, there exist nonnegative integers a, b, and c such that $\delta(p, L^aM^bR^c) = q$ and $0 \leqq a + b + c \leqq 15$.*

Proof. To get from state p to state q we must make sure that each gate is given the right sense to form pattern q. To make sure that gate A has the proper sense, $a = 0$ or 1. Gates B, D, and F can then be given their proper senses by dropping marbles into the middle hole, M. From Lemma 3, this can be done in less than 8 marble drops, hence $b \leqq 7$. Similarly, gates C, E, G and H can be made to have their proper senses to form gate pattern q by drops R. Again by Lemma 3, $c \leqq 7$. Hence we have $0 \leqq a + b + c \leqq 15$.

The construction of the proof gives us a method of getting from one state (gate pattern) to another. It does not guarantee that this is the smallest number of marble drops needed to effect this change in state. We say that $L^aM^bR^c \in \Sigma^*$ is *minimal* if for all $L^xM^yR^z \in \Sigma^*$ such that $L^aM^bR^c \equiv L^xM^yR^z$ we have

$$a + b + c \leqq x + y + z.$$

To determine necessary and sufficient conditions for a set of marble drops to be minimal we need several lemmas.

LEMMA 5. *If $L^aM^bR^c \equiv L^xM^yR^z$, then $a \equiv x \pmod 2$, $b \equiv y \pmod 2$, and $c \equiv z \pmod 2$.*

Proof. Assume $a \not\equiv x \pmod 2$, then the sense of gate A in the terminal pattern of the marble drop $L^aM^bR^c$ is not the same as the sense of gate A in the terminal pattern of $L^xM^yR^z$. That is, for any $p \in Q$, if $a \not\equiv x \pmod 2$, then $\delta(p, L^aM^bR^c) \neq \delta(p, L^xM^yR^z)$. In a similar way, the other two equations can be proven.

LEMMA 6. $L^2M^2R^2 \equiv \varepsilon$.

Proof.

$$\delta(0, L^2M^2R^2) = \delta(8, M^2R^2) = \delta(16, R^2) = 0.$$

This along with the fact that A_e is state independent yields the desired result.

Using Lemmas 3 and 6 we have

COROLLARY. *$L^2M^2 \equiv R^6$, $L^2R^2 \equiv M^6$, and $M^2R^2 \equiv L^6$.*

Proof (for the first equation).

$$L^2M^2 \equiv L^2M^2R^8 \equiv L^2M^2R^2R^6 \equiv \varepsilon R^6 \equiv R^6.$$

COROLLARY. *$L^4M^4 \equiv R^4$, $L^4R^4 \equiv M^4$, and $M^4R^4 \equiv L^4$.*

We are now prepared to improve upon the limit established in Lemma 4.

THEOREM 3. *For $p, q \in Q_e$, there exist nonnegative integers a, b, and c such that $\delta(p, L^aM^bR^c) = q$ and*

$$0 \leqq a + b + c \leqq 9.$$

Proof. Let $L^aM^bR^c \in \Sigma^*$ be minimal and let $p, q \in Q_e$ be such that $\delta(p, L^aM^bR^c) = q$. The first corollary to Lemma 6 implies that $\max(a, b, c) \leqq 5$. To see this, assume that at least one of a, b, or c is greater than 5. For definiteness assume it is a. Then by this corollary

$$L^aM^bR^c \equiv L^{a-6}L^6M^bR^c \equiv L^{a-6}M^2R^2M^bR^c \equiv L^{a-6}M^{b+2}R^{c+2}$$

and

$$|L^aM^bR^c| > |L^{a-6}M^{b+2}R^{c+2}|$$

which contradicts the assumption that $L^aM^bR^c$ is minimal.

Next, observe that at most one of a, b, and c is greater than 3. For example, if both a and b were greater than 3, then by using the second corollary of Lemma 6 we have

$$L^aM^bR^c \equiv L^{a-4}M^{b-4}L^4M^4R^c \equiv L^{a-4}M^{b-4}R^{c+4}.$$

But

$$|L^aM^bR^c| > |L^{a-4}M^{b-4}R^{c+4}|$$

which contradicts the assumption that $L^aM^bR^c$ is minimal.

With a similar argument using Lemma 6 we have that at least one of a, b, or c is less than 2, that is, $\min(a, b, c) \leqq 1$. Hence we have the desired result.

The number of combinations of a, b, and c satisfying the conditions listed in the proof of Theorem 3 is 128, the same as the number of states in A_e. Hence there is one and only one minimal move associated with every ordered pair of states in Q_e. We formalize this as

THEOREM 4 (The Minimalization Theorem for THINK-A-DOT). *Given $L^aM^bR^c \in \Sigma^*$, $L^aM^bR^c$ is a minimal move if and only if:*

(i) $\text{Max}(a, b, c) \leqq 5$;

(ii) *At most one of* a, b, *and* c *is greater than* 3;
(iii) $\text{Min}(a,b,c) \leqq 1$.

To illustrate the use of the lemmas and corollaries to determine the minimal move associated with an ordered pair of states, consider the attempt to move THINK-A-DOT from gate pattern 0 to gate pattern 39. It is easy to determine that $\delta(0, LM^5R^7) = 39$. Now,

$$\begin{aligned} LM^5R^7 &\equiv LM^5L^2M^2R && \text{(first corollary to Lemma 6)} \\ &\equiv L^3M^7R && \text{(the automaton is abelian)} \\ &\equiv L^3L^2R^2MR && \text{(first corollary to Lemma 6)} \\ &\equiv L^5MR^3 && \text{(the automaton is abelian)} \end{aligned}$$

which satisfies all of the minimality conditions of Theorem 4.

Similarly, to get from pattern 0 to pattern number 145, $\delta(0, L^6M^6R^3) = 145$ and

$$\begin{aligned} L^6M^6R^3 &\equiv M^2R^2M^6R^3 && \text{(first corollary to Lemma 6)} \\ &\equiv M^8R^5 && \text{(the automaton is abelian)} \\ &\equiv R^5 && \text{(Lemma 3)} \end{aligned}$$

which is minimal.

7. Generalized THINK-A-DOT. There are three obvious generalizations of the game THINK-A-DOT. We refer to them as the (m,n)-concave, (m,n)-convex, and the (m,n)-treelike game. The (m,n)-concave game has n rows where n is an odd positive integer and m is the number of gates in the first row, $m \geqq 2$. The even numbered rows have $m-1$ gates while the odd numbered rows have m gates. The channel connections from one row to the next are like those in the actual THINK-A-DOT game. THINK-A-DOT is an example of a (3,3)-concave game (Figure 5).

Figure 6 is an example of a convex game, the (3,3)-convex game. In general, m is

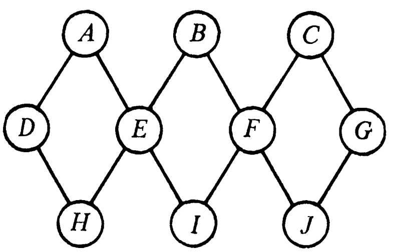


FIG. 6.

the number of gates in the first row and n is an odd positive integer, the number of columns. There are m gates in the odd numbered rows and $m+1$ gates in the even numbered rows. The (m,n)-treelike game has n rows, m gates in the first row and

$m + k - 1$ gates in the kth row. Figure 7 illustrates the (3,3)-treelike game.

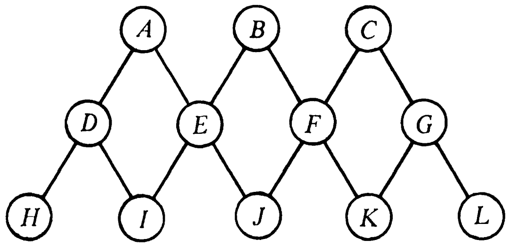


FIG. 7.

As we know, the patterns of the (3,3)-concave game fall into two classes, the odd patterns and the even patterns. The following questions are posed about the other THINK-A-DOT-like games:

1. The number of pattern classes in the (m, n)-concave game is $2^{[(n-1)/2]}$;
2. Prove or disprove that the number of pattern classes in the (3,3)-convex game is 4;
3. Prove or disprove that the number of pattern classes in the (m, n)-convex game is $4^{[(n-1)/2]}$;
4. What is the number of pattern classes in the (3,3)-treelike game? In the (m, n)-treelike game?
5. Effectively solve the minimum move problem for one of the generalized games.

References

1. B. L. Schwartz, Mathematical theory of THINK-A-DOT, this MAGAZINE, 40 (1967) 187–193.

2. A. C. Fleck, Isomorphism Groups of Automata, J. Assoc. Comput. Mach., 9 (1962) 469–476.

3. R. Bayer, Automorphism Groups and Quotients of Strongly Connected Automata and Monadic Algebras, Report No. 204, Department of Computer Science, University of Illinois, 1966.

4. Z. Bavel, On the Structure and Automorphism of Automata, Report 192, Department of Computer Science, University of Illinois, 1965.

5. THINK-A-DOT Instructions, E. S. R. Inc., Montclair, N. J.

FINITE GROUPS ACTING ON SETS WITH APPLICATIONS

LOUIS W. SHAPIRO, Howard University

1. The concept of a group acting on a set is a small generalization of the idea of a permutation group and provides a viewpoint that is useful in attacking a wide variety of problems. The basic concepts and some of the applications are presented here as a series of problems. None are exceptionally difficult although many are motivated by or follow from previous problems. For most an elementary knowledge of group theory will suffice. After the first two sections all sections can be read independently. With moderate supervision this might be useful as an independent study project for any student who has completed a first course in modern algebra.

$m + k - 1$ gates in the kth row. Figure 7 illustrates the (3,3)-treelike game.

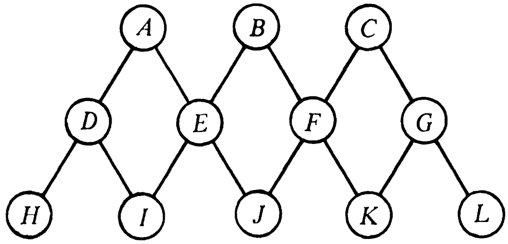


FIG. 7.

As we know, the patterns of the (3,3)-concave game fall into two classes, the odd patterns and the even patterns. The following questions are posed about the other THINK-A-DOT-like games:

1. The number of pattern classes in the (m, n)-concave game is $2^{[(n-1)/2]}$;
2. Prove or disprove that the number of pattern classes in the (3,3)-convex game is 4;
3. Prove or disprove that the number of pattern classes in the (m, n)-convex game is $4^{[(n-1)/2]}$;
4. What is the number of pattern classes in the (3,3)-treelike game? In the (m, n)-treelike game?
5. Effectively solve the minimum move problem for one of the generalized games.

References

1. B. L. Schwartz, Mathematical theory of THINK-A-DOT, this MAGAZINE, 40 (1967) 187–193.

2. A. C. Fleck, Isomorphism Groups of Automata, J. Assoc. Comput. Mach., 9 (1962) 469–476.

3. R. Bayer, Automorphism Groups and Quotients of Strongly Connected Automata and Monadic Algebras, Report No. 204, Department of Computer Science, University of Illinois, 1966.

4. Z. Bavel, On the Structure and Automorphism of Automata, Report 192, Department of Computer Science, University of Illinois, 1965.

5. THINK-A-DOT Instructions, E. S. R. Inc., Montclair, N. J.

FINITE GROUPS ACTING ON SETS WITH APPLICATIONS

LOUIS W. SHAPIRO, Howard University

1. The concept of a group acting on a set is a small generalization of the idea of a permutation group and provides a viewpoint that is useful in attacking a wide variety of problems. The basic concepts and some of the applications are presented here as a series of problems. None are exceptionally difficult although many are motivated by or follow from previous problems. For most an elementary knowledge of group theory will suffice. After the first two sections all sections can be read independently. With moderate supervision this might be useful as an independent study project for any student who has completed a first course in modern algebra.

DEFINITION. *A group, $(G,\circ)$ acts on a set S if each g in G is a function from S to S and*

(a) $(g \circ h)(s) = g(h(s))$ *for all g, h in G, s in S*

(b) $I(s) = s$ *for all s in S, where I is the identity of G.*

Exercise 1–1. If G acts on S each g in G is a permutation of S.

DEFINITION. *If S is a set,* $\mathrm{Sym}(S)$ *is the group of all permutations of S and* $\mathrm{Alt}(S)$ *is the set of all even permutations.* $\mathrm{Sym}(S)$ *is called the symmetric group on S and* $\mathrm{Alt}(S)$ *the alternating group on S.*

Except for informational items enclosed in parentheses, simple affirmative statements in the exercises are to be proved.

Exercise 1–2 *and definition.* If G acts on a set S then there is a homomorphism from G into $\mathrm{Sym}(S)$. This homomorphism will always be denoted θ and will be called the *homomorphism of G acting on S* or the action homomorphism.

Exercise 1–3 *and definition.* If $t \in S$ then $G_t = \{g \mid g \in G,\ g(t) = t\}$ is called the *stability subgroup* of t. Show that G_t is actually a subgroup of G.

DEFINITION. *If $t \in S$ where G acts on S, then the orbit of t under G is the set of all $g(t)$ where g ranges through the elements of G. This orbit is denoted $\mathcal{O}_G(t)$ or $\mathcal{O}(t)$.*

Exercise 1–4. If $S = \{1,2,3,4\}$ and G is $\mathrm{Sym}(S)$ (in such cases we will denote G as $\mathrm{Sym}(4)$) then find G_4 and $\mathcal{O}(4)$.

Exercise 1–5. If $S = \{1,2,3,4,5,6,7\}$ and G is the group of permutations I, $(1234)(56)$, $(13)(24)$, $(1432)(56)$ then find G_1, G_5, G_7, $\mathcal{O}_1, \mathcal{O}_5$, and $\mathcal{O}_7$.

Exercise 1–6. Let $f_1(z) = z, f_2(z) = -1/(1+z), f_3(z) = -(1+z)/z$. Show that $(G,\circ)$ is a group where $G = \{f_1, f_2, f_3\}$ and $\circ$ is compositions of functions. Then show that G acts on the complex plane with 0 and -1 deleted. Find G_1, G_i, and G_ω where $\omega = e^{(2\pi i/3)}$.

We still need an example of a group acting on a set which is not a permutation group. Such examples will be plentiful in Sections 2 and 5 but for now we provide the following:

Exercise 1–7. Let $G = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \;\middle|\; ac \neq 0 \right\}$ be the group of nonsingular upper triangular matrices with real coefficients. If $f = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ then let $f(x) = (ax+b)/c$. Show that this defines an action of G on the real line, $\mathbb{R}$. What is the kernel of the action homomorphism? What are G_0 and $\mathcal{O}(0)$?

Exercise 1–8. If G acts on S then $s \in \mathcal{O}(t)$ if and only if $\mathcal{O}(s) = \mathcal{O}(t)$.

Exercise 1–9. If G acts on S we say $s \sim t$ if $s = g(t)$ for some $g \in G$ where s and t

are in S. Show that $\sim$ is an equivalence relation and that the equivalence classes of $\sim$ are the orbits of G.

Perhaps the next exercise shows the origin of the term orbit.

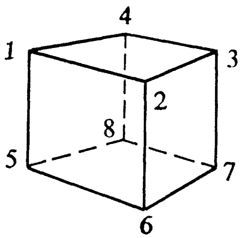
Exercise 1–10. Let G be the group of all 2×2 matrices of the form $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. If $f = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ then let $f(x,y) = (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta)$. Thus f rotates the point (x, y) through an angle θ around the origin. Show that G is a group and that G acts on $\mathbb{R}^2$. What are $G_{(1,0)}$, $G_{(0,0)}$, $\mathcal{O}(1,0)$, $\mathcal{O}(0,0)$? Describe the orbits of G. These obviously partition $\mathbb{R}^2$ which we knew *a priori* by Exercise 1–8.

Exercise 1–11. If G acts on S and $g(s) = t$ then $G_s = g^{-1}G_t g$. In particular G_s and G_t are isomorphic.

We now assume G is finite and we can use the partition to do counting in S.

Exercise 1–12. The Basic Theorem. *If a finite group G acts on a set S then $|G| = |G_t|\ |\mathcal{O}(t)|$ for any t in S. (Here $|A|$ denotes the cardinality of the set A.)*

Exercise 1–13. Take a cube and label all its vertices, say, as follows.



Any rotation of the cube into itself can be represented by the permutation it affects on the vertices. $\alpha = (1234)\ (5678)$, $\beta = (1265)\ (4378)$, $\gamma = (12)\ (46)\ (35)\ (78)$ and $\delta = (254)\ (683)$ are all rotations of the cube. Show $\mathcal{O}(1) = \{1, 2, \cdots, 8\}$ and $R_1 = \{I, \alpha, \alpha^2\}$ and thus that $|R| = 24$ where R is the group of all rotations of the cube into itself.

2. If $G_t = G$ or equivalently $g(t) = t$ for all g in G then t is a *fixed point* of the action of G on S. An oft-occurring situation which guarantees fixed points is set up in the next few exercises.

Exercise 2–1 *and definition.* If P is a finite p-group acting on set S then every orbit has length $1, p, p^2, \cdots, |P|$. The *length* of the orbit $\mathcal{O}(t)$ is just $|\mathcal{O}(t)|$.

Exercise 2–2 (Basic p-group Theorem). *If P is a finite p-group acting on a set S with $p \nmid |S|$ then P has at least one fixed point.*

Exercise 2–3. If G is a group of order 55 acting on a set, S, of order 18, show that G must have a fixed point (in fact at least 2).

Exercise 2–4. Returning to the rotation of a cube discussed in Exercise 1–13 we have a group, R, of order 24. Any element of order 3 generates a cyclic group of order 3. Conversely any subgroup of order 3 is cyclic and thus generated by an element. Show that any element of order 3 in R has 2 fixed points. Geometrically these points must be diametric. How many elements of order 3 are there in R?

We now set up another action. Let G be a group and also let $S = G$. We let G act on G by conjugation. That is, if g is in G then $f_g(x) = gxg^{-1}$ for all x in G. $f_gf_h(x) = f_g(hxh^{-1}) = g(hxh^{-1})g^{-1} = ghx(gh)^{-1} = f_{gh}(x)$, and also $f_I(x) = IxI^{-1} = x$, so indeed this is an action. If g is an element of G, the function f_g is called an *inner automorphism*.

Exercise 2–5. Show that the kernel of the action of G acting on G by conjugation is $\mathbb{Z}(G)$, the center of G.

Exercise 2–6. If H is a subgroup of G let H act on the set G by conjugation. Show that g is a fixed point of H if and only if g is in $C_G(H)$, the centralizer of H in G.

This case of G acting on G by conjugation is of sufficient interest that a special terminology has developed. An orbit is called a *conjugate class* and the stabilizer of a is just the centralizer $C_G(a)$ of a in G. By Exercise 8 the conjugate classes partition G. This action of G on G by conjugation in general does not make G into a permutation group on itself.

Exercise 2–7 (The class equation). Let the conjugate classes of the finite group G be $Cl(a_1)$, $Cl(a_2), \cdots, Cl(a_m)$, $Cl(a_{m+1}), \cdots, Cl(a_k)$ with $|Cl(a_1)| = |Cl(a_2)| = \cdots = |Cl(a_m)| = 1$ and $|Cl(a_i)| > 1$ $\forall i > m$. Then

$$|G| = \sum_{i=1}^{k} |Cl(a_i)| = |\mathbb{Z}(G)| + \sum_{i=m+1}^{k} |Cl(a_i)|$$

$$= |\mathbb{Z}(G)| + \sum_{i=m+1}^{k} \frac{|G|}{|C_G(a_i)|} = \sum_{i=1}^{k} \frac{|G|}{|C_G(a_i)|}.$$

Exercise 2–8. Let the finite p-group P act on the set $P^{\#}$ of all nonidentity elements of P by conjugation. Since $p \nmid |P^{\#}|$ show that $|\mathbb{Z}(P)| > 1$.

Exercise 2–9. Every group of order p^2 (p a prime) is abelian.

Exercise 2–10. If N is a normal subgroup of a finite p-group then show that P acts on N by conjugation. Also show that $p \nmid |N^{\#}|$, that P has fixed points in $N^{\#}$, and that $|N \cap \mathbb{Z}(P)| > 1$. (Taking $N = P$ we obtain Exercise. 2.8 as a special case.)

Exercise 2–11. Find all finite groups G with exactly 1, 2 or 3 conjugate classes.

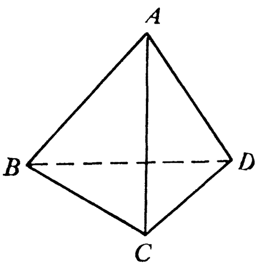
3. In this section we discuss the Pólya-Burnside Theorem and a few applications of it in some counting problems. The Pólya-Burnside Theorem appears in Burnside [**2**] page 189. Pólya realized its applicability and extended its uses in [**10**].

Exercise 3–1 (Pólya-Burnside). *If a finite group G acts on a finite set S and* $\chi(g)$

is the number of elements in S fixed by g then $(1/|G|)\sum_{g\in G}\chi(g)$ = *the number of orbits of G acting on S.* Hint: count the number of pairs (g,s) where $g(s)=s$ two different ways and compare.

Using this theorem we now can consider a counting problem. Let us assume we have wires set up as a regular tetrahedron and at each of the six edges we can attach our choice of a 100–ohm resistor, a 75 watt light bulb, or a capacitor. In our supplies are at least 6 of each of these components. We want to know how many essentially different contraptions we can make if we allow rotations of the tetrahedrons. First we take care of two preliminaries.

Preliminary 1. It is straightforward to count the different (ignoring rotations) contraptions available. We have 3 choices available at each of six locations so we have a set, S, of $3^6 = 729$ possible contraptions.



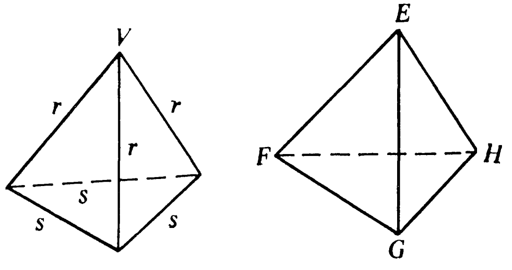
Preliminary 2. What does the group of rotations of a regular tetrahedron look like? If we label the vertices of a tetrahedron we find that the elements are I, (ABC), (ACB), (ABD), (ADB), (ACD), (ADC), (BCD), (BDC), $(AB)(CD)$, $(AC)(BD)$, and $(AD)(BC)$. Thus there is one element of order 1, 8 elements of order 3 each fixing one vertex and 3 elements of order 2 each with no fixed vertices. Let us call this group G.

Returning now to our problem we find that we have a count from Preliminary 1 but it is too high. For instance, there are 6 different contraptions with 5 resistors and 1 light bulb but these are not essentially different if we allow rotations. In fact, these six elements of S form just one orbit under G. Upon further thought each orbit under G gives us just one "essentially different" contraption so the Pólya-Burnside Theorem is exactly what is needed here.

We now need only to compute the $\chi(g)$ for each g in G. $\chi(I) = 3^6$ since the identity fixes every contraption in S. If g is an element of order 3, (a 120° rotation about an axis through one vertex, V, and the center of the opposite triangle) then g fixes only elements of the form below where r and s are any of the 3 choices. Thus

$$\chi(g) = 3^2.$$

If h has order 2 and thus interchanges two pairs of vertices we can assume $h = (EF)(GH)$. Then we have an arbitrary choice of 3 for the edges EF and HG. Side EG is taken to FH (and conversely) so we have a free choice for EG but then FH must be the same choice. Similarly we have a free choice for EH but then no choice for FG. Thus



$\chi(h) = 3^4$ here. Putting this all together we obtain:

$$\text{Answer} = \text{number of orbits} = \frac{1}{|G|} \Sigma_{g \in G} \chi(g)$$

$$= \frac{1}{12}(\chi(I) + 8\chi(g) + 3\chi(h)) = \frac{1}{12}(3^6 + 8 \cdot 3^2 + 3 \cdot 3^4)$$

$$= 87.$$

Exercise 3–2. Redo this example with n choices available at each side instead of 3. This should incidentally give you a somewhat elaborate proof that $(1/12)\cdot(n^6 + 8n^2 + 3n^4)$ is an integer for all positive integers n.

Exercise 3–3. Analyze the dihedral group of order 12 (that is, the group of symmetries of a regular hexagon), as a permutation group on its 6 vertices. Analyze here means first find the number of elements of each order, then subdivide these either as to number of fixed points or geometrically.

Exercise 3–4. If at each carbon atom in a benzene molecule either a $-NH^3$, a $COOH$, or a $-OH$ radical can be attached, how many different compounds are possible?

Exercise 3–5. If each side of a regular hexagon can be painted red, yellow, black, or green, how many essentially different designs are possible allowing all symmetries of a regular hexagon?

Exercise 3–6. If each side and both ends of a regular triangular prism can be painted one of 6 colors, how many essentially different combinations are possible?

Exercise 3–7. If a tape contains 10 digits each either 0 or 1 but the tape can be read indiscriminately from either end, then how many essentially different messages can be recorded on this tape?

For further developments of this material see Pólya [**10**], Liu [**8**], or Harary [**6**]. This theorem can be extended to give a count of the number of graphs with n vertices. See Harary's book for this and for various other important counting problems in combinatorics. His book also contains an extensive list of unsolved counting problems.

There are some other results related to the Polya counting theorem.

DEFINITION. *If a group G acts on a set S and the only orbit of this action is S itself then G is said to be transitive. If for every two pairs of points $\{s_1, s_2\}$ and $\{t_1, t_2\}$ in S there is a g in G such that $g(s_1) = t_1$ and $g(s_2) = t_2$ then G is doubly transitive.*

Exercise 3–8. *If G is transitive on S then $|S| \mid |G|$.*

Exercise 3–9. If *G is transitive on S then G is doubly transitive if and only if* G_t is transitive on $S - \{t\}$.

Exercise 3–10. If G is transitive on S then

$$\sum_{g \in G} (\chi(g))^2 = t|G|$$

where $\chi(g)$ is the number of fixed points of g and t is the number of orbits of G_s. In particular if G is doubly transitive

$$\sum_{g \in G} (\chi(g))^2 = 2|G|.$$

The following exercise is related.

Exercise 3–11*.

$$\sum_{g \in G} (\rho_2(g))^2 = \alpha|G|$$

where $\rho_2(g) = |\{x \mid x \in G,\ x^2 = g\}|$ and α is the number of conjugate classes, C, in G such that if x is in C so is x^{-1}.

Exercise 3–12** (unsolved). Generalize 3–11 for other numbers than 2. (See Research Problems, March 1971, AMERICAN MATHEMATICAL MONTHLY.)

4. In this section we consider the group $GL(2,F) = G$ of all 2×2 nonsingular matrices acting on various sets. Let the matrix $\binom{ab}{cd}$ take the point (x, y) to $(x, y)\binom{ab}{cd} = (ax + cy, bx + dy)$.

We can take S to be points in $F \times F$, lines in $F \times F$ through $(0,0)$, rays emanating from the origin, all subsets of $F \times F$, all finite subsets of $F \times F$, or a variety of other sets.

Exercise 4–1. Let $G = GL(2,F)$ act on the lines through $(0,0)$ in F^2 as above and let M be an element of G. Then a line, l, is fixed by M if and only if every point on l is an eigenvector of M.

Exercise 4–2. Find the subgroup of $GL(2,\mathbb{R})$ that takes all points on the hyperbola $xy = 1$ to other points on the same hyperbola. Prove that this is a maximal subgroup of $SL^*(2,\mathbb{R})$, the matrices of determinant ± 1.

Let $\mathbb{Z}_p$ denote the field of p elements where p is a prime. Let $S = \mathbb{Z}_p \times \mathbb{Z}_p$ and let $S' = \mathbb{Z}_p \times \mathbb{Z}_p - (0,0)$.

Exercise 4–3. If P is a p-subgroup of $GL(2, \mathbb{Z}_p)$ then P has a fixed point in S'. Therefore, there is a nonzero vector v in $\mathbb{Z}_p \times \mathbb{Z}_p$ such that v is an eigenvector with eigenvalue 1 for all the matrices in P. (Show this is also true if P is a p-subgroup of $GL(n, \mathbb{Z}_p)$.)

This last result can be generalized from $\mathbb{Z}_p$ to all fields of characteristic p. See Gorenstein [**5**] page 31.

Exercise 4–4. If $G = SL(n, F)$ denotes the group of all $n \times n$ matrices with determinant 1 and S is the set of all lines through the origin of $F_{(n)}$ then show that G acts on S. Show also that $\mathcal{O}(l) = S$ for any line l in S. Show that the kernel of the action homomorphism is $\mathbb{Z}(G)$. The group $G/\mathbb{Z}(G)$ is thus a permutation group on S and is called $PSL(n, F)$, the projective special linear group over F. (Except when F has order 2 or 3, $PSL(n, F)$ is a simple group.)

Exercise 4–5. Calculate the orders of $GL(2, \mathbb{Z}_p)$, $SL(2, \mathbb{Z}_p)$, $\mathbb{Z}(SL(2, \mathbb{Z}_p))$ and $PSL(2, \mathbb{Z}_p)$. Do the same for n instead of 2.

DEFINITION. *A finite group G is a Frobenius Group if N and H are proper subgroups of G such that $N \Delta G$, $NH = G$, and any $h \in H^*$ induces a fixed point free automorphism of n. That is $h^{-1}nh = n$ with $h \in H$, $n \in N$ implies $h = e$ or $n = e$.*

Frobenius groups are important in the theory of finite groups, division rings, projective geometry, and permutation groups. They have been classified quite thoroughly (see Passman [**9**]). It is known that N is always nilpotent and that H is solvable or involves $SL(2, \mathbb{Z}_5)$, N is called the kernel and H the complement. The following example is an interesting piece of folklore.

Exercise 4–6. Show that $SL(2, \mathbb{Z}_5)$ acts on $\mathbb{Z}_{11} \times \mathbb{Z}_{11}$ in a fixed point free manner, so that the semidirect product of $\mathbb{Z}_{11} \times \mathbb{Z}_{11}$ with $SL(2, \mathbb{Z}_5)$ is the smallest Frobenius group whose complement is not solvable. You may assume $SL(2,5)$ is a subgroup of $SL(2, \mathbb{Z}_{11})$ which follows from some computations with generators and relations (see Huppert [**7**]). This result will follow if we can show every element of $SL(2, \mathbb{Z}_5)$ of order 2, 3, or 5 is fixed point free which follows from the basic p-group theorem applied to the vectors in $\mathbb{Z}_{11} \times \mathbb{Z}_{11}$. Since $|(\mathbb{Z}_{11} \times \mathbb{Z}_{11})^*| = 120 = |SL(2, \mathbb{Z}_5)|$ this must be the smallest such example.

5. Many of the applications of groups acting on sets are in group theory itself. These come about by picking for the set S various sets of subsets of G. For instance back in Section 2 we let $S = G$ or $G - \{1\}$ and then let G act by conjugation.

Exercise 5–1. The Strong Cayley Theorem. *Let H be a subgroup of G and let $S = \{H, Hx, Hy, \cdots\}$ be the set of all right cosets of S. Let G act on S by right mul-*

tiplication so that $Hx \xrightarrow{g} Hxg$. *Verify that this is an action and that the kernel of the action homomorphism is* $\bar{H} = \bigcap_{x \in G} x^{-1}Hx$.

Exercise 5–2. If H is a subgroup of G show that $\bar{H} = \bigcap_{x \in G} x^{-1}Hx$ is the largest normal subgroup contained in H.

If H is of index n in G and θ is the action described above, we note that θ gives a homomorphism from G into Sym(n).

Exercise 5–3. If $H = \{I\}$, the identity subgroup, show that Cayley's Theorem results.

Using the Strong Cayley Theorem, we can obtain direct results saying that existence of a large subgroup guarantees the existence of a reasonably large normal subgroup. Conversely if normal subgroups are sparse so are large subgroups.

Exercise 5–4. If a group G has a subgroup H of finite index greater than 1, then G also has a normal subgroup of finite index greater than 1.

Exercise 5–5. If a group G has order 10,000, then G cannot be simple. [The 1st Sylow Theorem can be used here.]

Exercise 5–6. Using the fact that Alt(n) is simple for $n \geqq 5$ show that Alt(5) has no subgroups of order 15, 20 or 30. Show also that Alt(6) has no subgroups of prime index (it is the smallest group with this property).

Exercise 5–7. For $n \geqq 5$ show that Alt(n) has no subgroups of index $2, 3, \cdots, n-1$.

Exercise 5–8. If G has a subgroup of index 2, 3, or 4 show that G cannot be simple.

Exercise 5–9. For any positive integer, n, there are but a finite number of simple groups having a subgroup of index n.

It would be of great interest if this last result could be sharpened sufficiently to give a useful count. The next two problems are from recent issues of the AMERICAN MATHEMATICAL MONTHLY and are easy if set up with the proper group or subgroup acting on the correct set.

Exercise 5–10. If G is of order $p^n m$ where $m < 2p$ and p is prime, then G has a normal subgroup of order p^n or p^{n-1}.

Exercise 5–11. If G is a torsion group and H a subgroup of finite index m such that each nonidentity element of H has order $\geqq m$, then H is normal. [It is convenient to consider m prime and composite separately.]

Exercise 5–12. Show that the smallest symmetric group which contains a subgroup isomorphic to the quaternions is Sym (8).

Next we develop a short proof of the Sylow theorems using virtually no group theory. The standard proof due to Frobenius [**4**] can be found in many books such as Curtis and Reiner [3]. We start with some elementary results from ring theory. Our approach is due in part to Wielandt [13].

Exercise 5–13. Prove that the binomial theorem holds in any commutative ring.

Exercise 5–14. If p is a prime show

(a) $p \,\Big|\, \binom{p}{k}$ for $k = 1, 2, \cdots, p-1$.

(b) $(a+b)^p = a^p + b^p$ in any commutative ring of characteristic p.

(c) The Frobenius map $a \xrightarrow{\theta} a^p$ is a ring homomorphism in any commutative ring of characteristic p.

(d) $a \xrightarrow{\theta^k} a^{p^k}$ is a ring homomorphism in any commutative ring of characteristic p.

One application of Exercise 5–14 (c) is the following:

Exercise 5–15.

(I) $a^p = a$ for all $a \in \mathbb{Z}_p$, the integers modulo p.

(II) Equivalently $a^p \equiv a \pmod{p}$ for $a \in \mathbb{Z}$.

Exercise 5–16. Show that $\binom{p^q m}{p^q} = m\alpha$ where $\alpha \equiv 1 \pmod{p}$. If $p \nmid m$ this follows from Exercise 5–14(d). Otherwise it seems necessary to either expand the binomial coefficient or to check out what happens with a known group in the middle of Exercise 5–17.

Exercise 5–17 (The First Sylow Theorem). *Let G be a finite group of order $n = p^a m$ where $p \nmid m$ and let S be the set of all subsets of order p^b where $p^b \mid n$. Let G act on S by right multiplication and show that this actually is an action.*

Use the previous exercise to show that not every orbit has length divisible by p^{a-b+1}. Let $\mathcal{O}$ be one such nondivisible orbit and let T be one of the sets in $\mathcal{O}$. Show that G_T has order divisible by p^b. To finish let G_T act on the set T by right multiplication and use the left cancellation law to show $|G_T| \leqq p^b$ so that $St(T)$ is the desired subgroup.

Exercise 5–17.

(a) We now know that Alt(6) must have subgroups of order 1, 2, 4, 8, 3, 9, and 5. Write down explicitly one subgroup of each of these orders.

(b) Show that in any finite p-group the converse of Lagrange's Theorem is true.

Exercise 5–18. This proof of Sylow's Theorem is constructive in the sense that if the multiplication is known in a group the p-subgroups can be constructed. Try this for some group of very small order.

Exercise 5–19 (continued from Exercise 5–17). Show that p^{a-b+1} does not divide the orbit length of $\mathcal{O}$ if and only if $\mathcal{O}$ consists only of left cosets of subgroups of order p^b.

Exercise 5–20 (The 2nd Sylow Theorem). A group G of order $p^a m$ where $p \nmid m$ has $1 + kp$ subgroups of order p^a (which are called p-sylow subgroups). In fact the number of subgroups of order p^b is of the form $1 + kp$ for all $b \leqq a$.

Exercise 5–21 (The 3rd Sylow Theorem). The exercise is to fill in the details of the following argument. Let G be a finite group and let S be the set of all p-sylow subgroups of G. Assume that there are two distinct orbits $\mathcal{O}_1$ and $\mathcal{O}_2$ when G acts on S by conjugation. Let P_1 and P_2 be p-sylow subgroups, P_1 in $\mathcal{O}_1$, and P_2 in $\mathcal{O}_2$. If P_1 acts on $\mathcal{O}_2$ we see that $|\mathcal{O}|_2 = 0 \pmod p$ but P_2 acting on $\mathcal{O}_2$ yields $|\mathcal{O}|_2 = 1 \pmod p$. Thus G acting on S must have but one orbit and thus all the p-sylow subgroups of G are conjugate.

Exercise. 5–22 For each $p^n \,\Big|\, |P|$ where P is a finite p-group there are $1 + kp$ subgroups of order p^m.

Exercise 5–23. Show that the following 3 sets of matrices all with entries in Z_p are conjugate.

$$T_1 = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1 & e & 0 \\ 0 & 1 & 0 \\ f & g & 1 \end{bmatrix}, \quad T_3 = \begin{bmatrix} 1 & 0 & 0 \\ i & 1 & 0 \\ j & k & 1 \end{bmatrix}, \quad a,b,c,e,f,g,i,j,k \in \mathbb{Z}_p.$$

Exercise 5–24. If P is a p-sylow subgroup of a finite group G then the number of p-sylow subgroups in G is

$$\frac{|G|}{|N_G(P)|}.$$

Exercise 5–25. There are no simple groups of order 200.

Exercise 5–26. How many p-sylow subgroups are in $SL(2, \mathbb{Z}_p)$?

6. Exercises 6–1 and 6–2 are worth noting for anyone who has wanted to explain how modern algebra got its name (the algebra part) without explaining a substantial part of Galois Theory.

Exercise 6–1. If a and b are two elements of order 2 in a group G, then $\langle a, b \rangle$, the subgroup generated by a and b, is a dihedral group of order $2n$ where n is the order of the element ab.

We now want to examine the roots of polynomials over the rationals. If $f(x)$ is a polynomial of degree n it has at most n distinct roots in its splitting field, K. We are interested in Sym(n), the symmetric group on these n roots. In particular we are interested in two subsets of Sym(n), those elements that are in some way computable and the Galois group of K over $\mathbb{Q}$. Let us look at an example where there are a few computable elements of Sym(n) available.

Exercise 6–2. Is the polynomial $f(x) = x^8 + (x+1)^8 + 1$ irreducible? Hint: Note that $x \xrightarrow{a} 1/x$ and $x \xrightarrow{b} -1-x$ are elements permuting the 8 roots. (It can be shown that $f(x)$ has distinct roots by showing g.c.d. $(f(x), f'(x)) = 1$.)

[Use Exercise 6–1, some computations to find an element of order 3, Exercise 2–2, and the fact that complex roots of an equation with real coefficients come in complex pairs.]

Another interesting connection between computable elements and Galois groups is that if $x \xrightarrow{\alpha} \frac{ax+b}{cx+d}$ $(a,b,c,d \in \mathbb{Z})$ is a permutation of the roots of a polynomial, then α commutes with each element of the Galois group.

Exercise 6–3. Prove this assertion.

Exercise 6–4. Characterize those polynomials such that $x \xrightarrow{\alpha} 1/x$ permutes their roots.

Exercise 6–5. Show that $f(x) = \sum_{i=0} a_i x^i + \sum_{i=1}^{n} a_{n-i} x^{n+i}$ with $a_i \in \mathbb{Q}$ has a rational root. Hint: Use 2–2 and 6–4.

Exercise 6–6. Compute the centralizers of the following elements in Sym(n).
(a) $a = (123 \cdots n)$
(b) $b = (123)$
(c) $c = (12)(34) \cdots (n-1, n)$ where n is even.
Show that $C_{Sym(n)}(c)$ is solvable iff $n = 2, 4, 6,$ or 8.

Exercise 6–7. Show that the polynomial

$$a_0(x^{2n} + 1) + a_1(x^{2n-1} + x) + a_2(x^{2n-2} + x^2) + \cdots + a_n x^n$$

is solvable by radicals if $n \leqq 4$ where $a_i \in \mathbb{R}$.

References

1. E. Artin, Algebra I, Lecture notes in German from Universität Hamburg, 1961, pp. 9–14.
2. W. Burnside, Theory of Groups of Finite Order, 2nd ed., Dover, New York, 1955, pp. 189–191.
3. C. Curtis and I. Reiner, Representation Theory of Finite Groups and Associative Algebras, Interscience, New York, 1962, p. 17.
4. G. Frobenius, Gesammelte Abhandlungen II, Springer-Verlag, New York, 1968, p. 301.
5. D. Gorenstein, Finite Groups, Harper & Row, New York, 1968, p. 31.
6. F. Harary, Graph Theory, Addison-Wesley, Reading, 1969, pp. 178–197.
7. B. Huppert, Endliche Gruppen, Springer-Verlag, New York, 1969.
8. C. C. Liu, Introduction to Combinatorial Mathematics, McGraw-Hill, New York, 1968, pp. 126–166.
9. D. Passman, Permutation Groups, Benjamin, New York, 1968.
10. G. Polya, Kombinatorishe Anzahlbestimmungen für Gruppen, Graph und chemische Verbindungen, Acta Math., 68 (1937) 145–254.
11. C. H. Sah, Abstract Algebra, Academic Press, New York, 1967, pp. 54, 60, 79.
12. L. Sylow, Théorèmes sur les groupes de substitutions, Band V, (1872) 584–594.
13. H. Wielandt, Ein Beweis für den Existenz von Sylowgruppen, Arch. Math., 10 (1959) 401–402.

CHARACTERIZING MOTIONS BY UNIT DISTANCE INVARIANCE

RICHARD L. BISHOP, University of Illinois, Urbana

It has been proved by P. Zvengrowski [**1**, appendix to chapter II] that a map of the euclidean plane, $T: E^2 \to E^2$, which preserves unit distance is an isometry. His proof uses an approximation procedure based on the theorem of Kronecker that the ring generated over Z by $\sqrt{3}$ is dense in R. A weaker version of this theorem, where the map is assumed to be one-to-one, is proved in [**2**]. Some of the unproved assertions in [**2**] are not so obvious as one might think, since the proofs *we* have, anyway, do not generalize to higher dimensions. However, the ideas in [**2**] could be used to considerably simplify Zvengrowski's proof, eliminating the reliance on algebraic number theory. Here we present elementary proofs for all dimensions of this theorem.

THEOREM. *If* $T: E^p \to E^p$ *preserves unit distances, then* T *is an isometry* $(p > 1)$.

We say that T *preserves distance* r if $d(x, y) = r$ implies $d(Tx, Ty) = r$. Assuming that T preserves distance 1, our goal is to prove that T preserves distance r for all r.

LEMMA 1. [**1**] *If* T *preserves distance* r *and* n *is a positive integer* $\geqq 2$, *then* $d(Tx, Ty) \leqq nr$ *whenever* $d(x, y) \leqq nr$.

Proof. Let k be the integer such that $(k-1)r < d(x, y) \leqq kr$, so $k \leqq n$. On the segment xy lay off $k - 1$ points separated by distances r, starting with $x = x_1, x_2, \cdots, x_{k-1}$. Since $d(x_{k-1}, y) \leqq 2r$, there is a point x_k (not necessarily on the segment xy) such that $d(x_{k-1}, x_k) = d(x_k, y) = r$. If we map this configuration by T, we get a chain of k segments of length r connecting Tx and Ty. Hence $d(Tx, Ty) \leqq kr \leqq nr$.

LEMMA 2. *Suppose there are arbitrarily large and arbitrarily small distances preserved by* T. *Then* T *is an isometry.*

Proof. Let x, y be any two points, $a = d(x, y)$. Let $b > 0$ be such that T preserves distance $a + b$, and let z be the point at distance $a + b$ from x such that y is on the segment between x and z. Let $u = d(Tx, Ty)$, $v = d(Ty, Tz)$, and let y' be the point on the segment between Tx and Tz such that $d(Tx, y') = a$. We want to show that $u = a$.

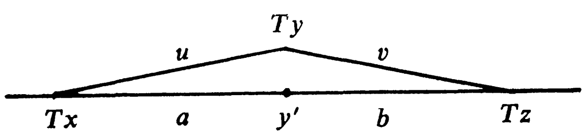


FIG. 1.

Suppose r is a distance preserved by T. Let m, n be the integers such that $(m-1)r < a \leqq mr$, $(n-1)r < b \leqq nr$. By Lemma 1, $u \leqq mr$ and $v \leqq nr$. Combining these inequalities we get $u - a < r$, $v - b < r$. But by the triangle inequality $a + b \leqq u + v$, which gives $a - u \leqq v - b < r$ and $b - v \leqq u - a < r$. Thus $|a - u|$ and $|b - v|$ are both bounded by r. By hypothesis we may take r arbitrarily small. Hence $a = u$ and $b = v$.

Now suppose that r is preserved. Consider the figure formed by two simplices having a common face and all edges of length r. We can represent such a figure analytically by letting one vertex of the common face be an origin O; the other vertices of the common face will be vectors $e_1, \cdots, e_{p-1}$ having length r; and the extreme points by vectors f_0, f_2 which will be reflections of each other in the hyperplane of O, $e_1, \cdots, e_{p-1}$. Since the length of $e_i - e_j, f_k - e_i$ must also be r for $i \neq j$, we can calculate the inner products of all the vectors involved: $\langle e_i, e_j \rangle = \langle f_k, e_i \rangle = r^2/2$. Since $(f_0 + f_2)/2$ must lie on the hyperplane of the common face and is symmetrically placed with respect to all the vertices, it must be their centroid: $(f_0 + f_2)/2 = (e_1 + \cdots + e_{p-1})/p$. Thus $f_0 - f_2 = f_0 + f_2 - 2f_2 = 2(e_1 + \cdots + e_{p-1} - pf_2)/p$. This enables us to calculate the inner product of $f_0 - f_2$ with itself:

$$\langle f_0 - f_2, f_0 - f_2 \rangle = 4r^2[p - 1 + p^2 + (p-1)(p-2)/2 - p(p-1)]/p^2$$

$$= 2r^2(p+1)/p.$$

Thus the distance between the extreme vertices of the figure is $r[2(p+1)/p]^{\frac{1}{2}} = 2r_p$.

Consider what happens to such a figure under T. The points separated by r must be mapped into points separated by r, so each simplex goes into a congruent simplex. But two congruent equilateral simplices having a common $p - 1$ face must either coincide or form a figure congruent to the one under consideration. That is, we must have either $Tf_0 = Tf_2$ or $d(Tf_0, Tf_2) = 2r_p$.

If we are given two points x, y at distance $2r_p$, then there is a figure congruent to the one above and having the extreme vertices as x, y. Thus every pair of points separated by $2r_p$ must either be mapped into the same point or mapped to a pair with separation $2r_p$. However, we can find a third point z such that $d(z,x) = r$, $d(z,y) = 2r_p$. If we had $Tx = Ty$, then $d(Tz, Tx) = r$ since T preserves r, but also $d(Tz, Tx) = d(Tz, Ty) = 0$ or $2r_p$ since $d(z,y) = 2r_p$. This is a contradiction, so we must have $d(Tx, Ty) = 2r_p$. We have proved

LEMMA 3. *If T preserves distance r, then T preserves distance* $2r_p = r[2(p+1)/p]^{\frac{1}{2}}$. *Hence there are arbitrarily large distances which T preserves.*

We have seen that T preserves a figure congruent to $\{0, e_1, \cdots, e_{p-1}, f_0, f_2\}$. Now we can generate more invariant figures by chaining together equilateral simplices face-to-face. Let us label $e_{p-1} = f_1$ and define f_k recursively by: f_{k+1} is the reflection of f_{k-1} in the hyperplane of 0, $e_1, \cdots, e_{p-2}, f_k$. The figures congruent to $\{0, e_1, \cdots, e_{p-2}, f_0, \cdots, f_k\}$ are mapped rigidly by T. Since $f_h - f_k$ is perpendicular to each e_i, the projection into the two-dimensional plane through O perpendicular to

$e_1, \cdots, e_{p-2}$ carries f_k into a sequence g_k which is equally spaced, with distance r from one to the next, on a central circle. All of the distances $d(f_0, f_k) = d(g_0, g_k)$ will be preserved by T, because we can build such a figure around two points at such a distance. When $p = 2$ we have $g_k = f_k$ and this sequence forms a regular hexagon, so gives no new information. For $p > 2$ a particular case, combined with Lemmas 2 and 3, completes the proof of the theorem:

LEMMA 4. *$d(f_0, f_5) = r\,|p^2 - 2p - 4|/p^2$ is preserved by T. When $p > 2$ this is less than r, so by iteration, T preserves arbitrarily small distances.*

Proof. We calculate the equal sides of the isosceles triangle O, g_0, g_1 and use a multiple angle formula to find the base of the isosceles triangle O, g_0, g_k. It is easy to check that $f_0 - (e_1 + \cdots + e_{p-2})/(p-1)$ is perpendicular to each e_i, so it must be g_0. Then

$$\begin{aligned}\langle g_0, g_0\rangle &= r^2[(p-1)^2 + p-2 - (p-1)(p-2) + (p-2)(p-3)/2]/(p-1)^2 \\ &= r_{p-1}^2 .\end{aligned}$$

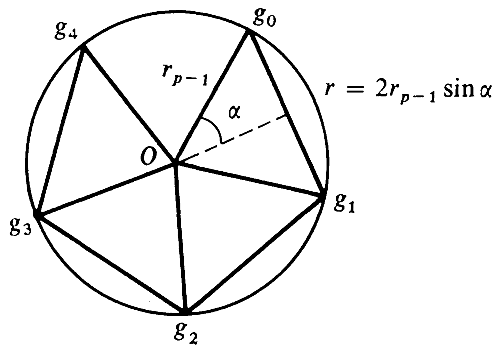


FIG. 2.

Hence the g_k all lie on a circle of radius r_{p-1}. Now the isosceles triangle having base b and equal sides a has vertex angle $2\alpha = \sin^{-1}(b/2a)$. Thus the angle α for Og_0g_1 has $\sin\alpha = r/2r_{p-1} = ([p-1]/2p)^{\frac{1}{2}}$ and hence $\cos\alpha = ([p+1]/2p)^{\frac{1}{2}}$. For the corresponding angle 5α for Og_0g_5 we have

$$\begin{aligned}e^{i5\alpha} &= [([p+1]^{\frac{1}{2}} + i[p-1]^{\frac{1}{2}})/[2p]^{\frac{1}{2}}]^5 \\ &= \frac{r}{2p^2 r_{p-1}}(2 - p^2 - 2[p+1]^{\frac{1}{2}} + i[4 + 2p - p^2]).\end{aligned}$$

The base b of this triangle is $b = 2r_{p-1}|\sin 5\alpha| = r|\,p^2 - 2p - 4|/p^2$.

Now we turn to the case $p = 2$. In this case the figures $\{0, f_0, f_1, \cdots, f_5\}$ are the vertices and center of a regular hexagon. We may permute O, f_0, f_1 and generate another hexagon whose vertex set is mapped rigidly by T. If we do this sort of thing maximally, and now assume $r = 1$, we get an *equilateral unit lattice*, which we abbreviate EUL. Every such EUL is mapped rigidly by T into another EUL, that is:

LEMMA 5. *If F is an* EUL, *then there is a euclidean motion M such that the restrictions of T and M to F are the same.*

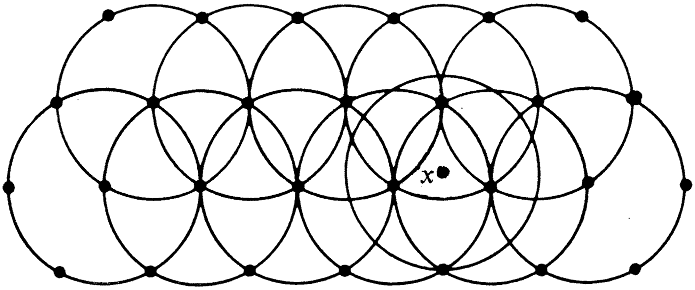


FIG. 3. An EUL with its unit circles.

Now we reduce to the case where T fixes all the points of an EUL, and our goal is to prove that T is the identity I. For if F_0 is an EUL and M_0 coincides with T on F_0, $M_0^{-1}T$ will fix all points of F_0 and will preserve distance 1. If $M_0^{-1}T$ can be proved to be I, then $T = M_0$. Thus we assume that T fixes an EUL F_0.

LEMMA 6. *The displacement $d(x, Tx)$ of every point is at most* 1.

Proof. Every x is within distance 1 of some point of F_0, so the result follows immediately from Lemma 1 with $n = 1$.

LEMMA 7. *The action of T on any* EUL *is a translation or the identity.*

Proof. The action of T on an EUL is either a line reflection, a rotation about a point, a glide reflection, a translation, or the identity. The only ones of these which do not have arbitrarily large displacements on some points of any EUL are the translations and the identity.

LEMMA 8. *The unit circle having a fixed point of T as its center is pointwise fixed under T.*

Proof. Any point on such a unit circle is on an EUL containing the center. Since the center is not translated, neither is the point.

LEMMA 9. *Every point lies on a unit circle whose center is fixed by T.*

Proof. The unit circles centered on F_0 consist of fixed points. But for every point the unit circle about it meets one of these fixed points.

This clearly finishes the proof of the theorem.

The author thanks John Wetzel for bringing his attention to Zvengrowski's theorem and particularly for pointing out the following references.

References

1. P. S. Modenov and A. S. Parkhomenko, Geometric Transformations, vol. 1, Academic Press, New York, 1965.

2. C. G. Townsend, Congruence-preserving mappings, this MAGAZINE, 43 (1970) 37–39.

AN INEQUALITY FOR ELLIPTIC AND HYPERBOLIC SEGMENTS

MICHAEL GOLOMB, Purdue University and HIROSHI HARUKI, University of Waterloo

Let P_0 be a point on the arc Γ of a plane curve and let P_-, P_+ be endpoints of a chord of Γ that is parallel to the tangent at P_0. If T is the area of the triangle with vertices P_-, P_0, P_+, while S is the area of the segment bounded by the arc and the chord with the common endpoints P_-, P_+ then

$$T = \tfrac{3}{4}S \tag{1}$$

provided the curve is a parabola. This equation is well known and was first proved by Archimedes (see [**1**]) by his method of exhaustion. Today, the proof of (1) is a simple exercise in elementary calculus. Less well known, if at all, is the fact to be proved below that 3/4 is the limit of the ratio of T to S as P_-, P_+ approach P_0 on an arbitrary curve Γ of continuous nonzero curvature. The main result to be proved in this note is that equation (1) is replaced by the inequality $T < \frac{3}{4}S$ if Γ is an ellipse and by $T > \frac{3}{4}S$ if Γ is one branch of a hyperbola.

Assume T is a planar simple arc, the continuous one-to-one image $P(t)$ of some interval. Suppose Γ has continuous nonzero curvature in the neighborhood of the interior point $P = P(t_0)$. Then there is a positive δ such that for every t_+ with $0 < t_+ - t_0 < \delta$ there is a unique t_- with $0 > t_- - t_0 > -\delta$ for which the chord with endpoints $P(t_-), P(t_+)$ is parallel to the tangent at P_0. Let $T(t_+)$ denote the area of the triangle with vertices $P(t_-), P_0, P(t_+)$, while $S(t_+)$ denotes the area of the segment bounded by the arc and the chord with the common endpoints $P(t_-), P(t_+)$. Clearly $T(t_+)$ is the maximum area of any triangle with the vertices $P(t_-), P(t), P(t_+)$, $t_- < t < t_+$. We prove

THEOREM 1. *Under the above hypotheses the following limit relation holds*:

$$\lim_{t \to t_0} \frac{T(t)}{S(t)} = \frac{3}{4}. \tag{2}$$

Proof. We may assume P_0 is the origin of the xy-plane and the arc is tangent to the x-axis at P_0. From the hypotheses it follows that a sufficiently small part of the arc with P_0 as interior point is given by

$$x = t - t_0, \quad y = ax^2 + o(x^2)$$

where $a > 0$ and $o(x^2)$, as usual, denotes a remainder term for which $x^{-2}o(x^2) \Rightarrow 0$ as $x \Rightarrow 0$. P_0 is the point with coordinates $(0,0)$ and if $P(t)$ has coordinates (x, y) then $P(t_-)$ has coordinates (x_-, y), where $y = ax_-^2 + o(x_-)$, hence

$$x_- = -x + o(x).$$

Therefore,

$$\begin{aligned} T(t) &= \tfrac{1}{2}(x - x_-)y = xy + o(x)y \\ &= ax^3 + o(x^3) \end{aligned} \tag{3}$$

and

$$S(t) = 2T(t) - \int_{x_-}^{x} y\,dx \tag{4}$$

$$= \frac{4}{3}ax^3 + o(x^3).$$

The assertion (2) follows immediately from (3) and (4).

We assumed that the arc has a continuous curvature near P_0 and that this curvature is not zero. That both of these hypotheses are necessary in Theorem 1 is seen from the following example. Let the arc be given by

$$x = t - t_0, \quad y = |x|^p. \tag{5}$$

For $p > 1$ this arc has the x-axis as tangent at P_0, and for each $t > t_0$ the points $P(t) = (x, y)$, $P(t_-) = (-x, y)$ bound a chord parallel to this tangent. But the arc has infinite curvature at P_0 if $p < 2$ and zero curvature if $p > 2$. As above, one finds

$$\lim_{t \to t_0} \frac{T(t)}{S(t)} = \frac{1}{2} + \frac{1}{2p} \tag{6}$$

which equals $\frac{3}{4}$ only if $p = 2$.

For a parabolic arc (that is, a piece of a quadratic parabola) the ratio $T(t)/S(t)$ is constant. Indeed, we prove

THEOREM 2. *If* Γ *is a parabolic arc then*

$$\frac{T(t)}{S(t)} = \frac{3}{4} \text{ for all } t. \tag{7}$$

Proof. We assume the arc is positioned so that P_0 is the origin and the axis of the parabola is parallel to the y-axis. Then the equation of the arc is

$$x = t - t_0, \quad y = ax^2 + bx$$

for some constants $a > 0$, b. If (x, y) are the coordinates of $P(t)$ then (x_-, y_-), where $x_- = -x$, $y_- = ax^2 - bx$, are those of $P(t_-)$. Hence

$$T(t) = \tfrac{1}{2}(xy_- - x_-y)$$

$$= ax^3,$$

$$S(t) = 2T(t) - \int_{-x}^{x} (ax^2 + bx)dx$$

$$= \frac{4}{3}ax^3,$$

and $T(t)/S(t) = \frac{3}{4}$ is proved.

THEOREM 3. *If* Γ *is an elliptic (hyperbolic) arc then*

$$\frac{T(t)}{S(t)} < \frac{3}{4} \left(\frac{T(t)}{S(t)} > \frac{3}{4}\right) \text{ for all } t \tag{8}$$

while

$$\sup_t \frac{T(t)}{S(t)} = \frac{3}{4} \left(\inf_t \frac{T(t)}{S(t)} = \frac{3}{4}\right). \tag{9}$$

Proof. Let the arc be positioned so that the x-axis is tangent at P_0 Then it is given by $x = f(t)$, $y = g(t)$ where $f(t_0) = g(t_0) = 0$ and

$$ax^2 + 2bxy + cy^2 - 2dy = 0. \tag{10}$$

Since there must be two distinct points of Γ that have the same ordinate $y \neq 0$, we conclude that $a \neq 0$ in (10) and we may assume $a = 1$. Suppose now $P(t_+) = (x_+, y_+)$ and $P(t_-) = (x_-, y_+)$, $y_+ > 0$, are the endpoints of a chord parallel to the x-axis. Since $x = x_+$ and $x = x_-$ satisfy equation (10) for $y = y_+$, we conclude that $-2by_+ = x_- + x_+$, hence

$$b = -\frac{x_- + x_+}{2y_+}. \tag{11}$$

If we set

$$b^2 - c = \delta, \tag{12}$$

where $\delta > 0$ if Γ is elliptic, $\delta < 0$ if Γ is hyperbolic, then c is determined by (11), (12) and d from (10), (11), and (12):

$$d = \frac{(x_- - x_+)^2}{8y_+} - \frac{\delta}{2} y_+. \tag{13}$$

We assert there is a unique parabola which passes through the points (x_+, y_+), (x_-, y_+) and is tangent to the x-axis at $(0,0)$. Indeed, if there is such a parabola, its points (x, y) must satisfy an equation of the form

$$x^2 + 2b_0xy + c_0y^2 - 2d_0y = 0. \tag{14}$$

As above, we find

$$b_0 = -\frac{x_- + x_+}{2y_+} = b \tag{15}$$

and in place of (12), (13) we have

$$c_0 = b_0^2 = c + \delta \tag{16}$$

$$d_0 = \frac{(x_- - x_+)^2}{8y_+} = d + \frac{\delta}{2} y_+. \tag{17}$$

Thus, the coefficients of (14) are uniquely determined.

We show next that the elliptic (hyperbolic) segment bounded by the chord with endpoints (x_-, y_+), (x_+, y_+) contains properly (is contained properly in) the parabolic segment determined by the same points (x_-, y_+), (x_+, y_+). Near $(0,0)$, equations (10) and (14) may be solved for y and give $y = h(x)$, $y = h_0(x)$, respectively. The functions h, h_0 have derivatives of all orders at 0, and $h(0) = h'(0) = h_0(0) = h_0'(0) = 0$. For the second order derivatives we find from (10) and (14):

$$h''(0) = \frac{1}{d}, \quad h_0''(0) = \frac{1}{d_0}. \tag{18}$$

Comparing (17) with (18) shows that $h''(0) < h_0''(0)$ if Γ is elliptic, and $h''(0) > h_0''(0)$ if Γ is hyperbolic. It follows that near $(0,0)$ the elliptic (hyperbolic) arc is below (above) the parabolic arc. But since two different conics that are tangent to each other cannot intersect in more than two other points we conclude that all of the elliptic (hyperbolic) arc from (x_-, y_+) to (x_+, y_+) lies below (above) the parabolic arc. The inequality (8) is an immediate consequence of this. But (8) together with (2) imply (9). Thus, the theorem is proved.

Reference

1. C. B. Boyer, A History of Mathematics, Wiley, New York, London, Sydney, 1968, pp. 142–144.

THE ALTITUDES OF A SIMPLEX ARE ASSOCIATED

LEON GERBER, St. John's University

It is well known that the altitudes of a triangle concur while the altitudes of an n-simplex ($n > 2$) do not concur unless each edge is perpendicular to the edges it does not meet. We give a simple proof of the following result of S. R. Mandan [**1**].

THEOREM. *Any $(n-2)$-dimensional flat which meets n of the altitudes of an n-simplex also meets the remaining altitude. Any set of $n+1$ mutually skew lines in n-space having this property is called an associated set of lines.*

A few remarks about associated lines will illustrate the strength of the theorem. In 3-space, given any three mutually skew lines there exist infinitely many lines associated with the first three; i.e., any line ((3–2)-flat) which meets the first three meets all the others. The lines associated with the given lines form one set of rulings of a hyperboloid of one sheet if there is no plane parallel to all three, and they form the rulings of a hyperbolic paraboloid if there is such a plane [**2**, pp. 14–15]. In 4-space, given any four mutually skew lines there is a unique fifth line associated with them [**3**, pp. 115–116]. An enumerative argument shows that for $n > 4$ there is in general no line associated with n given lines.

We develop some machinery for the proof. Let $\mathscr{I}$ represent the set $\{0, 1, \cdots, n\}$, $\mathscr{I}'$ the set $\{1, \cdots, n\}$, and let the ranges of Σ and Σ' be $\mathscr{I}$ and $\mathscr{I}'$ respectively. Let $\mathscr{A}$

We show next that the elliptic (hyperbolic) segment bounded by the chord with endpoints (x_-, y_+), (x_+, y_+) contains properly (is contained properly in) the parabolic segment determined by the same points (x_-, y_+), (x_+, y_+). Near (0,0), equations (10) and (14) may be solved for y and give $y = h(x)$, $y = h_0(x)$, respectively. The functions h, h_0 have derivatives of all orders at 0, and $h(0) = h'(0) = h_0(0) = h_0'(0) = 0$. For the second order derivatives we find from (10) and (14):

$$h''(0) = \frac{1}{d}, \quad h_0''(0) = \frac{1}{d_0}. \tag{18}$$

Comparing (17) with (18) shows that $h''(0) < h_0''(0)$ if Γ is elliptic, and $h''(0) > h_0''(0)$ if Γ is hyperbolic. It follows that near (0,0) the elliptic (hyperbolic) arc is below (above) the parabolic arc. But since two different conics that are tangent to each other cannot intersect in more than two other points we conclude that all of the elliptic (hyperbolic) arc from (x_-, y_+) to (x_+, y_+) lies below (above) the parabolic arc. The inequality (8) is an immediate consequence of this. But (8) together with (2) imply (9). Thus, the theorem is proved.

Reference

1. C. B. Boyer, A History of Mathematics, Wiley, New York, London, Sydney, 1968, pp. 142–144.

THE ALTITUDES OF A SIMPLEX ARE ASSOCIATED

LEON GERBER, St. John's University

It is well known that the altitudes of a triangle concur while the altitudes of an n-simplex ($n > 2$) do not concur unless each edge is perpendicular to the edges it does not meet. We give a simple proof of the following result of S. R. Mandan [**1**].

THEOREM. *Any $(n-2)$-dimensional flat which meets n of the altitudes of an n-simplex also meets the remaining altitude. Any set of $n+1$ mutually skew lines in n-space having this property is called an associated set of lines.*

A few remarks about associated lines will illustrate the strength of the theorem. In 3-space, given any three mutually skew lines there exist infinitely many lines associated with the first three; i.e., any line ((3–2)-flat) which meets the first three meets all the others. The lines associated with the given lines form one set of rulings of a hyperboloid of one sheet if there is no plane parallel to all three, and they form the rulings of a hyperbolic paraboloid if there is such a plane [**2**, pp. 14–15]. In 4-space, given any four mutually skew lines there is a unique fifth line associated with them [**3**, pp. 115–116]. An enumerative argument shows that for $n > 4$ there is in general no line associated with n given lines.

We develop some machinery for the proof. Let $\mathscr{I}$ represent the set $\{0, 1, \cdots, n\}$, $\mathscr{I}'$ the set $\{1, \cdots, n\}$, and let the ranges of Σ and Σ' be $\mathscr{I}$ and $\mathscr{I}'$ respectively. Let $\mathscr{A}$

be an n-simplex with corresponding vertices, faces and altitudes given by A_i, $\mathscr{A}_i$, $A_iH_i = h_i$, let $a_i = [a_{i1}, \cdots, a_{in}]$ be the position vector of A_i, and let $a_{i0} = 1$, $i \in \mathscr{I}$.

LEMMA. *Let*

$$D = \begin{vmatrix} 1 & a_{01} & \cdots & a_{0n} \\ 1 & a_{11} & \cdots & a_{1n} \\ & & \cdots & \\ 1 & a_{n1} & \cdot & a_{nn} \end{vmatrix}$$

and let D_{ij} be the cofactor of the ij element, $i, j \in \mathscr{I}$. Then

(i) $\Sigma_i D_{i0} = D = \pm n! \cdot v_n(\mathscr{A}) \neq 0$, *where v_n is the n-dimensional volume function.*

(ii) $\Sigma_j D_{ij} = 0$, $j \in \mathscr{I}'$.

(iii) *The vector $\boldsymbol{u}_i = [D_{i1}, \cdots, D_{in}]$ is normal to $\mathscr{A}_i$, $i \in \mathscr{I}$.*

Proof. Results (i) and (ii) are immediate. To prove (iii) let A_kA_l be any edge of $\mathscr{A}_i$. Then

$$(\boldsymbol{a}_k - \boldsymbol{a}_l) \cdot \boldsymbol{u}_i = \Sigma_j' (a_{kj} - a_{lj})D_{ij} = \Sigma_j (a_{kj} - a_{lj})D_{ij} = 0.$$

COROLLARY 1. *Let $\boldsymbol{f}_i$ be a vector normal to $\mathscr{A}_i$ in the outbound direction with magnitude $v_{n-1}(\mathscr{A}_i)$, $i \in \mathscr{I}$. Then $\Sigma \boldsymbol{f}_i = \boldsymbol{0}$.*

Proof. $\boldsymbol{u}_i$ is a scalar multiple of $\boldsymbol{h}_i$ and

$$\begin{aligned} \boldsymbol{h}_i \cdot \boldsymbol{u}_i &= \overrightarrow{A_iA_k} \cdot \boldsymbol{u}_i \text{ for any } k \neq i \\ &= \Sigma_j (a_{kj} - a_{ij})D_{ij} \\ &= -D, \end{aligned}$$

so the $\boldsymbol{u}_i$ are all outbound or all inbound normals and their lengths are proportional to $|D|/(n-1)!h_i = v_{n-1}(\mathscr{A}_i)$. The result is immediate from (ii).

COROLLARY 2. $v_{n-1}^2(\mathscr{A}_i) = [(n-1)!]^{-2}\Sigma_j' D_{ij}^2$.

COROLLARY 3. *If the edges A_0A_i of a simplex are mutually orthogonal $i \in \mathscr{I}'$ then $v_{n-1}^2(\mathscr{A}_0) = \Sigma' v_{n-1}^2(\mathscr{A}_i)$.*

Proof. We have $-\boldsymbol{f}_0 = \Sigma' \boldsymbol{f}_i$ and $\boldsymbol{f}_i \cdot \boldsymbol{f}_j = 0$, $1 \leqq i < j \leqq n$. Take the dot product of each side with itself.

Now for the proof of the theorem. An $(n-2)$-flat meets a line in a finite point or point at infinity if and only if the $(n-2)$-flat and the line are contained in some prime $(n-1)$-flat. The $(n-2)$-flat determined by the points $\boldsymbol{p}_i = \boldsymbol{a}_i + x_i\boldsymbol{u}_i$ of the altitude A_iH_i, $i = 2, 3, \cdots, n$, is coprime with A_kH_k, $k = 0, 1$, if and only if the set of vectors

$$\overrightarrow{A_kH_k}, \overrightarrow{A_kP_2}, \overrightarrow{A_kP_3}, \cdots, \overrightarrow{A_kP_n}$$

is dependent, that is

$$(1)\qquad \begin{vmatrix} D_{k1} & \cdots & D_{kn} \\ a_{21}+x_2D_{21}-a_{k1} & \cdots & a_{2n}+x_2D_{2n}-a_{kn} \\ & \cdots & \\ a_{n1}+x_nD_{n1}-a_{k1} & \cdots & a_{nn}+x_nD_{nn}-a_{kn} \end{vmatrix} = 0.$$

Adding the appropriate border we see that this is equivalent to

$$\begin{vmatrix} 1 & a_{k1} & \cdots & a_{kn} \\ 0 & D_{k1} & \cdots & D_{kn} \\ 1 & a_{21}+x_2D_{21} & \cdots & a_{2n}+x_2D_{2n} \\ & & \cdots & \\ 1 & a_{n1}+x_nD_{n1} & \cdots & a_{nn}+x_nD_{nn} \end{vmatrix} = 0.$$

With no loss of generality we may assume $a_{0j} = 0, j \in \mathscr{I}'$. Then $D_{00} = D$ and $D_{m0} = 0$, $m \in \mathscr{I}'$. Multiplying by the transpose of D we obtain for $k = 0$

$$\begin{vmatrix} 1 & 1 & 1 & & 1 \\ 0 & -D & -D & \cdots & -D \\ 1 & 1+\Sigma' a_{2j}a_{1j} & 1+\Sigma' a_{2j}a_{2j}+x_2D & \cdots & 1+\Sigma' a_{2j}a_{nj} \\ & & & \cdots & \\ 1 & 1+\Sigma' a_{nj}a_{1j} & 1+\Sigma' a_{nj}a_{2j} & \cdots & 1+\Sigma' a_{nj}a_{nj}+x_nD \end{vmatrix} = 0$$

and for $k = 1$

$$\begin{vmatrix} 1 & 1+\Sigma' a_{1j}a_{1j} & 1+\Sigma' a_{1j}a_{2j} & \cdots & 1+\Sigma' a_{1j}a_{nj} \\ 0 & D & 0 & \cdots & 0 \\ 1 & 1+\Sigma' a_{2j}a_{1j} & 1+\Sigma' a_{2j}a_{2j}+x_2D & \cdots & 1+\Sigma' a_{2j}a_{nj} \\ & & & \cdots & \\ 1 & 1+\Sigma' a_{nj}a_{1j} & 1+\Sigma' a_{nj}a_{2j} & \cdots & 1+\Sigma' a_{nj}a_{nj}+x_nD \end{vmatrix} = 0.$$

Adding D times the first row of the first determinant to the second row we see immediately that one is the negative of the other. Thus the given $(n-2)$-flat is coprime with A_0H_0 if and only if it is coprime with A_1H_1.

If some P_i, $i = 2, \cdots, n$, is a point at infinity, the only change necessary in (1) is that the ij element is x_iD_{ij}, $j \in \mathscr{I}'$ and the argument is identical.

References

1. S. R. Mandan, Altitudes of a general n-simplex, J. Australian Math. Soc., 5(1965) 409–415.
2. D. Hilbert and S. Cohn-Vossen, Geometry and the Imagination, Chelsea, New York, 1956.
3. H. F. Baker, Principles of Geometry IV, Cambridge University Press, 1940.

A NOTE ON CONJUGATE SPACES

GILBERT STEINER and RICHARD BRONSON, Fairleigh Dickinson University

1. Introduction. One of the primary uses of the Hahn-Banach theorem is to prove that the conjugate space of an arbitrary normed vector space contains nontrivial elements. In this note, we present an example which vividly demonstrates the pitfalls in attempting to extend finite dimensional methods to infinite dimensional spaces and, in doing so, illustrates the power of the Hahn-Banach theorem.

2. An example. First consider an n-dimensional normed vector space V over C, the complex numbers. Let $x_0 \in V$, $x_0 \neq 0$, and extend $\{x_0\}$ to a basis $\{x_0, x_1, \cdots, x_{n-1}\}$ for V. Define f: $V \to C$ by $f(x) = f(\alpha_0 x_0 + \sum_{\lambda=1}^{n-1} \alpha_\lambda x_\lambda) = \alpha_0$. Clearly f is a bounded linear functional on V.

We now generalize this construction to an infinite dimensional normed vector space V. Let $x_0 \in V$, $x_0 \neq 0$ and extend $\{x_0\}$ to a Hamel basis for V, $\{x_0\} \cup \{x_\lambda\}$, $\lambda \in \Lambda$, an indexing set. If $x \in V$, it follows that $x = \alpha_0 x_0 + \sum_{\lambda \in \Lambda} \alpha_\lambda x_\lambda$ where (1) only a finite number of the α's are nonzero, and (2) this representation is unique. Define f: $V \to C$ by $f(x) = \alpha_0$. Question: Is f a bounded linear functional on V?

In general, the answer is no. Specifically, consider $V = l_1$ and $x_0 = (1, 0, 0, \cdots)$. Let

$$x_k = (10^k,\ 1/k, 1/k^2, 1/k^3, \cdots),\ k = 2, 3, 4, \cdots.$$

The set $\{x_0, x_2, x_3, \cdots\}$ is linearly independent as the reader may readily verify. Therefore, it can be embedded into a Hamel basis. Now pick $y_k = (0, 1/k, 1/k^2, 1/k^3, \cdots)$; $y_k \in l_1$ and $\| y_k \| = 1/(k-1)$. But

$$|f(y_k)| = |f(-10^k x_0 + x_k)| = 10^k \geqq 10^k \left(\frac{1}{k-1}\right) = 10^k \| y_k \|.$$

Hence, $\| f \| \geqq 10^k$, and f is unbounded.

SELF-GENERATING INTEGERS

BENJAMIN L. SCHWARTZ, McLean, Virginia

1. Introduction. There are only four integers with the property that each one is equal to the sum of the factorials of the digits that represent it in the decimal system. One of these is 145; i.e., $145 = 1! + 4! + 5!$. This result is proved by G. D. Poole in [**2**]. There are also only finitely many integers with the property that each one is equal to the sum of the nth powers of the digits, where n is the number of digits. For example, 153 is a three-digit number; and $153 = 1^3 + 5^3 + 3^3$. The present author proved this in [**4**], and an earlier reference to the same result is in [**3**]. All proofs used similar methods. This paper presents a substantial generalization, encompassing the above two results, among others.

A NOTE ON CONJUGATE SPACES

GILBERT STEINER and RICHARD BRONSON, Fairleigh Dickinson University

1. Introduction. One of the primary uses of the Hahn-Banach theorem is to prove that the conjugate space of an arbitrary normed vector space contains nontrivial elements. In this note, we present an example which vividly demonstrates the pitfalls in attempting to extend finite dimensional methods to infinite dimensional spaces and, in doing so, illustrates the power of the Hahn-Banach theorem.

2. An example. First consider an n-dimensional normed vector space V over C, the complex numbers. Let $x_0 \in V, x_0 \neq 0$, and extend $\{x_0\}$ to a basis $\{x_0, x_1, \cdots, x_{n-1}\}$ for V. Define $f\colon V \to C$ by $f(x) = f(\alpha_0 x_0 + \sum_{\lambda=1}^{n-1} \alpha_\lambda x_\lambda) = \alpha_0$. Clearly f is a bounded linear functional on V.

We now generalize this construction to an infinite dimensional normed vector space V. Let $x_0 \in V$, $x_0 \neq 0$ and extend $\{x_0\}$ to a Hamel basis for V, $\{x_0\} \cup \{x_\lambda\}$, $\lambda \in \Lambda$, an indexing set. If $x \in V$, it follows that $x = \alpha_0 x_0 + \sum_{\lambda \in \Lambda} \alpha_\lambda x_\lambda$ where (1) only a finite number of the α's are nonzero, and (2) this representation is unique. Define $f\colon V \to C$ by $f(x) = \alpha_0$. Question: Is f a bounded linear functional on V?

In general, the answer is no. Specifically, consider $V = l_1$ and $x_0 = (1, 0, 0, \cdots)$. Let

$$x_k = (10^k,\ 1/k, 1/k^2, 1/k^3, \cdots),\ k = 2, 3, 4, \cdots.$$

The set $\{x_0, x_2, x_3, \cdots\}$ is linearly independent as the reader may readily verify. Therefore, it can be embedded into a Hamel basis. Now pick $y_k = (0, 1/k, 1/k^2, 1/k^3, \cdots)$; $y_k \in l_1$ and $\| y_k \| = 1/(k-1)$. But

$$|f(y_k)| = |f(-10^k x_0 + x_k)| = 10^k \geqq 10^k \left(\frac{1}{k-1}\right) = 10^k \| y_k \| .$$

Hence, $\| f \| \geqq 10^k$, and f is unbounded.

SELF-GENERATING INTEGERS

BENJAMIN L. SCHWARTZ, McLean, Virginia

1. Introduction. There are only four integers with the property that each one is equal to the sum of the factorials of the digits that represent it in the decimal system. One of these is 145; i.e., $145 = 1! + 4! + 5!$. This result is proved by G. D. Poole in [**2**]. There are also only finitely many integers with the property that each one is equal to the sum of the nth powers of the digits, where n is the number of digits. For example, 153 is a three-digit number; and $153 = 1^3 + 5^3 + 3^3$. The present author proved this in [**4**], and an earlier reference to the same result is in [**3**]. All proofs used similar methods. This paper presents a substantial generalization, encompassing the above two results, among others.

2. Notation. Let A denote an arbitrary natural number. Suppose that in positional notation with base b, A is represented by the sequence of digits $a_n a_{n-1} \cdots a_1 a_0$. That is $A = \sum_{i=0}^{n} a_i b^i$ where $1 \leqq a_n \leqq b-1$, and $0 \leqq a_i \leqq b-1$ for $i = 0, 1, \cdots, n-1$. If no base b is designated, base ten will be assumed. Let $\{f_n(\cdot)\}$ be a family of functions, $n = 1, 2, 3, \cdots$, with each f_n defined on (at least) the integers $0, 1, \cdots, b-1$. For a given integer A of n digits, let $S(A)$ denote $\sum_{i=0}^{n} f_n(a_i)$. An integer A is a *self-generating integer* (SGI) under $\{f_n\}$ if $A = S(A)$. Finally, let $F_n = \max_{0 \leqq i \leqq b-1} \{f_n(i)\}$.

3. Main result. THEOREM. *With the notation introduced above, suppose that*

$$\limsup_{n \to \infty} (n F_n / b^n) = k < 1 .$$

Then the set of SGI's *under* $\{f_n\}$ *is finite.*

Proof. Choose k' such that $k < k' < 1$. Let A be an n-digit integer, with n sufficiently large that $nF_n/b^n < k'$; and also $(n+1)/n < 1/k'$. Then

$$S(A) = \sum_{i=0}^{n} f_n(a_i) \leqq \sum_{i=0}^{n} F_n = (n+1)F_n$$

$$= \frac{n+1}{n}(nF_n) < (1/k')(k'b^n) = b^n .$$

But

$$A = \sum_{i=0}^{n} a_i b^i \geqq b^n , \text{ since } a_n \geqq 1 .$$

Therefore, for sufficiently large A, we have $S(A) < A$. The theorem follows.

4. Remarks.

(1) The two previously mentioned results are included in the general theorem. For Poole's case, use $f_n(x) = x!$. For the writer's previous result, use $f_n(x) = x^n$. In both cases, $k = 0$.

(2) It is possible to formulate the theorem in terms of the magnitude of A, rather than n, the number of digits in A. This is left to the reader. The form chosen for this paper was preferred, since it is easier to relate the generalization directly to the previous two special cases.

(3) The theorem is the strongest possible result, in the sense that the hypothesis on the value of k cannot be weakened. The reader will have no trouble finding a sequence of functions $\{f_n(\cdot)\}$ for which $\lim(nF_n/b^n) = 1$, and such that the set of SGI's under $\{f_n\}$ is infinite. However, it still might be possible to strengthen the theorem by casting the hypothesis in some form other than a limit condition.

(4) Currently, explicit methods for determining the members of the finite SGI sets are very crude. In [**2**], Poole used a computer to test all candidates after he had determined the upper bound on A. Similarly, in [**1**], Nelson used a computer to determine all SGI's under $\{x^n\}$ having ten digits or less. For this set of $f_n(\cdot)$ functions the best bound known on A is much larger, currently 58 digits.

Either additional examples of SGI's, or better methods to find them than brute force computation are desired.

(5) If we alter the definition of SGI so that the member of the family of functions $\{f_n\}$ which is used in the calculation of $S(A)$ does not depend on the number of digits in A, then it is much harder to characterize the set of SGI's. In [5], the author considered this question, using the class of functions $\{x^n\}$. With this formulation, the integer $A = 14{,}459{,}929$ might be considered self-generating, since

$$A = 1^7 + 4^7 + 4^7 + 5^7 + 9^7 + 9^7 + 2^7 + 9^7.$$

Notice that A has eight digits, but the exponent in the f function is 7. With this extended definition, the question is raised in [5] whether the set of SGI's under $\{x^n\}$ is finite. The answer is still unknown.

References

1. H. L. Nelson, More on PDIs, UCRL–7614, University of California, December 1963.

2. G. D. Poole, Integers and the sum of the factorials of their digits, this MAGAZINE, 44 (1971) 278–79.

3. Max Rumney, Digital invariants, Recreational Mathematics Magazine, **12** (1962) 6–8.

4. B. L. Schwartz, Finiteness of a set of self-generating integers, J. Recreational Mathematics, **2** (1969) 79–83.

5. ———, Finite bounds on digital invariants — some conjectures, J. Recreational Mathematics, **3** (1970) 88–92.

NOTES AND COMMENTS

Beginning with this issue Notes and Comments will appear regularly in each issue instead of just once a year in the December issue, provided of course that we receive items to include in it.

Daniel P. Giesy, Western Michigan University, writes that there is a typographical error in formula (2) in his paper *Still another elementary proof that* $\Sigma 1/k^2 = \pi^2/6$ in the May–June 1972 issue. Formula (2) should read

$$f_n(x) = \frac{\sin[(2n+1)x/2]}{2\sin(x/2)}. \tag{2}$$

David Singmaster, Polytechnic of the South Bank, London, SE 1, writes regarding Heuer's article *Continuous multiplication in* R^2 in the March 1972 issue that Heuer's Theorem 6.3B which asserts that for $G = Z_p$ there are exactly 8 nonisomorphic rings with $R^+ = Z_p \oplus Z_p$ may be combined with other known results (this MAGAZINE, 40 (1967) 83–85; Amer. Math. Monthly, 71 (1964) 449–50; L. Fuchs, *Abelian groups*, p. 263) to yield the following theorem: *For any prime p, there are exactly* 11 *distinct rings of order* p^2.

We know that there are exactly 2 distinct rings of order p. Now any ring of order n is a direct sum of its primary components so that if $n = p_1 p_2 \cdots p_r q_1^2 q_2^2 \cdots q_s^2$, then

Either additional examples of SGI's, or better methods to find them than brute force computation are desired.

(5) If we alter the definition of SGI so that the member of the family of functions $\{f_n\}$ which is used in the calculation of $S(A)$ does not depend on the number of digits in A, then it is much harder to characterize the set of SGI's. In [5], the author considered this question, using the class of functions $\{x^n\}$. With this formulation, the integer $A = 14{,}459{,}929$ might be considered self-generating, since

$$A = 1^7 + 4^7 + 4^7 + 5^7 + 9^7 + 9^7 + 2^7 + 9^7.$$

Notice that A has eight digits, but the exponent in the f function is 7. With this extended definition, the question is raised in [5] whether the set of SGI's under $\{x^n\}$ is finite. The answer is still unknown.

References

1. H. L. Nelson, More on PDIs, UCRL–7614, University of California, December 1963.
2. G. D. Poole, Integers and the sum of the factorials of their digits, this MAGAZINE, 44 (1971) 278–79.
3. Max Rumney, Digital invariants, Recreational Mathematics Magazine, **12** (1962) 6–8.
4. B. L. Schwartz, Finiteness of a set of self-generating integers, J. Recreational Mathematics, **2** (1969) 79–83.
5. ———, Finite bounds on digital invariants — some conjectures, J. Recreational Mathematics, **3** (1970) 88–92.

NOTES AND COMMENTS

Beginning with this issue Notes and Comments will appear regularly in each issue instead of just once a year in the December issue, provided of course that we receive items to include in it.

Daniel P. Giesy, Western Michigan University, writes that there is a typographical error in formula (2) in his paper *Still another elementary proof that* $\Sigma 1/k^2 = \pi^2/6$ in the May–June 1972 issue. Formula (2) should read

$$f_n(x) = \frac{\sin[(2n+1)x/2]}{2\sin(x/2)}. \tag{2}$$

David Singmaster, Polytechnic of the South Bank, London, SE 1, writes regarding Heuer's article *Continuous multiplication in* R^2 in the March 1972 issue that Heuer's Theorem 6.3B which asserts that for $G = Z_p$ there are exactly 8 nonisomorphic rings with $R^+ = Z_p \oplus Z_p$ may be combined with other known results (this MAGAZINE, 40 (1967) 83–85; Amer. Math. Monthly, 71 (1964) 449–50; L. Fuchs, *Abelian groups*, p. 263) to yield the following theorem: *For any prime p, there are exactly* 11 *distinct rings of order* p^2.

We know that there are exactly 2 distinct rings of order p. Now any ring of order n is a direct sum of its primary components so that if $n = p_1 p_2 \cdots p_r q_1^2 q_2^2 \cdots q_s^2$, then there are $2^r 11^s$ distinct rings of order n. Further, we now know all rings of order n for $n < 27$, except for $n = 8$.

Of the rings of order p^2:

3 have $R^+ = Z_{p^2}$ (all these are commutative and one has identity);
8 have $R^+ = Z_p \oplus Z_p$;
9 are commutative;
1 is a field;
2 have null multiplication.

He also suggests as a problem finding the rings of order p^3 or even just those of order 8 and cites Amer. Math. Monthly, 71 (1964) 918–20 and Nederl. Akad. Wetensch. Proc. Ser. A, 68 (1965) 632–45 and 69 (1966) 14–21 as references for related work on finite rings.

Robert E. Dressler, Kansas State University, notes regarding *Inequalities for* $\sigma(n)$ *and* $\phi(n)$ by Annapurna, (September 1972) that LeVeque ([**2**], Theorems 6–26 to 6–28) shows that

$$\frac{\phi(n)}{n} > \frac{c}{\log\log n}, \quad \text{for } n > 3, \tag{1}$$

and subsequently that

$$\frac{\sigma(n)}{n} < \frac{\log\log n}{c}, \quad \text{for } n > 3.$$

This answers the question raised by Annapurna whether the inequality, $\phi(n) \geqq \sqrt{n}$ except for $n = 2$ or $n = 6$, can be sharpened. Dressler also notes that c may be taken to be $1/122$ (by analyzing the proofs) and that (1) is best possible, as an exercise in LeVeque (p. 116) shows.

Roger Osborn, University of Texas at Austin, notes that the inequality $\phi(n) \geqq \sqrt{n}$ which Annapurna credits to Vaidya, Math. Student, 35 (1967) 79–80, appeared in an earlier paper *Two simple lower bounds for the Euler ϕ-function* by Kendall and Osborn, Texas Journal of Science, Vol. 17, No. 3, Sept. 1965, which gave the two lower bounds

$$\phi(n) > \sqrt{n} \quad \text{for } n > 6$$

and

$$\phi(n) > n^{2/3} \quad \text{for } n > 30.$$

From K. O. May, University of Toronto: "The bicycle problem, so concisely solved by D. E. Daykin in this MAGAZINE, 45 (1972) 1, may be illuminated still further from a kinematic point of view. When a bicycle moves along a level surface without free-wheeling, a point on one of the pedals describes a cycloid. This is evident because there must be some circle centered at the sprocket hub that rolls along a horizontal line. The curve is a "prolate" cycloid traced out by a point within this circle, a common cycloid with vertical cusps traced out by a point on the circumference, or a "curtate" cycloid with loops traced out by a point outside the circle, according as GR is greater than, equal to, or less than r, where G is the gear ratio of the bicycle (ratio of the angular velocity of the rear wheel to that of the pedal sprocket), R is the radius of the driving wheel, and r is the radius of the pedal. Since retrograde motion for a point on a cycloid occurs only in the curtate case, and $GR > r$ for the common bicycle, it is clear that a bicycle pedal always moves in the same direction as the bicycle, and hence that a backward push on a pedal cannot move the bicycle forward."

From Michael Golomb, Purdue University: In his article *A solution of Laplace's equation for a semi-infinite strip* (November 1972), C. R. Edstrom considers the problem $\nabla^2 u(x,y) = 0$ for $x > 0$, $0 < y < L$ with the boundary conditions $u(x,0) = u(0,y) = 0$, $u(x,L) = f(x)$. He shows that, although this problem does not belong to the class to which the usual method of separation of variables applies, it can be converted into such a problem if $f(x)$ satisfies the following condition: (E) There exists an integer N such that each of the derivatives of order $N+1$ and $N+2$ of the function $f(x)$ can be expressed as a linear combination of the function $f(x)$ and its first N derivatives.

The method requires the solution of a system of linear homogeneous second order differential equations for functions $G_0(y), \cdots, G_N(y)$, with boundary conditions $G_n(0) = 0$ $(n = 0, \cdots, N)$, $G_m(L) = 0$ $(m = 1, \cdots, N)$, $G_0(L) = 1$.

A much simpler and more transparent method is the following. We set $u(x,y) = U(x,y) + w(x,y)$ and determine $w(x,y)$ so that

$$(1) \qquad \nabla^2 w(x,y) = 0, \; w(x,0) = 0, \; w(x,L) = f(x), \; x > 0, \qquad 0 < y < L.$$

Then $U(x,y)$ is the solution of

$$(2) \quad \nabla^2 U(x,y) = 0, \; U(x,0) = U(x,L) = 0, \; U(0,y) = -w(0,y), \; x > 0, \, 0 < y < L$$

which is the usual problem for the method of separation of variables. We assert that (1) can be easily solved if $f(x)$ satisfies condition (E). This condition simply says that $f(x)$ is the solution of a linear homogeneous differential equation with constant coefficients of order $N+1$, hence is a linear combination of functions $x^m e^{\alpha x}$, where m is a nonnegative integer and α a (possibly complex) number. Since obviously superposition applies it suffices to assume $f(x)$ is of the form

$$(3) \qquad f(x) = P(x)e^{\alpha x}$$

where $P(x)$ is a polynomial of degree n. We first observe that if Q is any polynomial and we write

$$Q_1(x,y) = \frac{1}{2}[Q(x+iy) + Q(x-iy)],$$

$$Q_2(x,y) = \frac{1}{2i}[Q(x+iy) - Q(x-iy)]$$

then $\nabla^2 Q_1(x,y) = 0$, $\nabla^2 Q_2(x,y) = 0$. This is obvious to anyone familiar with the elements of complex variable theory, but is also immediately verified by differentiating

$$\frac{1}{2}[(x+iy)^m+(x-iy)^m],\ \frac{1}{2i}[(x+iy)^m-(x-iy)^m].$$

Therefore

$$w(x,y)=[Q_1(x,y)\sin\alpha y+Q_2(x,y)\cos\alpha y]e^{\alpha x} \tag{4}$$

satisfies $\nabla^2 w(x,y)=0$ and also $w(x,0)=0$, by definition of Q_2. If $\sin\alpha L\neq 0$ (the first case of Edstrom's method) then we choose the polynomial $Q(z)=\sum_{m=0}^{n} a_m z^m$ so that

$$Q_1(x,L)\sin\alpha L+Q_2(x,L)\cos\alpha L=P(x). \tag{5a}$$

This is an equation which gives the coefficients a_n, $a_{n-1},\cdots,a_0$ recursively (observe that the leading coefficient of $Q_1(x,L)$ is a_n, that of $Q_2(x,L)$ is 0). If $\sin\alpha L=0$, $L=p\pi$, then we choose the polynomial $Q(z)=\sum_{m=0}^{n} a_m z^{m+1}$ so that

$$Q_2(x,L)=(-1)^p P(x). \tag{5b}$$

Again the coefficients can be found recursively (that of x^n on the left being $(n+1)a_nL$). In either case, $w(x,y)$ as given by (4) is the solution of (1).

In the second example of Edstrom's article $f(x)=xe^{-x}$ and $L=p\pi$. Hence $\alpha=-1$, $\sin\alpha L=0$, $P(x)=x$, and we choose $Q(z)=a_1z^2+a_2z$. Then (5b) becomes

$$a_1[(x+iL)^2-(x-iL)^2]+a_2[(x+iL)-(x-iL)]=2i(-1)^p x$$

which gives immediately $a_1=(-1)^p/2L$, $a_2=0$. Therefore,

$$w(x,y)=\frac{(-1)^p}{2p\pi}[-(x^2-y^2)\sin y+2xy\cos y]e^{-x}.$$

BOOK REVIEWS

EDITED BY ADA PELUSO AND WILLIAM WOOTON

Materials intended for review should be sent to: Professor Ada Peluso, Department of Mathematics, Hunter College of CUNY, 695 Park Avenue, New York, New York 10021, or to Professor William Wooton, 1495 La Linda Drive, Lake San Marcos, California 92069.

A boldface capital **C** *in the margin indicates that a review is based in part on classroom use.*

In Mathematical Circles, Quadrants I and II. xvii + 136 pp.]
In Mathematical Circles, Quadrants III and IV. xi + 143 pp.] \$15.00
Mathematical Circles Revisited. xxii + 186 pp. \$10.00.
Mathematical Circles Squared. xxi + 186 pp. \$10.00.

By Howard W. Eves. Prindle, Weber & Schmidt, Boston, Massachusetts, 1969, 1969, 1971, 1972.

differentiating

$$\frac{1}{2}[(x+iy)^m+(x-iy)^m],\ \frac{1}{2i}[(x+iy)^m-(x-iy)^m].$$

Therefore

$$w(x,y)=[Q_1(x,y)\sin\alpha y+Q_2(x,y)\cos\alpha y]e^{\alpha x} \tag{4}$$

satisfies $\nabla^2 w(x,y)=0$ and also $w(x,0)=0$, by definition of Q_2. If $\sin\alpha L\neq 0$ (the first case of Edstrom's method) then we choose the polynomial $Q(z)=\sum_{m=0}^{n} a_m z^m$ so that

$$Q_1(x,L)\sin\alpha L+Q_2(x,L)\cos\alpha L=P(x). \tag{5a}$$

This is an equation which gives the coefficients $a_n, a_{n-1},\cdots,a_0$ recursively (observe that the leading coefficient of $Q_1(x,L)$ is a_n, that of $Q_2(x,L)$ is 0). If $\sin\alpha L=0$, $L=p\pi$, then we choose the polynomial $Q(z)=\sum_{m=0}^{n} a_m z^{m+1}$ so that

$$Q_2(x,L)=(-1)^p P(x). \tag{5b}$$

Again the coefficients can be found recursively (that of x^n on the left being $(n+1)a_nL$). In either case, $w(x,y)$ as given by (4) is the solution of (1).

In the second example of Edstrom's article $f(x)=xe^{-x}$ and $L=p\pi$. Hence $\alpha=-1$, $\sin\alpha L=0$, $P(x)=x$, and we choose $Q(z)=a_1z^2+a_2z$. Then (5b) becomes

$$a_1[(x+iL)^2-(x-iL)^2]+a_2[(x+iL)-(x-iL)]=2i(-1)^p x$$

which gives immediately $a_1=(-1)^p/2L$, $a_2=0$. Therefore,

$$w(x,y)=\frac{(-1)^p}{2p\pi}[-(x^2-y^2)\sin y+2xy\cos y]e^{-x}.$$

BOOK REVIEWS

EDITED BY ADA PELUSO AND WILLIAM WOOTON

Materials intended for review should be sent to: Professor Ada Peluso, Department of Mathematics, Hunter College of CUNY, 695 Park Avenue, New York, New York 10021, or to Professor William Wooton, 1495 La Linda Drive, Lake San Marcos, California 92069.

A boldface capital **C** *in the margin indicates that a review is based in part on classroom use.*

In Mathematical Circles, Quadrants I and II. xvii + 136 pp.] $15.00
In Mathematical Circles, Quadrants III and IV. xi + 143 pp.]
Mathematical Circles Revisited. xxii + 186 pp. $10.00.
Mathematical Circles Squared. xxi + 186 pp. $10.00.

By Howard W. Eves. Prindle, Weber & Schmidt, Boston, Massachusetts, 1969, 1969, 1971, 1972.

Someone has said that mathematics without history is mathematics stripped of its greatness. Mindful of this admonition, Howard Eves, the distinguished geometer and mathematical folklorist from "Down East," has served up a palatable smorgasbord of real and imaginary historical vignettes relating to mathematics and mathematicians, thereby providing extensive nourishment to students, teachers, and even the laity at almost all levels. In this handsomely bound four-volume set Professor Eves guides the reader three times around a circle. However, any deprecatory association of the act of "going around in circles" with the notion of futility would be ungraciously inappropriate.

Each trip has a different orientation and each is eminently worthwhile. In the first two volumes the development is chronological. *Mathematical Circles Revisited* is organized by subject matter, while *Mathematical Circles Squared* is to a large extent sectioned off with respect to national origins.

The term *smorgasbord* was used advisedly since the fare is highly diversified, ranging from the light touch — including "double entendres" (some pretty far out!) — to the serious, from "tall tales" to authentic history, from events to personalities, from animals to people.

Maine with its salty, invigorating climate seems to inspire zestful writing, e. g., E. B. White, Kenneth Roberts, et cetera, and Professor Eves honors this rule mostly in the observance and not in the breach.

The books are so full of so many tidbits (take three times three-sixty for an honest count!) that any attempt at an outline of contents is out of the question. In accordance with this reviewer's taste, a few of the highlights are: "Plimpton 322" and "Hindu Embroidery" in Vol. I; "What became of Karl Feuerbach?" in Vol. II; "Croutons for the French Soup" and "The Prince of Mathematicians" in *Mathematical Circles Squared.*

A recent article in the Mathematics Teacher has, with just cause, deplored the neglect of history, cultural contexts, and biographical settings in the teaching of mathematics, citing many reputable institutions as having no curricular offerings of any sort in this area. While this set of volumes is not intended to provide a scholarly background for the serious mathematics student (the author has done this in another book), it is certainly a "whetter of appetites."

Lest the potential reader be deterred by the prospect of a four-volume marathon, let him be reassured by this observation. Pick up any volume at any page and there is a savory morsel. You may read on rewardingly for three minutes or three hours!

J. E. YARNELLE, Hanover College, Hanover, Indiana

Applied Mathematics: An Introduction. By Harry Pollard. Addison-Wesley, Reading, Mass., 1972.

This book is an introduction to classical applied mathematics. It was written in an attempt "to convince the advanced undergraduate or beginning graduate student, at a critical point in his studies, that applied mathematics is interesting." This reviewer is in hearty agreement that applied mathematics is vitally important to society and

Someone has said that mathematics without history is mathematics stripped of its greatness. Mindful of this admonition, Howard Eves, the distinguished geometer and mathematical folklorist from "Down East," has served up a palatable smorgasbord of real and imaginary historical vignettes relating to mathematics and mathematicians, thereby providing extensive nourishment to students, teachers, and even the laity at almost all levels. In this handsomely bound four-volume set Professor Eves guides the reader three times around a circle. However, any deprecatory association of the act of "going around in circles" with the notion of futility would be ungraciously inappropriate.

Each trip has a different orientation and each is eminently worthwhile. In the first two volumes the development is chronological. *Mathematical Circles Revisited* is organized by subject matter, while *Mathematical Circles Squared* is to a large extent sectioned off with respect to national origins.

The term *smorgasbord* was used advisedly since the fare is highly diversified, ranging from the light touch — including "double entendres" (some pretty far out!) — to the serious, from "tall tales" to authentic history, from events to personalities, from animals to people.

Maine with its salty, invigorating climate seems to inspire zestful writing, e. g., E. B. White, Kenneth Roberts, et cetera, and Professor Eves honors this rule mostly in the observance and not in the breach.

The books are so full of so many tidbits (take three times three-sixty for an honest count!) that any attempt at an outline of contents is out of the question. In accordance with this reviewer's taste, a few of the highlights are: "Plimpton 322" and "Hindu Embroidery" in Vol. I; "What became of Karl Feuerbach?" in Vol. II; "Croutons for the French Soup" and "The Prince of Mathematicians" in *Mathematical Circles Squared.*

A recent article in the Mathematics Teacher has, with just cause, deplored the neglect of history, cultural contexts, and biographical settings in the teaching of mathematics, citing many reputable institutions as having no curricular offerings of any sort in this area. While this set of volumes is not intended to provide a scholarly background for the serious mathematics student (the author has done this in another book), it is certainly a "whetter of appetites."

Lest the potential reader be deterred by the prospect of a four-volume marathon, let him be reassured by this observation. Pick up any volume at any page and there is a savory morsel. You may read on rewardingly for three minutes or three hours!

J. E. YARNELLE, Hanover College, Hanover, Indiana

Applied Mathematics: An Introduction. By Harry Pollard. Addison-Wesley, Reading, Mass., 1972.

This book is an introduction to classical applied mathematics. It was written in an attempt "to convince the advanced undergraduate or beginning graduate student, at a critical point in his studies, that applied mathematics is interesting." This reviewer is in hearty agreement that applied mathematics is vitally important to society and to mathematics: to the first, as at least a component if not a source of solutions to vexing problems, and to the second, as a pointer to what is most vital and important for mathematics. For abstraction for abstraction's sake spawns much sterile mathematics, already dead when published. The most illustrious topologists, e. g., R. Bott and S. Smale, are close to the geometrically concrete and the world of natural phenomena.

But is it not strange that in 1972 an introduction to applied mathematics is written that is based on old-fashioned advanced calculus and includes only mathematics and applications known before 1930, indeed mostly before 1900, with the single exception of two pages devoted to linear programming? Falling bodies, path of quickest descent, the two-body problem, surface of revolution of least area, Hamilton's and Lagrange's equations and the n-body problem, Laplace's equation and the Dirichlet principle, the wave equation and the vibrating membrane, and the heat equation and the finite rod, infinite rod and cooling sphere: surely these are problems of a bygone age that unless given a contemporary twist may fail to stimulate the mind of the potentially interested student.

The section on the system

$$dx/dt = Ax - Bxy \qquad dy/dt = Cxy - Dy$$

is given an interpretation from Marxist economics. This is one of the more potentially interesting sections of the book, for this system and its generalizations have important applications to a variety of problems from animal prey-predator dynamics to biochemical reaction kinetics, including a model for life itself. Yet no mention of this is made. Simple examples from the theory of optimal control, for example, should have been included, even if one wished only to assume a background in old-fashioned engineering advanced calculus. To mention modern results on energy decay for solutions to the wave equation might pique more interest than Fourier's analysis of a vibrating string. But if one wishes to vibrate strings: how about vibrating a real string? The reviewer's point is not to have wished that his favorite topics were included in the book, but rather to point out that he would have desired any set with a contemporary flavor over a set almost entirely found in classical textbooks.

Mathematics, and applied mathematics especially, can only restore itself as the Queen of the Sciences in the eyes of both society and the student if it addresses contemporary problems from the social sciences, bio-medicine, and engineering. Fortunately, classical analysis, algebra and geometry have much to contribute here. Even more fortunately, new mathematics must and will be found to contribute to their solutions and understanding. The reviewer encourages the mathematical community to produce several little books called "Introduction to Applied Mathematics" written from this point of view.

N. D. KAZARINOFF, State University of New York at Buffalo

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles Valley College

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

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

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

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

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

To be considered for publication, solutions should be mailed before November 1, 1973.

PROBLEMS

866. *Proposed by Richard L. Breisch, Pennsylvania State University.*

Solve the congruence cryptarithm $LIFE \equiv SIZE \pmod{ELS}$ in base 6 with E, L and S nonzero.

867. *Proposed by L. Carlitz, Duke University.*

Let P be a point in the interior of the triangle ABC. Let R_1, R_2, R_3, denote the distances of P from the vertices of ABC and let r_1, r_2, r_3, denote the distances from P to the sides of ABC. Show that

$$(1) \qquad \Sigma r_1 R_2 R_3 \geqq 12 r_1 r_2 r_3$$

$$(2) \qquad \Sigma r_1 R_1^2 \geqq 12 r_1 r_2 r_3$$

$$(3) \qquad \Sigma r_2^2 r_3^2 R_2 R_3 \geqq 12 r_1^2 r_2^2 r_3^2 \,.$$

In each case there is equality if and only if ABC is equilateral and P is the center of ABC.

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

Two men undertake to arrive at a rendezvous independently, but each would arrive sometime between noon and x minutes after noon. One promises to wait y minutes, the other z minutes, but neither will stay beyond x minutes after noon. What is the chance that they will meet?

869. *Proposed by Roy Dubisch, East African Regional Mathematics Program, Addis Ababa, Ethiopia.*

Prove that in the expansion of $(x-1)(x-2)\cdots(x-k)$ the coefficient of x^{k-2} is $k(k+1)(k-1)(3k+2)/24$.

870. *Proposed by Everett Casteel, Bethel College, Minnesota.*

Given the number $2^{n-1}(2^n-1)$ where n is an odd integer greater than or equal to 3, prove that $2^{n-1}(2^n-1) \equiv 1 \pmod 9$.

871. *Proposed by Donald P. Minassian, Butler University, Indiana.*

A subgroup H of a fully ordered group G is convex if H contains all g in G such that $h \leqq g \leqq k$ whenever $h \leqq k$ are both in G. Let H be a proper subgroup of an abelian group G whose only element of finite order is the identity. Show that G admits an ordering under which H is not convex.

872. *Proposed by Warren Page, New York City Community College.*

Let $x = \sum_{i=1}^{n} \alpha_i u_i^2$, where u_i are integers and $\alpha_1 = 1$. Prove that for every natural number k one has $x^{2^k} = \sum_{i=1}^{n} \alpha_i w_i^2$, where w_i are integers.

ERRATA

On page 283 (November 1972) in problem 846 the fact that the railroad terminal, the opera house and the cathedral were collinear was omitted. Problem 850 should be summed to infinity. The last line of Problem 851 should read "show that p divides $(a^{2k}-1)/(a^2-1)$."

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.

Q567. If k is a positive integer, prove that $(6^{16k+2}/2) - 1$ is not a prime.

[*Submitted by Erwin Just*]

Q568. Solve the equation $x^{2n} + x^{2n-2} + \cdots + x^{2r} + \cdots + x^2 + 1 = x^n$.

[*Submitted by Murray S. Klamkin*]

Q569. Prove or disprove: If $\lim_{n\to\infty} n^p a_n = 0$ then there is an $\varepsilon > 0$ such that $0 \leqq \lim_{n\to\infty} n^{p+\varepsilon} a_n < \infty$.

[*Submitted by John Beidler*]

Q570. Let $f_{\bar{X}\bar{Y}}(x, y)$ be an absolutely continuous joint probability density function which is zero except in a connected nonrectangular closed region A of the plane. Prove that $\bar{X}$ and $\bar{Y}$ must be dependent random variables.

[*Submitted by L. Franklin Kemp*]

Q571. Show that every convergent net in a topological space $\bar{X}$ has a unique limit iff the space is a Hausdorff space.

[*Submitted by Simeon Reich*]

Q572. Show that if n and k are positive integers then

$$x^n + y^n = z^{n+1/k}$$

always has solutions in integers x, y, z.

[*Submitted by Norman Schaumberger*]

(Answers on page 176)

SOLUTIONS

Late Solutions

Jorge Andres, Brooklyn, New York: **873**; *Gladwin Bartel, La Junta, Colorado:* **873**; *V. S. Blanco, University of South Alabama:* **873**; *M. G. Greening, University of New South Wales, Australia:* **831, 837**; *Heiko Harborth, Braunschweig, Germany:* **833**; *Karl Heuer, Moorehead, Minnesota:* **835, 837, 838**; *J. A. H. Hunter, Toronto, Canada:* **844**; *Lew Kowarski, Morgan State College, Maryland:* **843**; *J. E. Mueller and P. Hartung, Bloomsburg State College, Pennsylvania:* **844**; *Albert J. Patsche, Rock Island Arsenal, Illinois:* **837**; *J. H. Nagaraja Ras, Andhra University, Waltair, India:* **833**; *Ben Sapolsky, Bok Vocational-Technical School, Philadelphia, Pennsylvania:* **835, 837**; *David R. Stone, Georgia Southern College:* **833, 834, 835**; *Bessie F. Tarpley, Bennett College, North Carolina:* **837**; *James Tattersall, Attleboro, Massachusetts:* **833, 835, 837**; *Paul Theil, Thomas More College, Kentucky:* **833**; *Phil Tracy, Liverpool, New York:* **844.**

Depleting the Bank Account

838. [September, 1972] *Proposed by Stanley Rabinowitz, Far Rockaway, New York.*

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

Solution by Kenneth M. Wilke, Topeka, Kansas.

Let B_0 denote the original amount in the account, w_i the ith withdrawal and B_i

the balance remaining after the ith withdrawal. These quantities are connected by the following relations:

$$B_i = B_{i-1}\left(1 - \frac{i}{100n^2}\right) = B_0 \prod_{j=1}^{i} (1 - j/100n^2)$$
$$w_i = B_{i-1} - B_i = B_{i-1}\frac{i}{100n^2}.$$

An examination of numerical examples reveals that w_i increases in the beginning; hence the largest withdrawal w_k occurs when $w_k > w_{k+1}$. This is equivalent to

$$B_{k-1}\frac{k}{100n^2} > B_k\frac{k+1}{100n^2} = B_{k-1}\frac{k+1}{100n^2}\left(1 - \frac{k}{100n^2}\right)$$

or $k^2 + k > 100n^2$ which requires $k \geqq 10n$ since n is an integer. A similar examination of $w_{k-1} < w_k$ reveals $k \leqq 10n$. Hence the size of the withdrawals increases until a maximum withdrawal is made on the $10n$th trip (the last trip of the 10th day) and decreases thereafter.

Also solved by Gladwin Bartel, La Junta, Colorado; Richard A. Gibbs, Fort Lewis College, Colorado; Michael Goldberg, Washington, D. C.; Vaclav Konecny, Jarvis Christian College, Texas; Lew Kowarski, Morgan State College, Maryland; R. Shantaram, University of Michigan-Flint; Frank Vertosick, Natrona Heights, Pennsylvania; and the proposer.

Probability of No Change

839. [September, 1972] *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey.*

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

Solution by Thomas Spencer, Trenton State College, New Jersey.

A moment's reflection will show that the only events which will leave the color composition of all three boxes unchanged are the choices white, white, white or red, red, red. Their probabilities by the multiplication rule are:

$$\left(\frac{w}{w+r}\right)\left(\frac{w+1}{w+r+1}\right)\left(\frac{w+1}{w+r+1}\right) = \frac{w(w+1)^2}{(w+r)(w+r+1)^2}$$

and

$$\left(\frac{r}{w+r}\right)\left(\frac{r+1}{w+r+1}\right)\left(\frac{r+1}{w+r+1}\right) = \frac{r(r+1)^2}{(w+r)(w+r+1)^2}$$

respectively. These events are disjoint and thus the required probability is their sum:

$$\frac{w(w+1)^2+r(r+1)^2}{(w+r)(w+r+1)^2}.$$

Note: In the case of N boxes the solution would be:

$$\frac{w(w+1)^{N-1}+r(r+1)^{N-1}}{(w+r)(w+r+1)^{N-1}}.$$

Also solved by Gladwin Bartel, La Junta, Colorado; Melvin Billick, Midland High School, Michigan; J. L. Brown, Jr., Pennsylvania State University; Joseph B. Browne, Oklahoma State University; Daniel L. Calloway, Ashville, North Carolina; Abraham L. Epstein, Hanscom Field, Massachusetts; George Fabian, Park Forest, Illinois; Michael Goldberg, Washington, D. C.; Kathleen Harris, New Hampton, Iowa; Karl Heuer, Moorhead, Minnesota; John M. Howell, Littlerock, California; Vaclav Konecny, Jarvis Christian College, Texas; Lew Kowarski, Morgan State College, Maryland; Michael W. O'Donnell, Carnegie-Mellon University; George Pfeiffer, Old Dominion University, Virginia; Louisa Russo, Michigan Technological University; R. Shantaram, University of Michigan-Flint; and the proposer.

Triangular Palindromes

840. [September, 1972] *Proposed by Charles W. Trigg, San Diego, California.*

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

Solution by E. P. Starke, Plainfield, New Jersey.

Let $T_r = \frac{1}{2}r(r+1)$, $r = 1,2,\cdots$ be the triangular numbers, and let a (*all* numerals will be written in the scale of 5) be the n-digit number

$$a = 222\cdots 2 = \tfrac{1}{2}(10^n - 1).$$

Then $T_a = 30303\cdots 303 = 3(10^{2n}-1)/(10^2-1)$ is a palindromic triangular number for all values of n.

The calculations are simple: $(a+1) = (10^n+1)/2$ and then

$$T_a = \frac{a(a+1)}{2} = \frac{(10^n-1)(10^n+1)}{13} = \frac{3(10^{2n}-1)}{44} = \frac{3(10^{2n}-1)}{10^2-1}.$$

These are by no means all the palindromic triangular numbers. E.g., $T_3 = 11$, $T_{13} = 121$, $T_{102} = 3003$, $T_{1303} = 1130311, \ldots$.

Also solved by Larla W. Hashek, Montebello, California; Thomas Moore, Plymouth, Maine; and the proposer.

A Generalization of Clairaut's Equation

841. [September, 1972] *Proposed by Murray S. Klamkin, Ford Motor Company, Dearborn, Michigan.*

Solve the following generalization of Clairaut's equation:

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

where $p = dy/dx$.

Solution by the proposer.

Let $r = 1 + \sqrt{1 + xG(p)}$ and differentiate (1) with respect to x, giving

$$xp' + rp'F' + r'F = 0.$$

Now replace r' by $p'\dfrac{dr}{dp}$ and x by $(r^2 - 2r)/G$ to give

$$p'\left\{(r^2 - 2r)/G + rF' + F\frac{dr}{dp}\right\} = 0. \tag{2}$$

If the first factor is zero, i.e., $p' = 0$, we get $y = cx + F(c)\,\{1 + \sqrt{1 + xG(c)}\}$ provided that $F(c)G(c) = 0$.

The other factor can be rewritten as

$$\left\{D_p + \frac{2}{FG} - \frac{F'}{F}\right\}\frac{1}{r} = \frac{1}{FG}.$$

Whence,

$$\frac{1}{r} = F\exp\left\{-2\int\frac{dp}{FG}\right\}\int\frac{\exp\left\{2\int\frac{dp}{FG}\right\}dp}{F^2G}. \tag{3}$$

(3) gives x as a function of p. Then (1) gives y also as a function of p. These two latter equations give the solution in parametric form.

Also solved by Kim R. Penrose, Montana State University; and W. Weston Meyer, General Motors Research Laboratories, Warren, Michigan.

A Property of Quadratic Monic Polynomials

842. [September, 1972] *Proposed by Kenneth Fogarty and Erwin Just, Bronx Community College, New York.*

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

Solution by Leonard Carlitz, Duke University.

Put

$$f(x) = (x - a)(x - b) = x^2 - (a + b)x + ab.$$

Then each of the equations

$$x^2 - (a + b)x + ab = a, \quad x^2 - (a + b)x + ab = b$$

has rational integral solutions. Hence

$$\begin{cases}(a+b)^2-4(ab-a) = u^2\\(a+b)^2-4(ab-b) = v^2,\end{cases}$$

where u, v are rational integers. Thus

$$\begin{cases}(a-b)^2+4a = u^2\\(a-b)^2+4b = v^2,\end{cases}$$

so that $a-b=rs$, where

$$u+v=2r'$$

$$u-v=2s'$$

and r', s' are rational integers. It follows that

$$4a=(r'+s'+r's')(r'+s'-r's'),$$

so that r', s' are both even. Put $r'=2r$, $s'=2s$. Then

$$\begin{cases}a = (r+s+2rs)(r+s-2rs)=(r+s)^2-4r^2s^2\\b = (r+s+2rs)(r+s-2rs)-4rs=(r-s)^2-4r^2s^2.\end{cases}$$

Therefore the four roots are the roots of

$$x^2-2(r^2+s^2-4r^2s^2)x+(r^2+s^2-4r^2s^2)^2=(r+s)^2.$$

Thus the roots are

(*) $$r^2+s^2-4r^2s^2+r+s.$$

For example for $r=1$, $s=2$ we get

$$-8,\ -10,\ -12,\ -14.$$

Remark. It follows from (*) that no root can be positive. If we assume a positive root then (*) implies (with $r, s>0$)

$$4r+4s>(4r^2-1)(4s^2-1)\geqq(4r-1)(4s-1)=16rs-4r-4s+1,$$

$$0>16rs-8r-8s+1,$$

$$4>(4r-2)(4s-2)+1.$$

If $rs=0$, the roots are not distinct.

Also solved by John O'Neill, University of Detroit; E. P. Starke, Plainfield, New Jersey; Phil Tracy, Liverpool, New York; Kenneth M. Wilke, Topeka, Kansas; Kenneth L. Yocom, South Dakota State University; and the proposers.

Circles on a Parabola

843. [September, 1972] *Proposed by Leon Bankoff, Los Angeles, California.*

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

Solution by Michael Goldberg, Washington, D.C.

With w_0 as origin, and with w_0F as the x-axis, the equation of the parabola is $y = 2\sqrt{p_0 x}$. The coordinates of w_1 are $(p_0 - p_1, 2\sqrt{p_0^2 - p_0p_1})$. Then, $(w_0w_1)^2 = (p_0 + p_1)^2 = (p_0 - p_1)^2 + 4p_0^2 - 4p_0p_1$, from which $4p_0p_1 = 4p_0^2 - 4p_0p_1$, and $p_1 = p_0/2$. Similarly, $(w_1w_2)^2 = (p_1 - p_2)^2 + \{2\sqrt{p_0^2 - p_0p_2} - 2\sqrt{p_0^2 - p_0p_1}\}^2$, from which

$$4(p_0 - p_2)(p_0 - p_1) = (2p_0 - p_2 - p_1 - p_1p_2/p_0)^2,$$

and in general

$$4(p_0 - p_{k+1})(p_0 - p_k) = (2p_0 - p_{k+1} - p_k - p_{k+1}p_k/p_0)^2.$$

From the latter equation, the values of p_{k+1} can be determined in terms of p_0 and p_k.

If we take $p_0 = 1$, then by means of the equation we find that $p_1 = 1/2$, $p_2 = 1/3^2$, $p_3 = 1/2 \cdot 5^2$, $p_4 = 1/17^2$, $p_5 = 1/2 \cdot 29^2$, $p^6 = 1/99^2$, $p_7 = 1/2 \cdot 169^2$, $p_8 = 1/577^2$, etc. The ratio of successive radii approaches $3 - 2\sqrt{2}$ as a limit.

A simpler recursion formula is $(1/p_{k+1})(1/p_{k-1}) = (1 + 1/p_k)^2$.

Note that the circle of center F and radius $2p_0$ is also tangent to all the circles w_k.

Also solved by Phil Tracy, Liverpool, New York; and the proposer.

A Diophantine Solution

844. [September, 1972] *Proposed by Gregory Wulczyn, Bucknell University, Pennsylvania.*

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

Solution by G. E. Bergum and K. L. Yocom (jointly), South Dakota State University.

The Diophantine equation to be solved is

$$(x - 1)^2 + 2x^2 + 3(x + 1)^2 = y^2,$$

or

$$6x^2 + 4x + 4 = y^2.$$

Now y is even and x is even so set $x = 2u$, $y = 2v$ getting

$$6u^2 + 2u + 1 = v^2.$$

Solving for u,

$$u = \frac{-1 \pm \sqrt{6v^2 - 5}}{6}.$$

Setting $6v^2 - 5 = r^2$ we have the Pell equation

$$r^2 - 6v^2 = -5,$$

which has infinitely many solutions in integers (see Nagel's *Number Theory*, p. 204). The fundamental solution is $(r, v) = (1, 1)$. Now for $u = (-1 \pm r)/6$ to be an integer we must have $r \equiv \pm 1 \pmod 6$. But from the Pell equation $r^2 \equiv -5 \equiv 1 \pmod 6$ and hence $r \equiv \pm 1 \pmod 6$. Thus the original problem has infinitely many solutions.

Also solved by Joseph Arkin, Spring Valley, New York; M. T. Bird, California State University at San Jose; Michael Goldberg, Washington, D. C.; M. J. Knight, California Institute of Technology; John O'Neill, University of Detroit; Paul Volpe, B. N. Cardozo High School, Bayside, New York; Kenneth M. Wilke, Topeka, Kansas; and the proposer.

Comment on Problem 791

791. [March, 1971, and January, 1972] *Proposed by D. Rameswar Rao, Secunderabad, India.*

Show that the only solution in positive integers of $x^3 + y^3 + z^3 = u^3$ with x, y, z, u in arithmetic progression is $x = 3$, $y = 4$, $z = 5$, and $u = 6$.

Comment by S. O. Schachter, Philadelphia, Pennsylvania.

Solution by David A. Rozen, Buffalo, New York, ends by saying, "That is $(3d)^3 + (4d)^3 + (5d)^3 = (6d)^3 \cdots$. In fact, "$d$" may be any positive integer (e.g., for "d" $= 2$, $6^3 + 8^3 + 10^3 = 12^3$)."

Allow me to point out that "d" may be any constant "k" — rational, irrational, plus or minus, greater than, less than, or equal to zero.

Comment on Q536

Q536. [January, 1972, and September, 1972] Show that the square roots of three distinct prime numbers cannot be the terms of a common geometric progression.

[Submitted by Murray S. Klamkin]

Comment by the proposer.

In his comment on Q536 (September, 1972), Wernick does not solve the given problem since he assumes that the three terms are consecutive terms of a geometric progression.

Comment on Q543

Q543. [May, 1972] Show that for all natural numbers $n \geqq 4$, $(n-1)^n > n^{n-1}$.

[*Submitted by Alexander Zujus*]

Comment by Frederick Ficken, New York University.

Replace $n-1$ by k and consider

$$a_k = k^{k+1} - (k+1)^k = k^k[k - (1 + 1/k)^k].$$

We know that $(1 + 1/k)^k$ increases with k, converging to $e < 3$. Hence $a_k > 0$ for $k \geqq 3$. Obviously $a_k \to \infty$.

A similar comment was submitted by Sidney Spital, California State University, Hayward.

Comment on Q549

Q549. [September, 1972] Find a sequence $\{a_n\}$ of real numbers such that $\{a_n\}$ converges to zero and $\{na_n\}$ converges to a transcendental number.

[*Submitted by Michael Jones*]

I. *Comment by Peter A. Lindstrom, Genesee Community College.*

[1]. If $a_n = \dfrac{\left(1 + \dfrac{1}{n}\right)^n}{n}$, where $0 < a_n < \dfrac{3}{n}$ for $n \geqq 1$, then $\{a_n\} \to 0$ and $\{na_n\} \to 1$.

[2]. If $a_n = \dfrac{\sum_{i=0}^{n} \dfrac{1}{i!}}{n}$, where $0 < a_n < \dfrac{3}{n}$ for $n \geqq 1$, then $\{a_n\} \to 0$ and $\{na_n\} \to 1$.

II. *Comment by Richard Gibbs, Fort Lewis College, Colorado.*

A much simpler sequence would be

$$\{a_n\} = \left\{\frac{\pi}{n}\right\}.$$

Similar comments were submitted by Joseph B. Browne, Oklahoma State University, and Shen Lin, Bell Laboratories, Murray Hill, New Jersey.

Comment on Q551

Q551. [September, 1972] Prove that the diophantine equation $5^x + 2 = 17^y$ has no solutions.

[*Submitted by Erwin Just*]

Comment by Shen Lin, Bell Laboratories, Murray Hill, New Jersey.

Read $5^x + 2 = 17^y \pmod 4$ produces $1 + 2 \equiv 1 \pmod 4$ which is a contradiction.

ANSWERS

A567. Since $(6^{16k+2}/2) - 1 = 6^2/2(6^{16k} - 1) + (6^2/2 - 1) = 18[(6^k)^{16} - 1] + 17$ and Fermat's theorem guarantees that $(6^k)^{16} - 1$ is divisible by 17, it follows that $(6^{16k+2}/2) - 1$ is divisible by 17.

A568. Summing the left hand side, we get $\dfrac{x^{2n+2} - 1}{x^2 - 1} = x^n$ or equivalently $(x^n - 1)(x^{n+2} + 1) = 0$. Thus $x = e^{i\theta}$ where

$\theta = 2\pi m/n$, $m = 1, 2, \cdots n - 1$ (excluding $\theta = \pi$ if m is even)

$\theta = \pi(2m + 1)/n$, $m = 1, 2, \cdots n + 1$ (excluding $\theta = \pi$ if n is odd).

More generally one can treat

$$\sum_{i=0}^{n} x^{ri} = x^{rn/2} \text{ in a similar way.}$$

A569. The statement is false. Let $p = 2$ and $a_n = (2^n \ln n)^{-1}$. Then $\lim_{n\to\infty} n^k a_n = 0$ but for every $\varepsilon > 0$, $\lim_{n\to\infty} n^{p+\varepsilon} a_n = \infty$.

A570. The solution follows by observing that $f_X(x)f_Y(y)$ is zero except in the smallest rectangle R containing A; therefore, there is a nonzero area $R - A$ where $f_X(x)f_Y(y) > f_{XZ}(x, y) = 0$.

A571. The space X is a Hausdorff space iff the diagonal is a closed subset of $X \times X$. That is, iff whenever the net (x_α, x_α) converges to (a, b) then (a, b) belongs to the diagonal. In other words, iff $x_\alpha \to a$ and $x_\alpha \to b$ implies $a = b$.

A572. The equation may be rewritten as

$$z = [(x/z)^n + (y/z)^n]^k.$$

If a and b are integers then setting $x/z = a$ and $y/z = b$

$$z = [a^n + b^n]^k$$

and hence

$$x = a[a^n + b^n]^k, \; y = b[a^n + b^n]^k.$$

(Quickies on pages 167–168)