Mathematical Spectrum

Volume 8 1975/76 *Number 2*

A Magazine of *Contemporary Mathematics*

Published by the *Applied Probability Trust*

Mathematical Spectrum is a magazine for the instruction and entertainment of student mathematicians in schools, colleges and universities, as well as the general reader interested in mathematics. It is published by the Applied Probability Trust, a non-profit making organisation established in 1963 with the support of the London Mathematical Society. The object of the Trust is the encouragement of study and research in the mathematical sciences.

Volume 8 of *Mathematical Spectrum* will consist of three issues, of which this is the second. The first issue was published in September 1975, and the third will appear in May 1976.

Articles published in *Mathematical Spectrum* deal with the entire range of mathematical disciplines (pure mathematics, applied mathematics, statistics, operational research, computing science, numerical analysis, biomathematics). Both expository and historical material may be included, as well as elementary research and information on educational opportunities and careers in mathematics. There is also a section devoted to problems.

Articles are normally commissioned by the Editors; the Editorial Committee also welcomes the submission of suitable material, including correspondence, queries and solutions to problems, for publication in *Mathematical Spectrum*. All correspondence about the contents should be sent to:

The Editor, *Mathematical Spectrum*,
Hicks Building, The University, Sheffield S3 7RH.

International Mathematical Olympiad, 1975

J. H. DURRAN
Winchester College

Shake the kaleidoscope. Visas and traveller's cheques; the anxious meeting at Cromwell Road (so these are our companions for the next twelve days); the Danube from 29,000 ft; Bulgarians trotting into Sofia airport lounge by the OUT door; the pound–stokinki exchange rate (look again! *stokinki*?) 2·72 to the pound, Euler's e; the heat; the friendly interpreter-companion; the traumatic isolation of team from jury at Burgus; the free pocket-money issue.

The final wording of the lucky six questions, chosen from dozens; prams and citizens promenading the Burgus boulevards in the cool of the evening; the open churches and a wedding; eating in a PECTOPAHT (transliteration from Cyrillic left as an exercise for the reader); never a dog.

The Opening Ceremony; flags of the nations; the heat; delay from lights failure; the heat; the start of the first four-hour session; tension in the jury room as questions on the questions are brought up to be answered or shrugged off; the heat; ice-cream, Coca-Cola, coffee rounds; only an hour to go.

See blue Black Sea (black Blue Danube we never see close); the inscrutable Mongolians, the smiling Vietnamese (they have something to smile about); sunflowers by the square mile (sorry: hectare); beaches, waves, sand, Golden Sands, Sunny Beach, tower hotels. The second session.

Marking, coordinating, discussing prizes; the jury's turn on the rack; three days of that; but beaches again for the teams; and again the beaches.

The two-day cavalcade to Sofia; the competitors' buses, police-escorted, holding the middle of the road; headlights glaring at noon; into the verge with oncoming traffic. The Balkan range and age-old passes, the invader's way; yoghourt; the Turks and 1875; the welcomes in the towns; bouquets of carnations, roses; children with gifts; platforms, speeches (marguerites in the smaller towns); attar of roses, Thracian tomb. Turnovo, the Ancient Capital, with cyclopean gatehouse and Baldwin's Tower; (breathless? weren't we all?); fruit juices *versus* Coca-Cola; Roman inscriptions (a far cry from Hadrian's Wall, the bounds of Empire, but Babel now); the Russian learner speaking his Russian to real live Russians with dust on their shoes.

Sofia; The Holy Wisdom and many-domed, golden-domed Alexander Nevsky; 1875 and the Turks; pink and peach stucco; Party HQ; cool mosque, the last of its tribe; Lenin's frown; ancient churches; frescoes; bookshops; a dog (on a lead). The Closing Ceremony (already?); the prizes, hand-tooled leather folders; photographs; hands clasped; the partings of friends, photographs; the final dinner, 200 seated in a mountain-top restaurant; a double-bill farewell speech with American speaking Russian and Russian English (the night of space-craft rendezvous); dancing, more speeches; more wine (beer for the boys).

Sofia airport again; Vienna, only the airport (too far to the city, but a pastry, surely, and a coffee? At over £1?); IMO's to be here next year, perhaps they'll be free then? Heathrow; the bonds slip away; the crowds suck us in; 'See you in October', 'Come and visit me', 'We'll look at my slides'.

What was the mathematics like? The what? Oh, the mathematics; there were puzzles, tricks, in-jokes, chess; the official part was like this:

1. Let x_i, y_i ($i = 1$ to n) be real numbers such that

$$x_1 \geqslant x_2 \geqslant \cdots \geqslant x_n \quad \text{and} \quad y_1 \geqslant y_2 \geqslant \cdots \geqslant y_n.$$

Prove that, if $z_1, z_2, \ldots, z_n$ is any permutation of $y_1, y_2, \ldots, y_n$, then

$$\sum_{i=1}^{n} (x_i - y_i)^2 \leqslant \sum_{i=1}^{n} (x_i - z_i)^2. \qquad \text{(6 pts)}$$

2. Let $a_1, a_2, \ldots, a_r \ldots$ be any infinite sequence of strictly positive integers such that $a_r < a_{r+1}$ for $1 \leqslant r$. Prove that infinitely many a_n can be written in the form

$$a_n = xa_i + ya_j$$

with x, y strictly positive, and $i \neq j$. (7 pts)

3. On the sides of an arbitrary triangle ABC, triangles ABR, BCP, CAQ are constructed externally with

$$\angle PBC = \angle CAQ = 45°,$$

$$\angle BCP = \angle QCA = 30°,$$

$$\angle ABR = \angle RAB = 15°.$$

Prove that $\angle QRP = 90°$ and $QR = RP$. (7 pts)

4. When 4444^{4444} is written in decimal notation, the sum of its digits is A. Let B be the sum of the digits of A. Find the sum of the digits of B. (A, B are written in decimal notation.) (6 pts)

5. Determine, with proof, whether one can find 1975 points on the circumference of a circle of unit radius such that the distance (along the chord) between any two of them is a rational number. (6 pts)

6. Find all polynomials P in two variables with the following properties:

(i) for some positive integer n and all real t, x, y,

$$P(tx, ty) = t^n P(x, y)$$

(that is P is homogeneous of degree n), and

(ii) for all real a, b, c

$$P(a + b, c) + P(b + c, a) + P(c + a, b) = 0,$$

and

(iii) $P(1, 0) = 1$. (8 pts)

Countries of origin:
1. Czechoslovakia, 2. Great Britain, 3. Netherlands, 4. USSR, 5. USSR, 6. Great Britain.

Our scores (out of 40) were 40, 40, 36, 32, 25, 24, 23, 19; we got 2 first prizes (only 8 competitors scored 40), 2 second prizes, 3 third prizes. As a team (team? the Olympic spirit?) we were 5th, 19 points behind the Hungarian winners, and the 7th team trailed us by 47 points.

Enter the NMC*, O reader; get chosen for the BMO†; try for the XVIIIth IMO in Austria, 1976.

* National Mathematical Contest.
† British Mathematical Olympiad.

Solving Polyomino Covering Problems by Computer

R. J. STAMMERS* and R. N. MADDISON
The Open University

1. Introduction

In 1953 S. W. Golomb, an American mathematician, defined the polyomino as 'any flat figure formed by joining unit squares along their edges'. He posed the following kind of problem, which is like a jigsaw puzzle with straight-edged square-cornered pieces. Given a number of small polyomino pieces and a large polyomino shape, whose area is equal to the sum of the areas of the pieces, find all the ways that the pieces can be placed so as to cover the shape exactly. We, the authors, are interested in devising methods of solving such problems by using a digital computer.

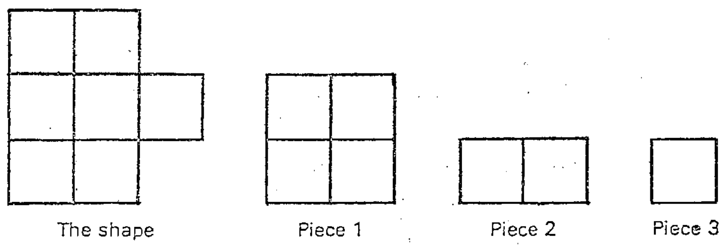


Figure 1. A simple example in which the problem is to find all ways of placing the pieces to cover the shape.

* Now at the University of Essex.

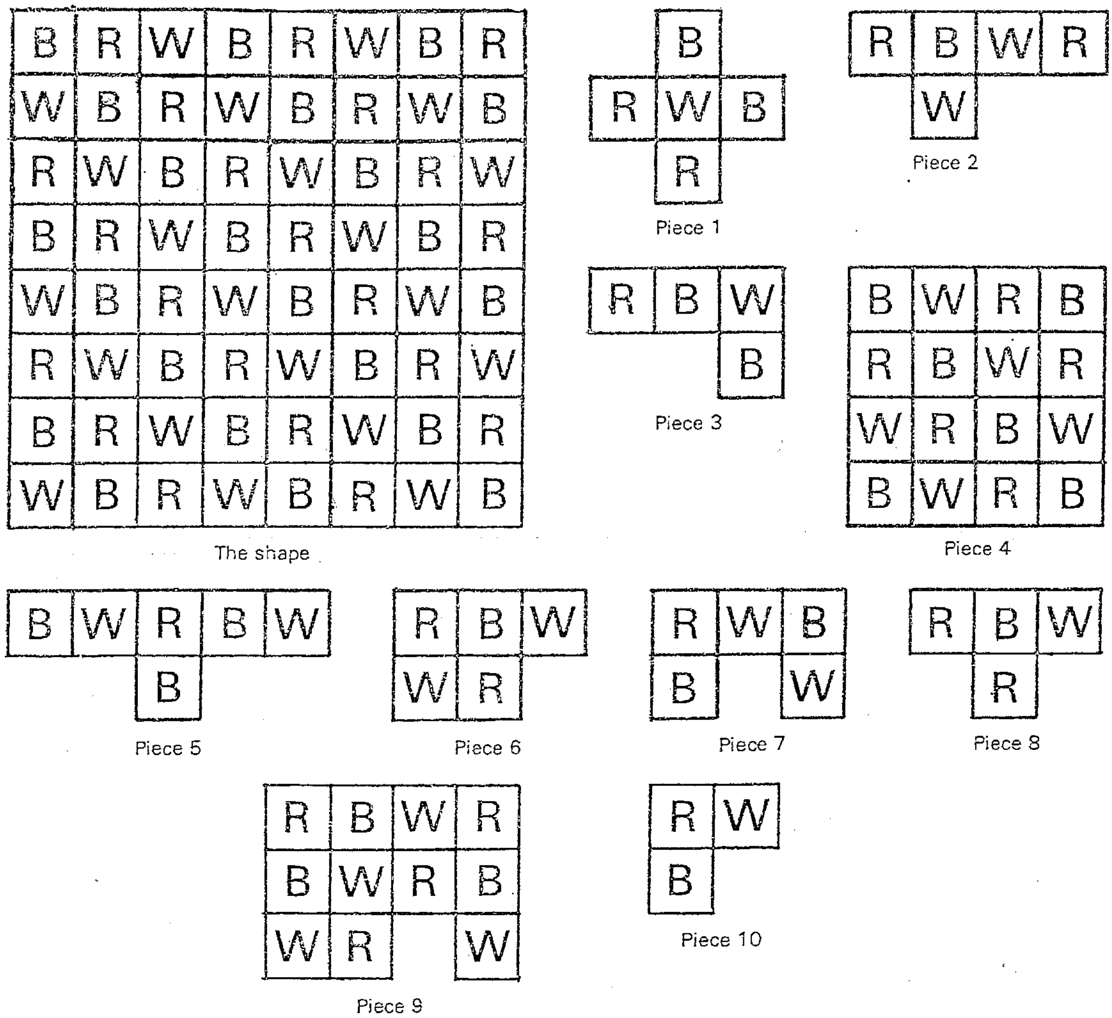


Figure 2. A problem actually solved in six minutes on an ICL 1903T computer. Each unit square is Red, White or Blue and the pieces are coloured both sides and may be rotated and/or turned over. There are eight solutions, related by the diagonal symmetry of the shapes of pieces 1 and 4.

A very simple instance of the problem is shown in Figure 1. This example is used to illustrate our method of solution or working throughout this article and is necessarily extremely simple. A more typical sort of problem (which we have succeeded in solving using our computer method) is shown in Figure 2. The reader may like to construct some problems like these from paper or cardboard and try them.

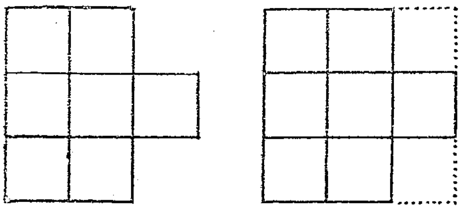
Despite the seeming simplicity of this kind of problem, it has defied almost all attempts to find a general method of solution. To be successful, the method of solution must work for every possible problem of the kind described, must be programmable to be done automatically by computer, and must—within a finite amount of computer time—either yield the solution(s) or discover that there are none.

Methods for solving these problems basically involve searching 'combination trees' for solutions. Each branch represents a move or operation or possible placing of a piece. After any particular move there may be a choice of many possible moves or operations or positions for placing the next piece, corresponding to many sub-branches leading from a particular branch. Unfortunately, the nature of the problems is such that these trees grow very rapidly as one searches—a combination or search explosion effect.

Our aim was to try to find a method that could be programmed and that reduced the amount of searching to such an extent that the solution of sizeable problems—such as that of Figure 2—is practicable and feasible. We also found that the methods are relevant to a much more general class of problems.

2. Character representation

The first stage is to translate the problem into a representation suitable for computers. Figure 3 shows how a polyomino can be represented as a character matrix and hence as a character string. The polyomino shape is 'squared-up' from the top left corner to size $n \times n$ say, after which it can be described as an $n \times n$ matrix by using 1's to denote an element present and 0's to denote an element not being present. Coloured polyominoes can be described by using other characters, such as letters. The polyomino pieces are similarly represented by means of matrices of the same size as the matrix used to represent the shape.



$$\begin{matrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{matrix} \qquad 110111110$$

Character matrix — Character string

Figure 3. The shape, or a piece, can be represented by a string of characters corresponding to the elements of an enclosing square.

Having devised a method for representing the polyominoes in a computer we wish to devise algorithms corresponding to the 'real-life' operations that can be performed with polyomino pieces.

Reflecting or rotating the matrix describing a polyomino piece yields another matrix which describes the reflected/rotated polyomino. Figure 4 shows a horizontal reflection and a clockwise rotation through 90°. Using these two operations we can obtain all eight possible orientations of the polyomino. To conform to our standard representation the matrices are 'top left justified' after reflection and rotation.

The character matrices are in fact held as character strings inside the computer. Duplicate versions of pieces can thus be identified by testing for identical character strings.

3. Testing for fits of a piece on the shape

It is necessary to develop an algorithm to test whether a piece will 'fit' a shape in any particular position. An attempt to fit a piece to a shape in such a way that

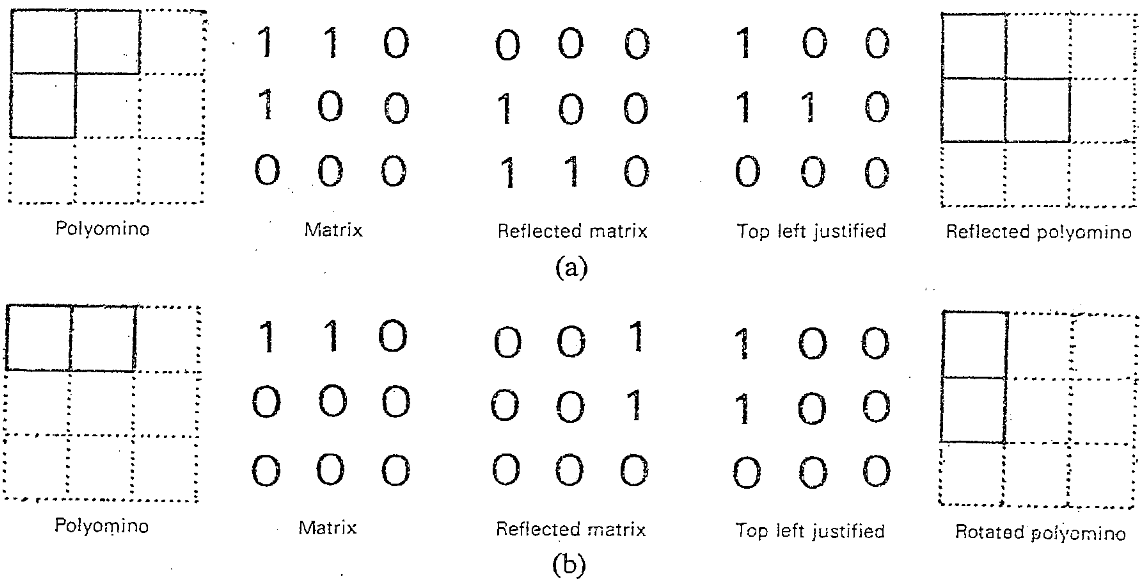


Figure 4. Forming the representations corresponding to (a) turning over about a horizontal line, and (b) rotating 90° clockwise.

element (1,1) of the piece covers element (k, l) of the shape is described as an 'attempt to fit the piece at (k, l)'. The rule for testing whether a piece will fit at (1,1) is as follows. 'Every non-zero character in the string describing the piece must be identical with the corresponding character in the string describing the shape'. The character string describing the residual shape can be deduced from a similar sort of rule. Suppose the shape and the piece were represented by $n \times n$ matrices. Then to test for a fit at (k, l) we first shift the string representing the piece $(k - 1)n + (l - 1)$ positions right and apply the same rules. Examples of pieces that fit and do not fit are given in Figures 5 and 6.

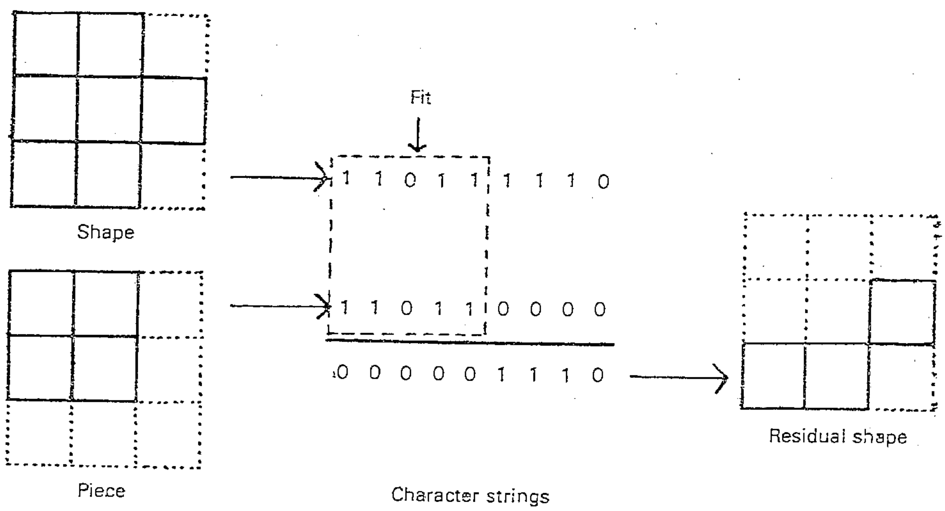


Figure 5. Discovering that piece 1 fits at the top left of the shape, and obtaining the representation of the residual shape that would have to be fitted by the remaining pieces if that first possible fit of piece 1 is done.

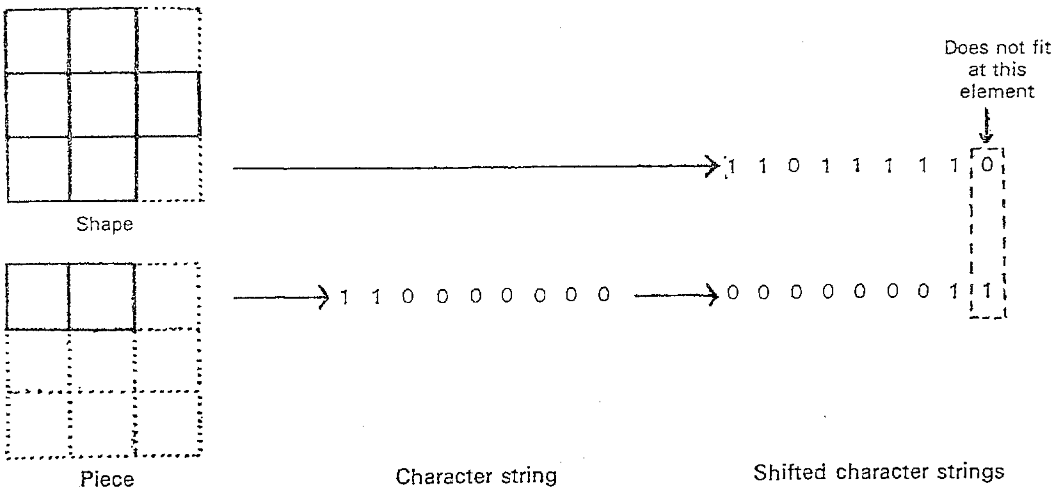


Figure 6. Discovering that piece 2 does not fit with its top left corner in row 3, column 2.

4. Abstracting the problem still further

The algorithms to 'jiggle around' with polyominoes can be used to abstract the problem further. After the input of descriptions of the shape and the pieces a list is produced of all the distinct versions (i.e. reflections and rotations) of all the pieces. From these a further list may be made of all the ways in which any piece can be fitted to the shape. In principle this may be produced by trying all possible distinct versions of the pieces at all possible positions on the shape. In this list of possible fits, entries 1 to 20 might be the 20 possible ways of placing piece 1 on the shape; entries 21 to 35 the 15 possible ways of placing piece 2 (assuming piece 1 unplaced); and so forth.

Finally a remaining fits matrix, or refits matrix, may be constructed. Its purpose is to show all the fits that remain still possible after any one given fit has been performed. Element (r, s) of the refits matrix shows whether fits s can be performed after fit r has been performed: the entries can be coded '1' for yes, '0' meaning no. For this type of polyomino problem, the refits matrix is symmetric. Figure 7 shows how this information is built up for the example of Figure 1.

We wrote 'may be produced' above because neither the complete list of fits nor the complete refits matrix need necessarily be calculated at the outset. We may defer calculating parts of the list of fits or the refits matrix until we decide those parts are needed. For the purpose of the next section, assume that everything has been calculated.

Finding a set of fits for the pieces which exactly cover the shape then reduces to finding a chain of the right length through the refits matrix. Taking the example shown in Figure 7, select any of the fits 1–17 to start the process. Suppose that fit 1 was chosen. From the refits matrix, fits 6, 15, 16 and 17 would then remain. If fit 6 was chosen, the fits that would remain are the remaining fits after *both* fit 1 (6, 15, 16 and 17) *and* fit 6 (1, 11, 12, 13, 14 and 15). In this case the only remaining fit that is compatible with both of these is fit 15. This yields a chain of length 3: fit 1, fit 6, fit 15. When a chain is found with as many items as there are pieces, then a solution has been found. Hence fits (1, 6, 15) are a solution to the example.

				Remaining fits matrix																
Piece	Version number	Row and column coordinates of top left	Fit number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1	1 1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1
	1	2 1	2	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0
2	2	1 1	3	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
	2	2 1	4	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1
	2	2 2	5	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1
	2	3 1	6	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0
	3	1 1	7	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1
	3	1 2	8	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1
	3	2 1	9	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	1
	3	2 2	10	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0
3	4	1 1	11	0	1	0	1	1	1	0	1	1	1	0	0	0	0	0	0	0
	4	1 2	12	0	1	0	1	1	1	1	0	1	1	0	0	0	0	0	0	0
	4	2 1	13	0	0	1	0	1	1	0	1	0	1	0	0	0	0	0	0	0
	4	2 2	14	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0
	4	2 3	15	1	1	1	1	0	1	1	1	1	1	0	0	0	0	0	0	0
	4	3 1	16	1	0	1	1	1	0	1	1	0	1	0	0	0	0	0	0	0
	4	3 2	17	1	0	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0

Figure 7. The remaining fits matrix. After doing fit number 1, fit numbers 6, 15, 16 and 17 remain still possible.

5. Generalising the polyomino covering problem

Polyomino covering problems are members of a much wider class of combinational problems, characterised as follows.

(a) Some given objective has to be achieved. (In this case the shape has to be covered with the pieces.)

(b) The objective comprises a set of sub-objectives and achieving the objective means achieving all the sub-objectives. (In this case the sub-objectives are the fitting of each of the pieces and the covering of each of the elements of the shape.)

(c) A repertoire of allowable operations is available with which to achieve the objective. (In this case an allowable operation is the positioning of a piece on the shape in a way that is allowable when no other pieces have been placed anywhere.)

(d) Finding a solution means finding a sequence of allowable operations which together achieve all the sub-objectives.

(e) At each stage of searching for or selecting the next operation in a sequence, the operations that can be selected are dependent on the operations chosen earlier in the sequence. (In the case of the polyominoes each time we fit a piece onto the shape we restrict the ways in which pieces can subsequently be placed on the shape.)

It may already be apparent that one advantage conferred by the remaining fits matrix is that it enables us to perform all the 'polyomino juggling' once only at the outset, and thereafter it is possible to search for solutions without repeating this work on each branch down the combination tree. What is perhaps more important, however, is that, once the remaining fits matrix has been produced, *the task of finding solutions has been reduced to a form which is independent of polyomino covering problems, but common to the general class of problems.*

6. Searching for solutions

Despite our rearrangement of the problem, the task of finding solutions still involves searching large combination trees. For problems of any complexity, the trees are so enormous that it is neither practicable nor feasible to search the entire tree exhaustively. In examples like Figure 2, there may be about 250 possible placings or fits for piece 1, for each of these there may be over 200 compatible fits of piece 2, for each of these 250 $\times$ 200 combinations about 200 fits of piece 3 and so forth. Our tree has already got 10,000,000 branches and we are only down to level 3. Try to imagine the size at level 12! The big question is, how can a computer be programmed to reduce the amount of searching to such an extent that the solution of sizeable problems is practicable and feasible? Some possible approaches are outlined below.

Suppose that it is possible to establish some necessary conditions that the shape and the pieces must satisfy in order for solutions to exist. These conditions can be used to reduce the area of the tree that is actually searched. The entries in the remaining fits matrix can be tested, and useless items pruned out. In addition it is possible to apply the conditions repeatedly whilst working through the combination tree, thus enabling profitless paths to be identified more quickly.

These conditions operate in a 'trade-off' situation—they are only useful if the search time saved is greater than the time spent testing the conditions. We find, when looking for solutions to different problems, that conditions which are excellent for one particular problem are useless for another. Watching humans solve these problems suggested two simple conditions that have been found to work well for all the examples tried.

At any stage during the search,

1. there must remain at least one fit for each of the unfitted pieces;

and

2. there must remain at least one fit to cover each of the uncovered elements of the shape.

In other words, for each remaining sub-objective there must remain at least one operation which achieves that sub-objective. The structure of the remaining fits matrix is such that these conditions can be tested very briskly by the use of binary logic operations.

The way in which the searching is organised, or scheduled, is very important.

Systematically examining all the alternatives in a pre-determined sequence does not work well. Here again, watching humans solve these problems has suggested a method for organising the search which has been found to work well. If $L - 1$ operations have previously been performed, then we shall say that we are level at L in the search. The method is now as follows:

At any level L in the search, identify the most critical sub-objective. This is the sub-objective for whose achievement the smallest number of options is available at level L.

Select this as the sub-objective on which to branch; in other words at level L consider in turn each of the operations which achieve the most critical sub-objective.

For each operation considered at level L, establish the situation resulting from using this operation. Then repeat this procedure at level $L + 1$, and hence at levels $L + 2$, $L + 3$, etc., either until a chain of operations of the required length is discovered, which gives a solution, or until it is impossible to continue, in which case that branch or situation is abandoned and the next alternative is considered.

A procedure such as this, which recalls itself, is known as a recursive procedure. Let us give an example.

Suppose that in the process of solving a larger problem we have reached level L and the situation is as shown in Figures 1 and 7. The remaining fits are fits 1, 2, 3, . . ., 17. By means of the refits matrix the table shown in Figure 8 is constructed.

Sub-objective	List of remaining fits that accomplish the sub-objective	Total number in list
1. Cover (1,1)	1, 3, 7, 11	4
2. Cover (1,2)	1, 3, 8, 12	4
3. Cover (2,1)	1, 2, 4, 7, 9, 13	6
4. Cover (2,2)	1, 2, 4, 5, 8, 10, 14	7
5. Cover (2,3)	5, 15	2*
6. Cover (3,1)	2, 6, 9, 16	4
7. Cover (3,2)	2, 6, 10, 17	4
8. Fit piece 1	1, 2	2*
9. Fit piece 2	3, 4, 5, 6, 7, 8, 9, 10	8
10. Fit piece 3	11, 12, 13, 14, 15, 16, 17	7

* Most critical because of lowest frequency.

Figure 8. Initial selection of the most critical sub-objective from the list giving the possible fits that achieve each sub-objective.

From this table, the two most critical sub-objectives are 'fit piece 1' and 'cover element (2,3)', both of which are achieved by only two of the remaining operations. So we make an arbitrary choice and branch on the second of these, 'cover element (2,3)'. This means that at level L we are going to examine operations 5 and 15 in turn. Consider first operation 5, placing piece 2 at (2,2). After this operation, the remaining possible operations are 11, 12, 13, 16, 17. The remaining sub-objectives are 1, 2, 3, 6, 7, 8, 10. We can now identify that this situation is not going to yield any solutions because none of the remaining operations achieves the remaining sub-objective 8—in other words, after placing piece 2 at (2,2) there is then no way of placing piece 1. Hence we reject operation 5 at level L. The next operation to be tried at level L is operation 15—placing piece 3 at (2,3). After this operation the remaining possible operations are 1, 2, 3, 4, 6, 7, 8, 9, 10. The remaining sub-objectives are 1, 2, 3, 4, 6, 7, 8, 9. There does exist at least one remaining operation for each remaining sub-objective, and so we go on to level $L + 1$.

At level $L + 1$ we need to re-establish the sub-objective information for the new situation, in order to determine the most critical sub-objective. This information is shown in Figure 9.

Sub-objective	List of remaining fits that accomplish the sub-objective	Total number in list
1. Cover (1,1)	1, 3, 7	3
2. Cover (1,2)	1, 3, 8	3
3. Cover (2,1)	1, 2, 4, 7, 9	5
4. Cover (2,2)	1, 2, 4, 8, 10	5
5. Done		
6. Cover (3,1)	2, 6, 9	3
7. Cover (3,2)	2, 6, 10	3
8. Fit piece 1	1, 2	2*
9. Fit piece 2	3, 4, 6, 7, 8, 9, 10	7
10. Done		

* Most critical because of lowest frequency.

Figure 9. At a typical later stage, after trying fit 15, i.e. piece 3 at row 2, column 3 of the shape. The combinational or search explosion effect will be smallest if we try the two fits that achieve sub-objective 8 rather than the 3, 5 or 7 possible branches for some other sub-objective.

From this table we see that the most critical sub-objective is 'fit piece 1', which is achieved by only two of the remaining operations. Hence we branch on this sub-objective, which means that at level $L + 1$ we are going to examine operations 1 and 2 in turn. The entire search tree for this example is shown in Figure 10. The important point is the extent to which the method minimises the number of trial operations or branches to be considered.

Our present computer program, based on the method outlined above, found all 65 solutions to the problem shown in Figure 11 in 3 hours 12 minutes, using 58K

24-bit words of store of an ICL 1903T computer. Does this help us answer our big question, how can a computer be programmed to reduce the amount of searching to such an extent that the solution of sizeable problems is practicable and feasible? In the course of the search in Figure 10, six trial placings of a piece were made; of these, five were satisfactory in the sense that they actually later gave rise to

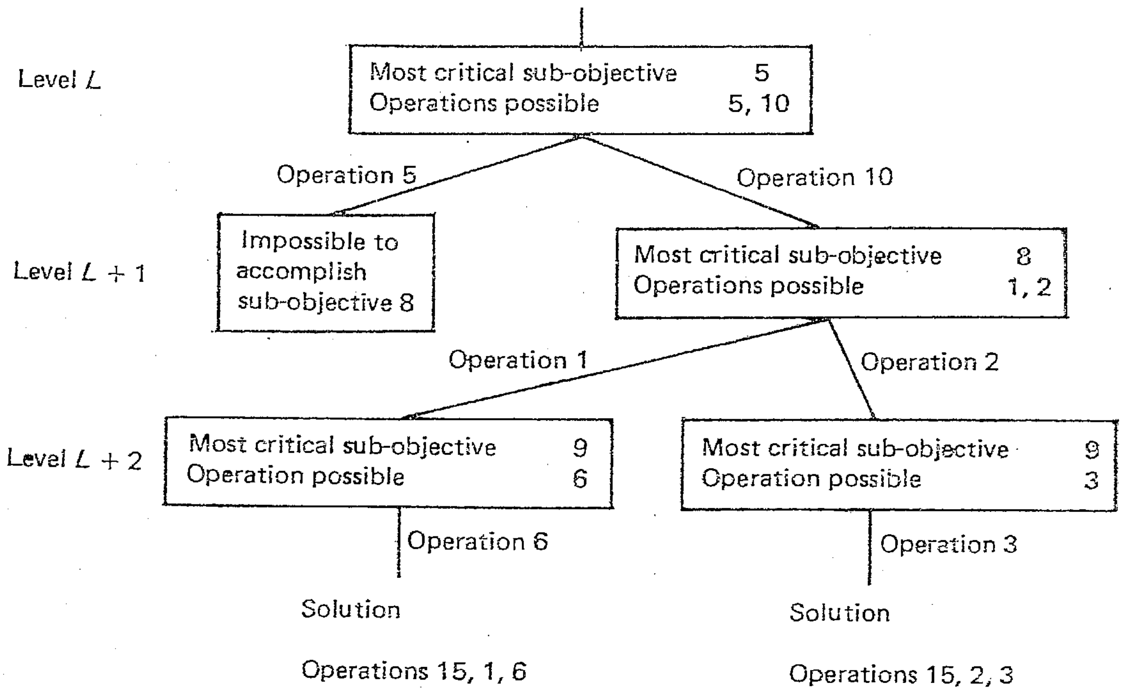


Figure 10. The tree of all the branches considered in solving Figure 1 by the method described.

successful solutions. The success ratio 5/6, or in general of total number of successful operations used to total operations tried, is a measure of the performance of the procedure. This ratio varies from one example to another. In general the success ratio decreases as examples increase in complexity. A measure of the complexity of an example is as follows.

Suppose that there are n pieces and the number of operations possible with piece i is p_i. Then the product $p_1 p_2 p_3 \ldots p_n$ is a measure of the complexity of the example.

For the example shown in Figure 2, this complexity is about 10^{30}, and the computer method had a success ratio of about 0·1 (300 successful operations out of 3,000 trials). Although it is very dependent on the speed of the actual computer used, the average amount of computer time per operation used is also important in evaluating the procedure. The average computer time per operation used also increases as the complexity of the problem increases. For this example the average time on the 1903T was 4 seconds per operation.

If one considers that the example in Figure 11 is a sizeable problem, then we can claim to have found a feasible method. Most people would take several days to find one solution to this problem, and several weeks to find all of them. So it is fair to say that, in human terms, the example does constitute a sizeable problem, and that this method is quicker than humans solving it. Our 1903T program could

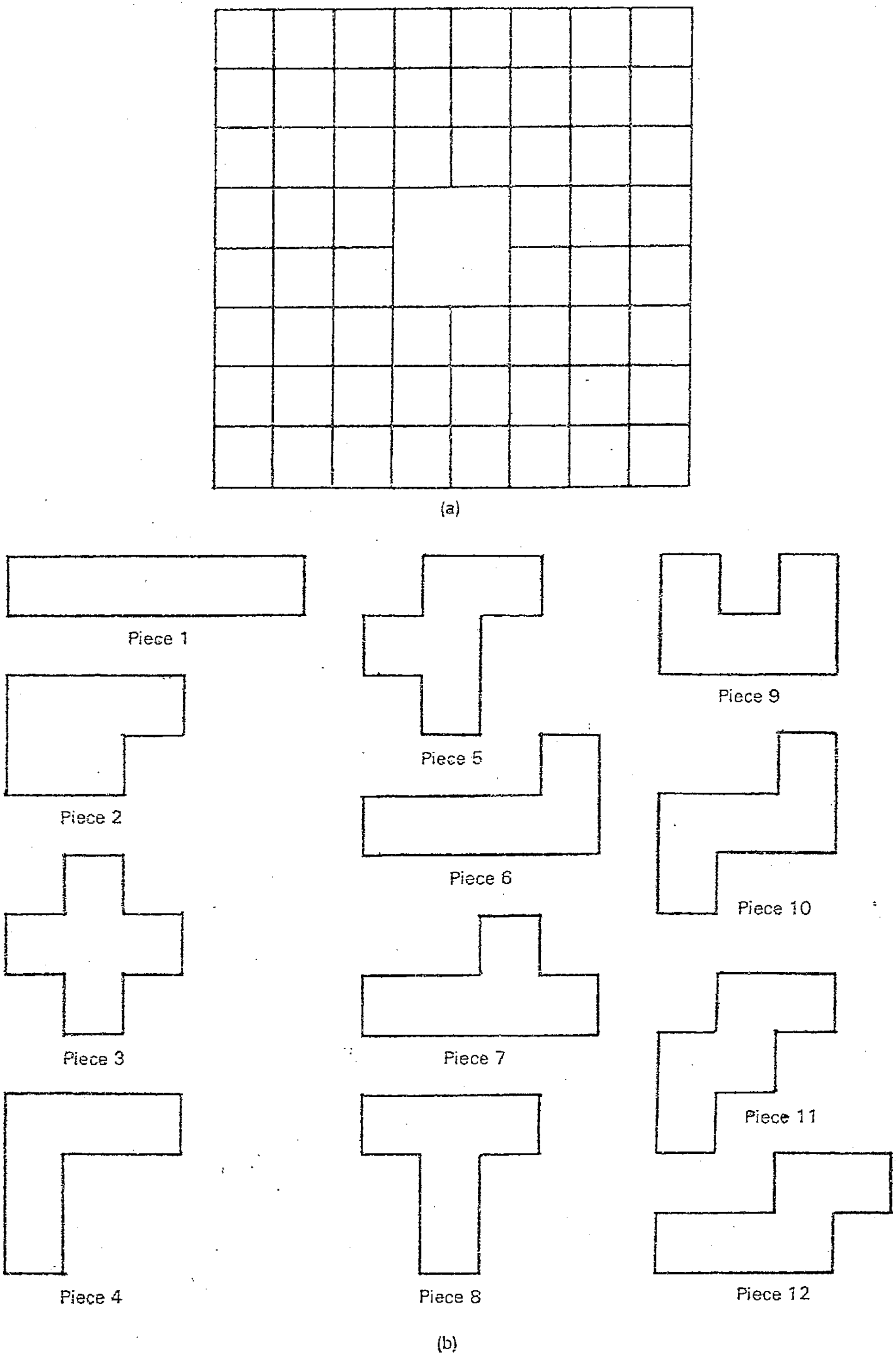


Figure 11. A well-known polyomino problem, using as pieces all the arrangements of five unit squares. (a) The shape. (b) The pieces, which may be turned round or over.

solve within a reasonable amount of computer time any polyomino covering problem that a human might reasonably expect to solve, up to say 20 pieces on a shape of 150 to 200 unit squares. However, some of the real-life problems, such as resource allocation, or school time-tabling, that are members of the general class of combinational problems, are beyond the scope of the present general procedure.

Acknowledgment

We are grateful to the Open University for computing facilities.

Reference

S. W. Golomb, *American Mathematical Monthly* **61** (1954), 675–682.

Time to Win, Time to Lose

FREDERICK STERN
San Jose State University

Consider a classical gambler of probability theory. He begins playing with an initial stake of z units, gambles one unit at a time with probability $p < 1$ of winning a unit and $q = 1 - p$ of losing one. He stops playing either when he is *ruined*, having lost all of his original stake, or when he *retires*, having doubled his original stake by accumulating $2z$ units. By symmetry, if $p = q = \frac{1}{2}$, the gambler has a probability one-half either of being ruined or of doubling his initial stake. In contrast, for p less than q, even if the game is only slightly against the player, and even if the initial stake is quite moderate, the probability of ruin is considerably higher than one-half.

The exact formula, which was known during the time of James Bernoulli (see reference 1), for the probability of ruin is $Q(z, p) = q^z/(p^z + q^z)$. For example, if $p = 0{\cdot}45$, $q = 0{\cdot}55$ and $z = 5$, $Q(5, 0{\cdot}45)$ is about $0{\cdot}73$. Interpreting this in terms of relative frequencies, in repeated plays our tireless gambler would lose his five-unit initial stake about 73% of the time and double it only about 27% of the time!

Does this disparity in ruin and retirement probabilities from the symmetric ($p = \frac{1}{2}$) to the non-symmetric ($p \neq q$) games also exist when we consider the length of the game? In symmetric and non-symmetric cases, how does the average number of steps compare in games which end respectively in ruin or in retirement? In symmetric games, as again seems apparent, the average length of games ending in ruin equals the average length of games ending in doubling the stake.

There are, moreover, two extreme cases which, despite the assumption that $p \neq q$, support the hypothesis of the equality of the average length of ruin and of retirement games. First, if the initial stake is one unit, then no matter what value is chosen for p, games ending in ruin end in one step. And so do games ending in doubling the stake. Secondly, suppose p is very small, say 0·01, and $q = 0{\cdot}99$. If the initial stake is five units, then with very high probability $[(0{\cdot}99)^5/((0{\cdot}01)^5 + (0{\cdot}99)^5) = 0{\cdot}999989]$ ruin will occur and, with almost as high probability $[(0{\cdot}99)^5 = 0{\cdot}950990]$ ruin will occur in exactly five steps. On the other hand, there is a very slight probability $[(0{\cdot}01)^5/((0{\cdot}01)^5 + (0{\cdot}99)^5) = 0{\cdot}000011]$ that doubling the stake will occur.

Most of this probability rests in those games with five consecutive wins $[(0{\cdot}01)^5 = 0{\cdot}000010]$. Games ending with doubling of the stake should average in length slightly over five steps which should also be the approximate average length of the games in which ruin occurs.

Let us pose this question more precisely and answer it. Let $u_n(z, p)$ be the probability that a game with an initial stake of z units and a probability p of winning one unit ends in ruin in exactly n steps. Let $v_n(z, p)$ be the corresponding probability that the game ends with doubling that stake in n steps. For the probability of ruin we have

$$Q(z, p) = \sum_{n=z}^{\infty} u_n(z, p)$$

while

$$1 - Q(z, p) = \sum_{n=z}^{\infty} v_n(z, p)$$

is the probability of doubling the stake. Define the conditional expectation of the number of steps given that ruin occurs as

$$E(z, p) = \sum_{k=z}^{\infty} k u_k(z, p)/Q(z, p)$$

and the conditional expectation of the number of steps given that doubling of the stake occurs as

$$F(z, p) = \sum_{k=z}^{\infty} k v_k(z, p)/[1 - Q(z, p)].$$

Our question becomes 'How do $E(z, p)$ and $F(z, p)$ compare for different values of p?' The answer is, simply, that for each value of p and z, $E(z, p)$ and $F(z, p)$ are equal, i.e., the long-run average length of games ending in ruin is the same as that for games which end in doubling the stake.

One way to prove this statement is to consider, for a given initial stake z, a sequence of wins and losses which leads to ruin. Such a sequence could be represented by a list of W and L symbols (for example $(W, L, L, W, L, L, L, L, L)$ if z is five), which always ends with L, has n W's and $(n + z)$ L's where n is any non-negative integer. Such ruin sequences, furthermore, are in one-to-one correspondence with sequences which lead to a doubling of the stake. To $(W, L, L, W,$

L, L, L, L, L) corresponds ($L, W, W, L, W, W, W, W, W$) and, in general, to go from a sequence of one kind to a sequence of the other, one need only change each W to an L and *vice versa*. The corresponding stake-doubling sequences end with W, have n L's and $(n + z)$ W's. The probability of a typical ruin sequence is $p^n q^{n+z} = q^z(pq)^n$ and that for the corresponding stake-doubling sequence is $p^z(pq)^n$. We can write

$$u_{2n+z}(z, p) = \sum_{S_{2n+z}} q^z(pq)^n = q^z(pq)^n|S_{2n+z}|$$

where S_m represents the set of all ruin sequences of length m and $|S_m|$ represents their number. Likewise,

$$v_{2n+z}(z, p) = \sum_{S'_{2n+z}} p^z(pq)^n = p^z(pq)^n|S'_{2n+z}|$$

where S'_m represents the set of all stake-doubling sequences of length m. Since S_{2n+z} and S'_{2n+z} are in one-to-one correspondence, $|S'_{2n+z}| = |S_{2n+z}|$. Substituting these into the defining equations, we have

$$E(z, p) = q^z \sum_{n=0}^{\infty} (2n + z)(pq)^n|S_{2n+z}| \Big/ q^z \sum_{n=0}^{\infty} (pq)^n|S_{2n+z}|$$

and

$$F(z, p) = p^z \sum_{n=0}^{\infty} (2n + z)(pq)^n|S'_{2n+z}| \Big/ p^z \sum_{n=0}^{\infty} (pq)^n|S'_{2n+z}|.$$

Since $|S_m| = |S'_m|$ we find that $E(z, p) = F(z, p)$ as asserted.

You might be curious about numerical values for these expectations. By using difference equations and generating functions (see reference 2) it is possible to find that if $p = \frac{1}{2}$, the conditional expectations are given by z^2 and if $p \neq q$, they are given by

$$z\left[1 - \left(\frac{q}{p}\right)^z\right] \Big/ (p - q)\left[1 + \left(\frac{q}{p}\right)^z\right].$$

You could obtain these results directly from the expressions for $E(z, p)$ or $F(z, p)$ given above, if you could find a way to write a formula for $|S_{2n+z}|$ and sum the series in which these expressions appear.

References

1. See I. Todhunter, *A History of the Mathematical Theory of Probability* (Chelsea Publishing Company, 1949) pp. 61–63. For a summary of this and related results see W. Feller, *An Introduction to Probability Theory* Vol. I (3rd Ed., Wiley, 1968), pp. 344–349.
2. F. Stern, Conditional expectation of the duration in the classical ruin problem, *Mathematics Magazine* 48 (1975), 200–203.

Magic Squares

A. D. MISRA
Gauhati University

1. Introduction

An $n \times n$ matrix whose elements are the integers $1, 2, 3, \ldots, n^2$ is called a *magic square* if the sums of the elements in all rows and in all columns are the same (and are therefore all equal to $\frac{1}{2}n(n^2 + 1)$). Various methods for obtaining such squares appear in references [1], [2] and [4]. Johnson [3] describes a simple construction for magic squares of odd order. An even simpler method for constructing such magic squares is given in Section 3; and magic squares of order $4n (n = 1, 2, \ldots)$ are discussed in Section 4.

2. Notation

Let n and p be integers such that

$$2 \leqslant p \leqslant n.$$

We denote by $A_{n,p}$ the $n \times n$ matrix each of whose elements is one of the integers $0, 1, \ldots, n - 1$ and which is constructed as follows. The first row of $A_{n,p}$ is the vector

$$(0, 1, \ldots, n - 1),$$

and the second row is the vector

$$(p - 1, p, \ldots, n - 1, 0, 1, \ldots, p - 2).$$

Generally, every row is of the form

$$(q - 1, q, \ldots, n - 1, 0, 1, \ldots, q - 2); \tag{1}$$

and the $(i + 1)$th row is the vector of the form (1) which begins with the pth element of the ith row. Thus $A_{n,n}$ is the usual 'circulant'

$$\begin{bmatrix} 0 & 1 & \ldots & n-2 & n-1 \\ n-1 & 0 & \ldots & n-3 & n-2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 2 & \ldots & n-1 & 0 \end{bmatrix},$$

while $A_{n,2}$ is the matrix

$$\begin{bmatrix} 0 & 1 & \ldots & n-2 & n-1 \\ 1 & 2 & \ldots & n-1 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ n-1 & 0 & \ldots & n-3 & n-2 \end{bmatrix}.$$

It follows from the definition that, if

$$A_{n,p} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

and if

$$a_{ij} = m,$$

then

$$a_{i+1,j} = \begin{cases} m + p - 1 & \text{if} \quad m + p - 1 \leqslant n - 1, \\ m + p - 1 - n & \text{if} \quad m + p - 1 \geqslant n. \end{cases}$$

Hence the jth column of $A_{n,p}$ is the vector

$$(r_1, r_2, \dots, r_n)$$

(written horizontally instead of vertically) in which $r_1 = j - 1$ and, generally, r_i is the remainder after the division of $(j - 1) + (i - 1)(p - 1)$ by n. Thus, for $i = 1, 2, \dots, n$,

$$(j - 1) + (i - 1)(p - 1) = \alpha_i n + r_i, \tag{2}$$

where α_i is an integer and $0 \leqslant r_i \leqslant n - 1$.

The matrix $B_{n,p}$ is constructed from the elements $1, 2, \dots, n$ in the same way as $A_{n,p}$ was constructed from $0, 1, \dots, n - 1$. For instance

$$B_{n,n} = \begin{bmatrix} 1 & 2 & \dots & n \\ n & 1 & \dots & n-1 \\ \dots & \dots & \dots & \dots \\ 2 & 3 & \dots & 1 \end{bmatrix}, \quad B_{n,2} = \begin{bmatrix} 1 & 2 & \dots & n \\ 2 & 3 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ n & 1 & \dots & n-1 \end{bmatrix}.$$

3. Odd order magic squares

Lemma 1. Let n, p be integers such that $2 \leqslant p \leqslant n$ and the pair $(n, p - 1)$ is coprime (i.e. the highest common factor of n and $p - 1$ is 1). Then every column of $A_{n,p}$ contains all the integers $0, 1, \dots, n - 1$; and every column of $B_{n,p}$ contains all the integers $1, 2, \dots, n$.

Proof. Clearly we need only consider the matrix $A_{n,p}$. The jth column of $A_{n,p}$ is the vector

$$(r_1, r_2, \dots, r_n)$$

with r_i given by (2). Hence, when $1 \leqslant i \leqslant k \leqslant n$,

$$\begin{aligned} r_i = r_k &\Leftrightarrow (j-1) + (i-1)(p-1) - \alpha_i n = (j-1) + (k-1)p - \alpha_k n \\ &\Leftrightarrow (k-i)(p-1) = (\alpha_k - \alpha_i)n \\ &\Leftrightarrow n \text{ divides } (k-i)(p-1). \end{aligned}$$

Since n and $p-1$ are coprime, the last statement holds if and only if n divides $k-i$. But we have $0 \leqslant k-i \leqslant n-1$ and therefore n can divide $k-i$ only if $k-i=0$. Thus $r_i = r_k$ if and only if $i=k$; in other words, the elements of $(r_1, r_2, \ldots, r_n)$ are all distinct. As every r_i is one of the integers $0, 1, \ldots, n-1$, the vector $(r_1, r_2, \ldots, r_n)$ is simply a rearrangement of $(0, 1, \ldots, n-1)$.

In the matrix

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \tag{3}$$

the jth *left-to-right diagonal* (where $1 \leqslant j \leqslant n$) is defined to be the vector

$$(a_{1j}, a_{2,j+1}, \ldots, a_{n-j+1,n}, a_{n-j,1}, \ldots, a_{n,j-1})$$

with n components. The case $n=5, j=4$ is represented diagrammatically below:

$$\begin{bmatrix} \cdot & \cdot & \cdot & * & \cdot \\ \cdot & \cdot & \cdot & \cdot & * \\ * & \cdot & \cdot & \cdot & \cdot \\ \cdot & * & \cdot & \cdot & \cdot \\ \cdot & \cdot & * & \cdot & \cdot \end{bmatrix}.$$

Lemma 2. Let (n, p) be integers such that $2 \leqslant p \leqslant n-1$ and the pair (n, p) is coprime. Then every left-to-right diagonal of $A_{n,p}$ contains all the integers $0, 1, \ldots, n-1$; and every left-to-right diagonal of $B_{n,p}$ contains all the integers $1, 2, \ldots, n$.

Proof. We again consider $A_{n,p}$ only. If the jth left-to-right diagonal is

$$(s_1, s_2, \ldots, s_n),$$

then

$$s_1 = j-1$$

and

$$s_2 = \begin{cases} j-1+p & \text{if} \quad j-1+p \leqslant n-1, \\ j-1+p-n & \text{if} \quad j-1+p \geqslant n. \end{cases}$$

The formula giving an arbitrary s_i is analogous to (2):

$$(j-1) + (i-1)p = \alpha_i n + s_i, \tag{4}$$

where α_i is an integer and $0 \leqslant s_i \leqslant n-1$.

An argument just like the one used in the proof of Lemma 1 now shows that, if n and p are coprime, $s_i = s_k$ if and only if $i = k$. It then follows as before that $(s_1, s_2, \ldots, s_n)$ is a rearrangement of $(0, 1, \ldots, n - 1)$.

Note. By (4), the jth left-to-right diagonal of $A_{n,n}$ has all its elements equal to $j - 1$. Similarly, all elements of the jth left-to-right diagonal of $B_{n,n}$ are equal to j.

In the matrix (3), the jth *right-to-left diagonal* is the vector

$$(a_{1j}, a_{2,j-1}, \ldots, a_{j1}, a_{j+1,n}, \ldots, a_{n,j-1}).$$

Lemma 3. Let n, p be integers such that $3 \leqslant p \leqslant n$ and the pair $(n, p - 2)$ is coprime. Then every right-to-left diagonal of $A_{n,p}$ contains all the integers $0, 1, \ldots, n - 1$; and every right-to-left diagonal of $B_{n,p}$ contains all the integers $1, 2, \ldots, n$.

Proof. The jth right-to-left diagonal of $A_{n,p}$ is

$$(t_1, t_2, \ldots, t_n),$$

where $t_1 = j - 1$; and all t_i are given by the relation

$$(j - 1) + (i - 1)(p - 2) = \alpha_i n + t_i \tag{5}$$

in which α_i is an integer and $0 \leqslant t_i \leqslant n - 1$. From this point the proof follows the same lines as those used in the proofs of Lemmas 1 and 2.

Note. By (5), the jth right-to-left diagonal of $A_{n,2}$ has all its elements equal to $j - 1$. All elements of the jth right-to-left diagonal of $B_{n,2}$ are equal to j.

At this point it is useful to pick out some of the situations in which the conditions of Lemmas 1, 2 and 3 are satisfied.

(i) For every integer $n \geqslant 2$, n and $n - 1$ are coprime. (If an integer $r \geqslant 2$ divides $n - 1$, then division of n by r produces the remainder 1.)

(ii) For every odd integer $n \geqslant 3$, n and $n - 2$ are coprime. (Neither of n, $n - 2$ is divisible by an even integer; and reasoning of the kind used in (i) shows that $3, 5, 7, \ldots$ cannot be common factors.)

(iii) When n is an odd prime, n and p are coprime for $2 \leqslant p \leqslant n - 1$.

Theorem 1. For every odd integer $n \geqslant 3$, the matrix

$$M_n = B_{n,2} + nA_{n,n}$$

is a magic square.

Proof. Each row sum of $B_{n,2}$ is $1 + 2 + \cdots + n = \frac{1}{2}n(n + 1)$; and each row sum of $A_{n,n}$ is $0 + 1 + \cdots + n - 1 = \frac{1}{2}(n - 1)n$.

Since the pairs $(n,1)$ and $(n, n - 1)$ are coprime, it follows from Lemma 1 that each column sum of $B_{n,2}$ is also $\frac{1}{2}n(n + 1)$ and each column sum of $A_{n,n}$ is $\frac{1}{2}(n - 1)n$.

Thus every row sum and every column sum of M_n is

$$\tfrac{1}{2}n(n + 1) + n.\tfrac{1}{2}(n - 1)n = \tfrac{1}{2}n(n^2 + 1),$$

as required.

It has still to be shown that the elements of M_n are distinct and that every element is one of the integers $1, 2, \ldots, n^2$. For this purpose we consider the left-to-right diagonals of M_n. By Lemma 2, each of these diagonals in $B_{n,2}$ contains all the integers $1, 2, \ldots, n$; and by the note after Lemma 2, the jth left-to-right diagonal of $A_{n,n}$ has all its elements equal to $j - 1$. Hence the elements of the jth left-to-right diagonal of M_n comprise the elements of the set

$$L_j = \{1 + n(j-1), 2 + n(j-1), \ldots, n + n(j-1)\}.$$

If $1 \leqslant j < k \leqslant n$, then every element of L_j is less than every element of L_k. For the largest element of L_j is $n + n(j-1)$, while the smallest element of L_k is $1 + n(k-1)$ and

$$n + n(j-1) = nj \leqslant n(k-1) < 1 + n(k-1).$$

Therefore, no two of the sets L_j have an element in common, and this means that all the n^2 elements of M_n are distinct.

Finally, every element of M_n is a positive integer not exceeding $n + n(n-1) = n^2$. This completes the proof of the theorem.

Theorem 2. Let n be an odd integer greater than or equal to 5, and let p be an integer such that $3 \leqslant p \leqslant n - 1$ and each of the pairs $(n, p-1)$, $(n, p-2)$ is coprime. Then the matrix

$$M_{n,p} = B_{n,2} + nA_{n,p}$$

is a magic square.

Proof. As in the proof of Theorem 1 we see that the row sums of $B_{n,2}$ and of $A_{n,p}$ are $\frac{1}{2}n(n+1)$ and $\frac{1}{2}(n-1)n$ respectively. Moreover Lemma 1 shows that, since the pairs $(n,1)$ and $(n, p-1)$ are coprime, the column sums of $B_{n,2}$ and $A_{n,p}$ are also $\frac{1}{2}n(n+1)$ and $\frac{1}{2}(n-1)n$ respectively. Therefore all the rows and columns of $M_{n,p}$ have the required sum $\frac{1}{2}n(n^2+1)$.

Next we consider right-to-left diagonals. Since the pair $(n, p-2)$ is coprime, Lemma 3 ensures that every such diagonal of $A_{n,p}$ contains all the integers $0, 1, \ldots, n-1$. On the other hand, by the note after Lemma 3, the jth right-to-left diagonal of $B_{n,2}$ has all its elements equal to j. Therefore the jth right-to-left diagonal of $M_{n,p}$ consists of the elements of the set

$$R_j = \{j + n.0, j + n.1, \ldots, j + n(n-1)\}.$$

If $1 \leqslant j < k \leqslant n$, then R_j and R_k cannot have a common element. For suppose that a belongs to both R_j and R_k. Then

$$a = j + ns, \qquad \text{where} \qquad 0 \leqslant s \leqslant n - 1$$

and also

$$a = k + nt, \qquad \text{where} \qquad 0 \leqslant t \leqslant n - 1.$$

It follows that

$$k - j = n(s - t),$$

i.e. that n divides $k - j$. But this is impossible since $1 \leqslant k - j \leqslant n - 1$. Hence no integer appears in more than one R_j and so all elements of $M_{n,p}$ are distinct. It is now easy to see that the elements of $M_{n,p}$ are precisely the integers $1, 2, \ldots, n^2$. Thus $M_{n,p}$ is a magic square.

Theorems 1 and 2 enable us to exhibit any number of magic squares of odd order. Two examples are

$$M_3 = B_{3,2} + 3A_{3,3} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} + 3 \begin{bmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 9 \\ 8 & 3 & 4 \\ 6 & 7 & 2 \end{bmatrix};$$

and

$$M_{5,4} = B_{5,2} + 5A_{5,4} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{bmatrix} + 5 \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 3 & 4 & 0 & 1 & 2 \\ 1 & 2 & 3 & 4 & 0 \\ 4 & 0 & 1 & 2 & 3 \\ 2 & 3 & 4 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 7 & 13 & 19 & 25 \\ 17 & 23 & 4 & 10 & 11 \\ 8 & 14 & 20 & 21 & 2 \\ 24 & 5 & 6 & 12 & 18 \\ 15 & 16 & 22 & 3 & 9 \end{bmatrix}.$$

4. Magic squares of the order $4n$

Let n be any positive integer. We begin by arranging the integers $1, 2, 3, \ldots, 4n$ as the vector

$$w = (2n + 1, 2n, 2n + 2, 2n - 1, \ldots, 4n - 1, 2, 4n, 1)$$

in which the 1st, 3rd, . . ., $(4n - 1)$th positions are occupied by the integers $2n + 1$, $2n + 2, \ldots, 4n$, and the 2nd, 4th, . . ., $4n$th are filled by $2n, 2n - 1, \ldots, 1$. We note that, for $k = 1, 2, \ldots, 2n$, the sum of the $(2k - 1)$th element and the $2k$th element is $4n + 1$. A vector x is obtained from w by taking the last two elements of w, making these the first two elements of x and leaving the relative positions of the other elements of w unchanged. Thus

$$x = (4n, 1, 2n + 1, 2n, 2n + 2, 2n - 1, \ldots, 4n - 1, 2).$$

The vector x' is constructed from x by interchanging the $(2k - 1)$th element and the $2k$th one (for $k = 1, 2, \ldots, 2n$). So

$$x' = (1, 4n, 2n, 2n + 1, 2n - 1, 2n + 2, \ldots, 2, 4n - 1).$$

We now define X_{4n} to be the $4n \times 4n$ matrix whose first n rows are x, whose middle $2n$ rows are x' and last n rows are x again. For example

$$X_4 = \begin{bmatrix} 4 & 1 & 3 & 2 \\ 1 & 4 & 2 & 3 \\ 1 & 4 & 2 & 3 \\ 4 & 1 & 3 & 2 \end{bmatrix}, \quad X_8 = \begin{bmatrix} 8 & 1 & 5 & 4 & 6 & 3 & 7 & 2 \\ 8 & 1 & 5 & 4 & 6 & 3 & 7 & 2 \\ 1 & 8 & 4 & 5 & 3 & 6 & 2 & 7 \\ 1 & 8 & 4 & 5 & 3 & 6 & 2 & 7 \\ 1 & 8 & 4 & 5 & 3 & 6 & 2 & 7 \\ 1 & 8 & 4 & 5 & 3 & 6 & 2 & 7 \\ 8 & 1 & 5 & 4 & 6 & 3 & 7 & 2 \\ 8 & 1 & 5 & 4 & 6 & 3 & 7 & 2 \end{bmatrix}.$$

Next we let

$$y = (0, 4n - 1, 1, 4n - 2, \ldots, 2n - 2, 2n + 1, 2n - 1, 2n)$$

and

$$y' = (4n - 1, 0, 4n - 2, 1, \ldots, 2n + 1, 2n - 2, 2n, 2n - 1),$$

so that, in each case, the sum of the $(2k - 1)$th element and the $2k$th is $4n - 1$. Then Y_{4n} is the $4n \times 4n$ matrix whose first $2n$ columns are y and last $2n$ columns are y'. Thus

$$Y_4 = \begin{bmatrix} 0 & 0 & 3 & 3 \\ 3 & 3 & 0 & 0 \\ 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 \end{bmatrix}, \quad Y_8 = \begin{bmatrix} 0 & 0 & 0 & 0 & 7 & 7 & 7 & 7 \\ 7 & 7 & 7 & 7 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 6 & 6 & 6 & 6 \\ 6 & 6 & 6 & 6 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & 5 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 3 & 3 & 3 & 3 \end{bmatrix}.$$

Theorem 3. For every integer $n \geqslant 1$, the $4n \times 4n$ matrix

$$M_{4n} = X_{4n} + 4nY_{4n}$$

is a magic square.

Proof. Each row of X_{4n} contains all the integers $1, 2, \ldots, 4n$ and their sum is $2n(4n + 1)$. A typical row of Y_{4n} contains $2n$ elements equal to p, say, and $2n$ equal to $4n - 1 - p$. Hence every row sum of Y_{4n} is $2n(4n - 1)$. Thus every row sum of M_{4n} is

$$2n(4n + 1) + 4n.2n(4n - 1) = \tfrac{1}{2}.4n[(4n)^2 + 1],$$

which is the required value.

A column of X_{4n} consists of $2n$ integers q, say, and $2n$ integers equal to $4n + 1 - q$. Therefore every column sum of X_{4n} is $2n(4n + 1)$. Every column of Y_{4n} consists of the integers $0, 1, \ldots, 4n - 1$ whose sum is $2n(4n - 1)$. Hence all column sums of M_{4n} are also equal to $\frac{1}{2}.4n[(4n)^2 + 1]$.

To prove that the elements of M_{4n} are the integers $1, 2, \ldots, (4n)^2$ we consider 'zigzag' columns. When j is odd, the jth zigzag column of an arbitrary $4n \times 4n$ matrix consists of the first n elements of column j, then the middle $2n$ elements of column $j + 1$ and finally the last n elements of column j. When j is even the construction of the jth zigzag column is similar, but the middle $2n$ elements come from column $j - 1$.

It is clear that the jth zigzag column of X_{4n} has all its elements equal to x_j, the jth element of the first row. On the other hand, every zigzag column of Y_{4n} is either y or y'. Hence the jth zigzag column of M_{4n} consists of all the elements from the set

$$Z_j = \{x_j + 4n.0, x_j + 4n.1, x_j + 4n.2, \ldots, x_j + 4n(4n - 1)\}.$$

As j varies from 1 to $4n$, x_j takes all the values between 1 and $4n$. An argument similar to the one employed for the right-to-left diagonals of R_j which figured in the proof of Theorem 2 now shows that no integer appears in more than one Z_j. This means that all elements of M_{4n} are distinct and it is easily seen that they are, in fact, the integers $1, 2, \ldots, (4n)^2$. Thus M_{4n} is a magic square.

The 4×4 and 8×8 magic squares obtained from Theorem 3 are

$$M_4 = \begin{bmatrix} 4 & 1 & 15 & 4 \\ 13 & 16 & 2 & 3 \\ 5 & 8 & 10 & 11 \\ 12 & 9 & 7 & 6 \end{bmatrix}, \quad M_8 = \begin{bmatrix} 8 & 1 & 5 & 4 & 62 & 59 & 63 & 58 \\ 64 & 57 & 61 & 60 & 6 & 3 & 7 & 2 \\ 9 & 16 & 12 & 13 & 51 & 54 & 50 & 55 \\ 49 & 56 & 52 & 53 & 11 & 14 & 10 & 15 \\ 17 & 24 & 20 & 21 & 43 & 46 & 42 & 47 \\ 41 & 48 & 44 & 45 & 19 & 22 & 18 & 23 \\ 32 & 25 & 29 & 28 & 38 & 35 & 39 & 34 \\ 40 & 33 & 37 & 36 & 30 & 27 & 31 & 26 \end{bmatrix}.$$

I am grateful to Professor J. Medhi for helpful discussions in the preparation of this note.

References

1. W. S. Andrews, *Magic Squares and Cubes*, 2nd ed. (Open Court Publishing, Chicago, 1917).
2. T. M. Apostol and H. S. Zuckerman, On magic squares constructed by the uniform step method, *Proc. Amer. Math. Soc.* **2** (1951), 557–565.
3. C. R. Johnson, A matrix theoretic construction of magic squares, *Amer. Math. Monthly*, **79** (1972), 1004–1006.
4. B. Rossner and R. J. Walker, The algebraic theory of diabolic magic squares, *Duke Math. J.* **5** (1939), 705–728.

Correction

Can You Contribute to Time Series Research?

O. D. ANDERSON
Volume 8, Number 1, pp. 21–23

Page 22, lines 18–20 should read:

> to be fully assimilated and hence A_{10}, say, will have unit effect on Z_{10}, θ_1 effect on Z_{11}, θ_2 effect on Z_{12} and so on.

Page 22, line 25 should read:

$$\rho_j = \mathrm{Cov}[Z_i, Z_{i-j}]/\mathrm{Var}[Z_i] \quad j > 1,$$

Page 22, line 28 should read:

> r_j, and hence hope to recognise any underlying process. To do this, the theoretical

Page 23, line 11 should read:

> (reference 1). Note that this agrees with (1) and (2), and for an MA(0), the

The Editor regrets that the errors listed above were inadvertently introduced into this article; he hopes that readers who were previously confused will now return to the corrected version.

Letters to the Editor

Dear Editor,

A simple method of linear curve fitting

I have developed a method of fitting a straight line to a given set of points which may be of interest to your readers. It is easier to use than the usual method of least squares, and in practice the line obtained is nearly the same as the least squares line. However, no formal justification of the method exists at present.

The procedure is best demonstrated by a worked example. We require an even number of points; suppose that the points are

$$(1,3),\ (2,5),\ (3,4),\ (4,5),\ (5,9),\ (6,7).$$

This set of points is divided into two of equal size, with the points in the *lower set* having smaller x-coordinates than the points in the *upper set*. Let A, B be the centroids of the points in the lower and upper sets respectively; thus $A = (2,4)$ (since 2 is the average of 1, 2, 3, while 4 is the average of 3, 5, 4) and $B = (5,7)$. We call A the *lower centroid* and B the *upper centroid*. The join of A and B, namely $y = x + 2$, is the required line.

The method of least squares fits a straight line in such a way that the sum of the squares of the vertical distances of the given points from the fitted line is minimised. The method of joining the upper and lower centroids ensures that the algebraic sum of the vertical distances is zero for both the lower and upper sets of points. In both methods the line passes through the centroid of all the points.

For comparison, the two methods were used to fit straight lines to a large number of sets of points. In each case the gradient, y-intercept and sum of squared vertical deviations for the two lines were calculated. The percentage by which the sum for the line joining the centroids exceeds the least sum is a measure of the effectiveness of the method. I have never yet found the excess to be greater than 10%. Below are two examples.

1. Set of points
   (0,0), (2,3), (4,9), (6,11), (8,16), (10,21), (12,25), (14,28), (16,31), (18,33).

	Join of centroids	Least squares line
Gradient	1·98	1·92
y-intercept	−0·12	0·38
Sum of squared deviations	13·23	12·21

13·23 exceeds 12·21 by 8·41%.

2. Set of points
   (8,1380), (9,1292), (10,1174), (11,1113), (12,1066), (13,978), (14,904), (15,820), (16,752), (17,661).

	Join of centroids	Least squares line
Gradient	−76·40	−77·19
y-intercept	1969·00	1978·80
Sum of squared deviations	2148·80	2097·60

2148·8 exceeds 2097·6 by 2·44%.

It can be seen that in practice the lines obtained by the two methods are very nearly the same.

Yours sincerely,

P. J. Bishop
(Golborne Comprehensive School, Golborne,
Warrington, Lancs.)

Dear Editor,

The divergence of the harmonic series

After Captain Draim's delightful article (Vol. 7, No. 1, 9–12), your readers may be interested in seeing two more proofs of the divergence of the harmonic series

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

These indirect proofs make use of the following simple facts:

(i) In a convergent series brackets may be inserted at will (i.e. the resulting series converges to the same sum).

(ii) A convergent series may be multiplied by a constant, term by term, and two convergent series may be added term by term (i.e. the resulting series have the 'natural' sums).

In both proofs we begin by supposing that the series (1) converges to sum S, say.

(A) Using (i) we have

$$S = \left(1 + \frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \cdots + \left(\frac{1}{2n-1} + \frac{1}{2n}\right) + \cdots.$$

Since

$$\frac{1}{2n-1} + \frac{1}{2n} > \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$$

for $n = 1, 2, \ldots$, it follows that

$$S > 1 + \frac{1}{2} + \cdots + \frac{1}{n} + \cdots.$$

Thus $S > S$, and the contradiction shows that the series (1) cannot converge.

(B) By (ii)

$$\frac{1}{2}S = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \cdots + \frac{1}{2n} + \cdots$$

and

$$S - \frac{1}{2}S = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1} + \cdots$$

$$> \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \cdots + \frac{1}{2n} + \cdots = \frac{1}{2}S.$$

Hence $\frac{1}{2}S > \frac{1}{2}S$ and we again have a contradiction.

Yours sincerely,

M. R. Chowdhury
(Jahangirnagar University, Dacca, Bangladesh)

Problems and Solutions

Sixth formers and students are invited to submit solutions to some or all of the problems below: the most attractive solutions will be published in subsequent issues. When writing to the Editorial Office, please state your full name and the postal address of your school, college or university.

Problems

8.4. Is it possible to partition the integers 1, 2, . . ., 13 into two subsets such that neither subset possesses three integers in arithmetic progression?

8.5. (Submitted by I. D. Macdonald, University of Stirling.) Let r, s be positive integers with $r > s$. Prove that

$$\frac{1}{r-s}+\frac{1}{r-s+1}+\cdots+\frac{1}{r}+\cdots+\frac{1}{r+s}>\frac{2s+1}{r},$$

and deduce that, if n is an integer greater than 1 and $m = (3^n - 1)/2$, then

$$1+\frac{1}{2}+\cdots+\frac{1}{m}>n.$$

(This provides another proof of the divergence of the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \cdots$ — see Volume 7 No. 1 and Volume 8 No. 1.)

8.6. (Submitted by B. G. Eke, University of Sheffield.) Show that the product of four consecutive positive integers cannot be a perfect cube.

Solutions to Problems in Volume 7, Number 3

7.7. Distinct points L and M are given in the plane, and k is a real number such that $0 < k < 1$. Then the locus of all points X in the plane such that $LX/MX = k$ is a circle (Apollonius' Circle). A tangent is drawn through M to touch the circle at T. Show that the angle TLM is a right angle.

Solution by M. Ram Murty and V. Kumar Murty (Carleton University, Ottawa)

Let H and K be the points of internal and external division respectively of the line LM such that LH:$MH = LK$:$MK = k$. Then H, K lie on the given circle, so L is in the interior of the circle. Furthermore, since $LH < MH$, M must lie in the exterior of the circle. Since T is a point on the circle, LT:$MT = k = LH$:MH, so that TH bisects $\angle LTM$ internally. Similarly, since LT:$MT = LK$:MK, TK bisects $\angle LTM$ externally. Hence $\angle HTL + \angle LTK = 90°$. But $\angle MTH = \angle TKL$ because MT is a tangent to the circle. Since $\angle HTL = \angle MTH$, this gives $\angle TKL + \angle LTK = 90°$. From ΔTLK, it follows that $\angle TLK$ is a right angle, so that $\angle TLM$ is also a right angle.

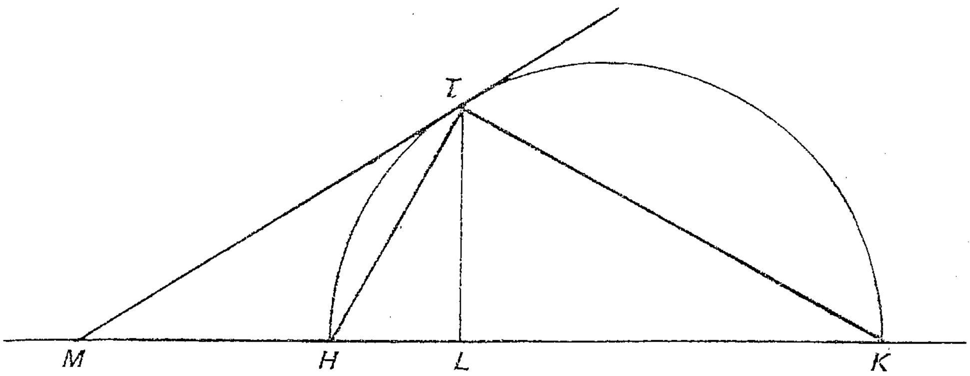
Also solved by Geraint Jones (Yale Sixth Form College, Wrexham) and Thomas Gedrich (Kingston Grammar School).

7.8. Use the identity

$$\frac{4}{1+t^2}=4-4t^2+5t^4-4t^5+t^6-\frac{t^4(1-t)^4}{1+t^2}$$

to show that

$$\frac{22}{7}-\frac{1}{1260}>\pi>\frac{22}{7}-\frac{1}{630}.$$



Solution by Thomas Gedrich

If we integrate both sides of the identity between 0 and 1, we obtain

$$\pi = \frac{22}{7} - \int_0^1 \frac{t^4(1-t)^4}{1+t^2}\,dt.$$

But

$$\int_0^1 \frac{t^4(1-t)^4}{2}\,dt < \int_0^1 \frac{t^4(1-t)^4}{1+t^2}\,dt < \int_0^1 t^4(1-t)^4\,dt$$

and

$$\int_0^1 t^4(1-t)^4\,dt = \frac{1}{630}.$$

The result follows.

Also solved by Geraint Jones, M. Ram Murty and V. Kumar Murty.

7.9. Let a be a positive integer and let b, c be integers. Suppose that $ax^2 + bx + c$ has two distinct roots in the range $0 < x < 1$. Show that $a \geqslant 5$ and find such a quadratic with $a = 5$.

Solution by Geraint Jones

Suppose first that the sum of the roots of the polynomial is less than or equal to 1. Then $0 < -b/a \leqslant 1$, or $0 < -b \leqslant a$. Since the roots are real and distinct, we also have $b^2 > 4ac$. Thus $a^2 \geqslant b^2 > 4ac$, so that $a > 4c$. But $c > 0$ because the roots are positive, so that $c \geqslant 1$ and $a > 4$, i.e. $a \geqslant 5$.

Now suppose that the sum of the roots is greater than 1. If we put $y = 1 - x$, we obtain the polynomial

$$ay^2 - (2a + b)y + (a + b + c),$$

and the sum of the roots of this polynomial will be less than 1. We can therefore apply the above considerations to this polynomial. Since its leading coefficient is a, this again gives that $a \geqslant 5$.

The quadratic

$$5x^2 - 5x + 1$$

has roots

$$\frac{1}{2} \pm \frac{1}{2(5)^{\frac{1}{2}}},$$

and so satisfies the conditions.

Also solved by Thomas Gedrich, M. Ram Murty and V. Kumar Murty.

Book Reviews

History of Mathematics: Topics for Schools. By WALTER POPP, translated by Maxim Bruckheimer. Transworld Student Library, Transworld Publishers Ltd, London, 1975. Pp. viii+150. £0·85 paperback.

The history of mathematics is likely to find a place in an increasing number of school mathematics courses in the near future. Such a development could provide much interest and background for the student and it is to be hoped that it will not trigger off a further batch of tedious examination questions. The new Open University History of Mathematics course (beginning early in 1976) will include material which could be used in schools and this book will be a useful link between this course and possible school-based work.

The book does not claim to be any more than a selection of topics from the history of mathematics and it will undoubtedly provide ideas and material for the teacher although he will probably need to consult other, more detailed, accounts at the same time. In Part I there are sections on Algebra, Arithmetic and Geometry described as being suitable for the lower and middle school. Part II, for the upper school, includes Series, Calculus and Analytic Geometry. The presentation and style of the book are rather dull; there is a total lack of humour and the odd anecdote from the history of mathematics would have done much to enliven the account. Any pupil below the sixth form would find the book difficult.

An important but disappointing book, it should nevertheless find a place on the teacher's shelf and it could well stimulate further reading and investigation.

University of Durham — M. L. CORNELIUS

An Introduction to Modern Algebra. By B. W. JONES. Collier Macmillan, London, 1975. Pp. 362. £4·75.

This book is aimed at first- and second-year undergraduates. The first four chapters cover material relevant to a first-year university course on algebra, starting with some group theory with examples taken from groups of matrices, the Orthogonal Group, groups of symmetries and permutations. Lagrange's theorem and Cayley's theorem are proved. However, the group theory is not taken very far, in particular no mention is made of the Sylow theorems.

Chapter 3 is a good chapter on Fields, Integral Domains and Rings. It includes the usual results on the integers. Polynomials and algebraic extensions of fields are introduced here. Chapter 4 is a treatment of elementary linear algebra. The basic theorems on finite dimensional vector spaces are covered, and determinants are discussed. However, the book is obviously not intended to be a text on linear algebra, and first-year undergraduates would need an additional treatment of this subject.

Chapter 5 is an elementary treatment of the concept of an ideal. The last chapter is very good. It is an introduction to Galois theory, including a discussion of constructibility and the solution of cubic and quartic equations.

Finally we note that the author chooses to write his mappings on the right. The book is made more interesting by the biographical notes describing some of the famous mathematicians of history.

University of Durham — R. S. ROBERTS

Calculus of Variations. By A. M. ARTHURS. Routledge and Kegan Paul, London, 1975. Pp. 80. £1·50 paperback.

Despite its brevity this book provides a straightforward introduction to the calculus of variations. The text is easy to read, and there are many worked examples with a good number of exercises at the end of each chapter. There are, however, one or two surprising omissions, although I suspect that this is inevitable in such a short book.

The first chapter introduces the calculus of variations through the usual classical problems, namely the shortest path, minimal surface, brachistrochrone and isoperimetric problems. The Euler–Lagrange equation is derived, but it seems unfortunate that the brachistochrone problem is *not* solved in the text, but only set as an exercise.

The second chapter provides extensions, including the important case of variable end points and an introduction to variational principles in mechanics. The third chapter, on minimal principles, provides a collection of rather non-related problems, such as an introduction to dynamic programming and the solution of isoperimetric problems. It does seem a shame that no mention is made here or earlier of the extension of variational calculus to optimal control. The final chapter is devoted to direct methods, concentrating on the Rayleigh–Ritz method.

The book is in the Library of Mathematics series and should be welcome to undergraduate students.

University of Newcastle D. BURGHES

Elementary Partial Differential Equations. By R. J. GRIBBEN. Van Nostrand Reinhold Company Ltd, London, 1975. Pp. 121. £1·50 paperback.

Dr Gribben's book is a brave attempt at trying to present a short and elementary approach to partial differential equations. Although it is well written and presented, I feel that most undergraduate students would find it difficult to cope with unless it is used as a companion to a good lecture course. In spite of a large number of exercises set at the end of each chapter, there are far too few worked examples in the text, particularly in the first chapter.

The first chapter contains an introduction full of important techniques, definitions and theorems. Partial differential equations are introduced through mathematical modelling, and throughout the book emphasis is rightly placed on the derivation of p.d.e.'s in various physical systems. Linearity, superposition of solutions, classification and uniqueness are all dealt with in this chapter, together with an introduction to the required results on Fourier series. Here I found a rather confusing notation, x' being used as a variable. It surely would have been better to use y, for instance.

The rest of the book is devoted to chapters on the wave equation, diffusion equation and Laplace's equation. I enjoyed these chapters and, with the above reservation, this book makes a welcome addition to the New Mathematics Library series.

University of Newcastle D. BURGHES

Introductory Vector Analysis. By C. D. COLLINSON. Edward Arnold, London, 1974. Pp. 161. £2·50 paperback; £5·50 hardback.

This book presents a rigorous treatment and hardly an introduction to vector analysis; it is rather more suitable for mathematicians than for scientists who might require only a *working* knowledge of the subject.

Much emphasis is placed on the use of index notation, which is introduced in chapter two. My experience has taught me that the 'summation convention', if not stressed when first introduced, will always cause difficulties throughout a vector analysis course. Dr Collinson gives a reasonable introduction, although I would have liked to see a few *simple* examples on its use.

Despite the moderate length of the book, all the expected topics are well covered. It includes chapters on the calculus of functions of several variables, line, surface and volume integrals, and a short introduction to tensors. Although there are usually plenty of worked examples, most chapters are surprisingly short of set exercises. I say surprisingly, since I get the impression on reading this text that Dr Collinson well understands the difficulties encountered by students.

University of Newcastle D. BURGHES

The Factor Book. By R. L. Hubbard. Hilton Management Services Ltd, Lytham St Annes, 1975. Pp. 6+100,000÷500. £3·00.

The theory of numbers is generally considered to be the 'purest' (and most useless) branch of pure mathematics. However, it has one feature in common with other sciences, namely the important rôle played by experiments, which take the form of testing and even guessing general theorems by numerical examples. Gauss used calculations extensively, and in recent times the possibility of experiment has been greatly enhanced by the power and potentiality of the computer.

The present volume lists the factorisation of numbers from 1 to 100,000. It is clearly set out in columns containing 500 entries on a page (whence the calculation offered in the details at the beginning of this review). The book is easy to handle. It should prove to be of immense value to number theorists, whether they are setting examination papers or just playing about with numbers.

College of St Hild and St Bede, Durham J. V. Armitage

Notes on Contributors

J. H. Durran is Head of the Mathematics Department at Winchester College. He became interested in statistics in the early 1960's when the SMP courses were being written. He is the author of *Statistics and Probability* (C.U.P.), and Chairman of the Committee on Statistical Education. As the second British jury member he went with the British team to the IMO in Bulgaria this year.

R. J. Stammers was employed in the Data Administration Team of the Data Processing Division of the Open University until 1974 when he became Senior Research Officer in the Department of Computing at the University of Essex. He is an external research student of the Open University studying for a higher degree.

R. N. Maddison is a Senior Lecturer in Mathematics at the Open University. Previously he was employed by British Railways Board and lectured at Glasgow and Sheffield Universities. His main interests are in computer systems and their many applications.

Frederick Stern is an Associate Professor of Mathematics at San Jose State University in San Jose, California. He is a graduate of the Courant Institute, New York University, and has done operational research for industry. His research interests include probability and statistics.

A. D. Misra is a Lecturer in Mathematics at Gauhati University in Assam, India. He obtained his first degree and M.Sc. from the same university, and is now working for a doctorate. His research interests are in the fields of linear algebra and mathematical programming.

Contents

PRICES (*postage included*)

Prices for Volume 8 (Issues Nos. 1, 2 and 3):

Subscribers in Britain and Europe: £0·70
Subscribers overseas: £1·40 (US$3.50; $A.2.40)
(These prices apply even if the order is placed by an agent in Britain.)
A discount of 10% is allowed on all orders for five or more copies.

Back issues:

Volume 1 is out of print. All other back issues are still available at the following prices:

Volumes 2, 3, 4, 5 and 6 (2 issues each volume):
£1·00 (US$2.50; $A.1.70) per volume.

Volume 7 (3 issues):
£1·40 (US$3.50; $A.2.40).

Enquiries about rates, subscriptions and advertisements should be directed to:

Editor—*Mathematical Spectrum,*
Hicks Building,
The University,
Sheffield S3 7RH, England.

Printed in England by Galliard (Printers) Ltd, Great Yarmouth