Mathematical Spectrum

Volume 9 1976/77 *Number* 1

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 9 of *Mathematical Spectrum* will consist of three issues, the second of which will be published in January 1977 and the third in May 1977.

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.

A Case Study in Inequalities

L. MIRSKY
University of Sheffield

Anyone who has advanced beyond the bare rudiments of arithmetic will have met inequalities, i.e. statements asserting that a particular number (or algebraic expression) is greater than some other number (or algebraic expression). Thus $10 > \pi^2$ or, to cite a less prosaic example,

$$3 + 4\cos\theta + \cos 2\theta \geqslant 0 \tag{1}$$

for every angle θ, and the sign $\geqslant$ in (1) cannot be replaced by $>$. (No self-respecting reader will wish to proceed further before convincing himself of the truth of these statements.)

Inequalities (of the most varied type) have always played a prominent role in mathematics, and especially in the part of mathematics that is covered by the general name of 'analysis'. Yet, for a long time, their study was completely haphazard, and it is only since the thirties that the topic has assumed the stature of a systematic discipline. Several books, both on inequalities in general and on inequalities of special kinds, are now available; and the subject is continually gaining in strength, richness of content and interest.

Very frequently, a common pattern can be recognized in the development of different investigations within the general area of inequalities. Some mathematician discovers a particular inequality—either by a fluke, or from love of the game, or (as likely as not) because he needs a special tool for a specific purpose, perhaps in geometry or in theoretical physics. Other mathematicians examine the result and find that it opens to them fresh vistas: they may be able to 'improve' in one way or another on the original inequality, they may generalize it, they may produce variants or analogues. In this manner, every one of the 'classical' inequalities has come to support a substantial superstructure.

In the present article we shall seek to illustrate this pattern by a simple case study involving one of the 'classical' inequalities. We shall present the proof of an inequality found in 1882 by the eminent Russian mathematician P. L. Chebycheff (1821–1894) and shall then indicate a number of possible extensions and variants. The moral of the story (as of all such stories) is that there is more in the original result than meets the eye. It may be of interest to add that Chebycheff was a professor at the University of St Petersburg (now Leningrad) and that he made important contributions to the theory of approximations, the theory of numbers, mechanics and, above all, mathematical probability. In fact, the inequality we propose to discuss was discovered in the course of research into probabilistic questions.

Let us consider two sequences of n (real) numbers each, say $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$. From them, we can construct a further sequence $a_1b_1, a_2b_2, \ldots, a_nb_n$;

and we shall be interested in the average values, or 'means' of the three sequences, i.e. in the numbers

$$A = \frac{1}{n}(a_1 + a_2 + \cdots + a_n), \qquad B = \frac{1}{n}(b_1 + b_2 + \cdots + b_n),$$

$$M = \frac{1}{n}(a_1b_1 + a_2b_2 + \cdots + a_nb_n).$$

We now ask: is it true that, for *all* choices of a's and b's, $AB \leqslant M$? Or, perhaps, is it true that, for *all* choices of a's and b's, $AB \geqslant M$? The answer in either case is 'No' as we can easily† see by choosing (for any given n) two sets of values of a's and b's such that for one of them $AB > M$ while for the other $AB < M$. Moreover, the answer to our question still remains 'No' even if we concern ourselves with positive numbers only.

So far, then, our search has been inconclusive. To obtain a valid result of the desired kind, we shall make some further assumptions about the a's and b's; we shall, in fact, suppose that both sequences are 'increasing', i.e.

$$a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n, \tag{2a}$$
$$b_1 \leqslant b_2 \leqslant \cdots \leqslant b_n. \tag{2b}$$

In that case the inequality $AB \leqslant M$ *is* valid; and we shall embody this result in a formal statement.

Main Theorem. If the relations (2a) *and* (2b) *are satisfied by the real numbers* $a_1, \ldots, a_n, b_1, \ldots, b_n$, *then*

$$(a_1 + a_2 + \cdots + a_n)(b_1 + b_2 + \cdots + b_n) \leqslant n(a_1b_1 + a_2b_2 + \cdots + a_nb_n). \tag{3}$$

This is the original 'Chebycheff inequality', and it should be noted that it asserts a relation between *averages* since (3) can be written in the form

$$\frac{1}{n}(a_1 + a_2 + \cdots + a_n) \cdot \frac{1}{n}(b_1 + b_2 + \cdots + b_n) \leqslant \frac{1}{n}(a_1b_1 + a_2b_2 + \cdots + a_nb_n).$$

A similar remark applies in a number of instances below and we shall not repeat it each time.

The proof of the Main Theorem is easier than one might, perhaps, expect. In view of (2a) and (2b), we have

$$(a_k - a_j)(b_k - b_j) \geqslant 0, \tag{4}$$

i.e.

$$a_kb_k + a_jb_j - a_kb_j - a_jb_k \geqslant 0$$

whenever $1 \leqslant k, j \leqslant n$. Summing this last inequality for all k, j in the stated range,

† A judicious reader should never rest content until he satisfies himself that, in using terms such as 'easy' or 'easily', the author is not bluffing.

we obtain

$$\sum_{k,j=1}^{n} a_k b_k + \sum_{k,j=1}^{n} a_j b_j - \sum_{k,j=1}^{n} a_k b_j - \sum_{k,j=1}^{n} a_j b_k \geqslant 0.$$

This can be written as†

$$\Big(\sum_{k=1}^{n} a_k b_k\Big)\Big(\sum_{j=1}^{n} 1\Big) + \Big(\sum_{k=1}^{n} 1\Big)\Big(\sum_{j=1}^{n} a_j b_j\Big) - \Big(\sum_{k=1}^{n} a_k\Big)\Big(\sum_{j=1}^{n} b_j\Big) - \Big(\sum_{k=1}^{n} b_k\Big)\Big(\sum_{j=1}^{n} a_j\Big) \geqslant 0,$$

i.e.

$$2n \sum_{k=1}^{n} a_k b_k - 2\Big(\sum_{k=1}^{n} a_k\Big)\Big(\sum_{k=1}^{n} b_k\Big) \geqslant 0;$$

and this is, of course, equivalent to (3).

Having now established the Main Theorem, we supplement it with a number of observations. In the first place, instead of considering increasing sequences, we can just as well consider decreasing ones, i.e. if (2a), (2b) are replaced by the inequalities

$$a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n, \qquad b_1 \geqslant b_2 \geqslant \cdots \geqslant b_n,$$

then (4) continues to hold, the argument above remains unchanged, and conclusion (3) retains its validity. (Indeed, we recognize more generally that for (3) to hold it suffices that the a's and b's should be 'similarly ordered'.)

Again, suppose that we replace (2a) and (2b) by the conditions

$$a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n, \qquad b_1 \geqslant b_2 \geqslant \cdots \geqslant b_n. \tag{5}$$

Does the inequality (3) continue to hold in this case? It does not, but it is easy to see what happens. In place of (4), we now have

$$(a_k - a_j)(b_k - b_j) \leqslant 0$$

whenever $1 \leqslant k, j \leqslant n$, and the earlier argument still applies but with all signs of inequality reversed. We conclude, therefore, that, if relations (5) are given, then

$$(a_1 + a_2 + \cdots + a_n)(b_1 + b_2 + \cdots + b_n) \geqslant n(a_1 b_1 + a_2 b_2 + \cdots + a_n b_n).$$

(An alternative method is to write $b'_k = -b_k$ for $1 \leqslant k \leqslant n$ and to apply the Main Theorem to the sequences $a_1, a_2, \ldots, a_n$ and $b'_1, b'_2, \ldots, b'_n$.)

We shall next glance at a related result. Let $p_1, p_2, \ldots, p_n$ be arbitrary *positive* numbers. Then an argument only insignificantly more complex than that used to

† An expression such as $\Sigma_{j=1}^{n} 1$ may possibly be unfamiliar to some readers, but it stands simply for $\Sigma_{j=1}^{n} c_j$, where $c_1 = c_2 = \cdots = c_n = 1$, and so has the value n. We also take the opportunity to recall that the 'label' used for the summation suffix is of no importance; thus, for example, $\Sigma_{j=1}^{n} a_j = \Sigma_{k=1}^{n} a_k$.

establish the Main Theorem leads to the conclusion that, for real numbers subject to (2a) and (2b), we have

$$(p_1a_1 + p_2a_2 + \cdots + p_na_n)(p_1b_1 + p_2b_2 + \cdots + p_nb_n) \leqslant (p_1 + p_2 + \cdots + p_n)(p_1a_1b_1 + p_2a_2b_2 + \cdots + p_na_nb_n). \quad (6)$$

(The reader is encouraged to make sure that he is still following by writing out a full proof of this result.) Our new conclusion that (2a) and (2b) imply (6) is a 'generalization' of the Main Theorem since the latter result can be obtained by a special choice of the values of $p_1, p_2, \ldots, p_n$ (namely $p_1 = p_2 = \cdots = p_n = 1$).

What we have just done is to improve the Main Theorem by deriving a stronger (or, to be precise, a more general) conclusion from the original hypotheses (2a) and (2b). We shall now, by way of contrast, mention another and very recent improvement of a different type: here the conclusion will remain unchanged while the hypothesis is weakened. Suppose, then, that in place of (2a) and (2b), we have

$$a_1 \leqslant \frac{1}{2}(a_1 + a_2) \leqslant \frac{1}{3}(a_1 + a_2 + a_3) \leqslant \cdots \leqslant \frac{1}{n}(a_1 + a_2 + \cdots + a_n), \quad (7a)$$

$$b_1 \leqslant \frac{1}{2}(b_1 + b_2) \leqslant \frac{1}{3}(b_1 + b_2 + b_3) \leqslant \cdots \leqslant \frac{1}{n}(b_1 + b_2 + \cdots + b_n). \quad (7b)$$

We might now say that the a's and b's 'increase on the average'. Before we go any further, however, let us be clear about the precise relation between (2a) and (7a) (or, equivalently, between (2b) and (7b)). The answer is very simple: if the a's satisfy (2a), they will certainly satisfy (7a); but the converse is not necessarily true, i.e. it is possible to choose a's—and let the reader not fail to do so—which satisfy (7a) but fail to satisfy (2a). In short, by postulating (7a) in place of (2a), we are imposing a weakened constraint on the a's (and *mutatis mutandis* for the b's).

We are now able to state our next result. It is to the effect that, for (3) to hold, we do not need to postulate (2a) and (2b): the *weaker* hypotheses (7a) and (7b) will suffice. Since, then, a weaker hypothesis is seen to lead to the same conclusion, we have now a *stronger* theorem. The proof of this strengthened form of the Main Theorem is not quite as obvious as that of the original version; but although it is not hard, we shall absolve the reader from its scrutiny. (See, however, the list of proposed problems at the end of the article.)

So far our discussion has been concerned with *two* sequences. It is, however, possible to formulate analogues of the Main Theorem involving a larger number; and the case of three sequences represents the general situation with sufficient generality. Let us, then, suppose that

$$a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n, \qquad b_1 \leqslant b_2 \leqslant \cdots \leqslant b_n, \qquad c_1 \leqslant c_2 \leqslant \cdots \leqslant c_n. \quad (8)$$

Is the inequality

$$(a_1 + a_2 + \cdots + a_n)(b_1 + b_2 + \cdots + b_n)(c_1 + c_2 + \cdots + c_n) \leqslant n^2(a_1b_1c_1 + a_2b_2c_2 + \cdots + a_nb_nc_n), \quad (9)$$

which is analogous to (3), then necessarily true? Not quite, and the reader is invited to construct a suitable numerical example (say for $n = 4$) which demonstrates that the desired conclusion does not follow. However, it is easy to retrieve the position by insisting that the a's, b's and c's should be not merely real but, in fact, non-negative. In that case (8) does, indeed, imply (9); and the proof is obtained readily by two successive applications of the Main Theorem. Further—and this is not at all surprising—if all inequalities in (8) are reversed, then, for non-negative a's, b's and c's, (9) still remains valid.

We shall next present a variant of a rather different kind, which is obtained by a transition from the 'discrete' to the 'continuous' case. A definite integral is, as we recall, the limiting value of a sum: it therefore behaves, in some respects, not unlike a sum. Perhaps it is even more appropriate to view the expression

$$\frac{1}{\beta - \alpha}\int_\alpha^\beta f(x)\,dx$$

(where $\alpha < \beta$) as an 'average' of the values of the function $f(x)$ in the range $\alpha \leqslant x \leqslant \beta$. Consequently, inequalities relating to averages of sequences might be expected to possess analogues in which integrals make their appearance. This is certainly so in the present case.

Let, then, $f(x)$, $g(x)$ be two increasing functions:† this postulate corresponds to (2a) and (2b) (and it is easily seen that two decreasing functions would do equally well). With this assumption, we have

$$\left(\int_\alpha^\beta f(x)\,dx\right)\left(\int_\alpha^\beta g(x)\,dx\right) \leqslant (\beta - \alpha)\int_\alpha^\beta f(x)g(x)\,dx, \tag{10}$$

an inequality which looks like and, indeed, corresponds to, (3). As for the proof of (10), this depends on nothing worse than an easy adaptation of the argument leading to (3), with integration now taking the place of summation. The integral inequality just stated is, once again, a generalization of the Main Theorem since, by a special choice of α, β and the functions $f(x)$, $g(x)$, we are actually able to recover that result. We propose to leave to the reader the task of finding out for himself how this may be done in detail and merely advise him to look at functions whose graphs consist of sets of straight segments parallel to the x-axis.

A reader with a taste for exploration will find plenty of scope in the preceding pages. We have continually challenged him to verify statements made but left unproved in the text, and it is easy to formulate many further questions closely related to but going beyond our discussion. Here is a short list of such questions.

(i) Under what circumstances can the sign $\leqslant$ in the Main Theorem be replaced by the sign $<$ of strict inequality?
(ii) State and prove an analogue for three sequences of the inequality (6).
(iii) Generalize the inequality (9) for the case of m ($\geqslant 3$) sequences.

† We say that a function $f(x)$ is increasing if $f(x_1) \leqslant f(x_2)$ whenever $x_1 < x_2$.

(iv) Let $a_1, a_2, \ldots, a_n$ be any positive numbers. Show, both directly and by means of the Main Theorem, that

$$(a_1 + a_2 + \cdots + a_n)\left(\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}\right) \geqslant n^2.$$

(v) Making use of the relation

$$\{f(x) - f(u)\}\{g(x) - g(u)\} \geqslant 0 \qquad (\alpha \leqslant x, u \leqslant \beta),$$

or otherwise, give a proof of the theorem asserting the validity of (10).

(vi) Formulate and prove an integral analogue of the inequality (6). Exhibit both (6) and (10) as special cases of this result.

(vii) Formulate and prove an integral analogue of the inequality (9).

(viii) Show that, if the real numbers $a_1, \ldots, a_n, b_1, \ldots, b_n$ satisfy (7a) and (7b), then the inequality (3) is valid. Also show that the converse inference is false.

(ix) Can the inequality (3) be established under a hypothesis weaker than that given by (7a) and (7b)?

To play the game according to the Queensberry rules, we ought to disclose the status of these questions. The answers to (i)–(vii) are well known and the solution of the problems calls for an understanding of the ideas discussed earlier but for no ingenuity. The proof of (viii) is easy enough given the 'right' idea; but to spot the right idea may require an imaginative effort. As regards (ix), here the situation is quite different for we do not even know whether a reasonable answer exists at all. This question therefore poses a genuine challenge.

Bias on Ballot Papers, or the Good Fortune of Basil Brush

G. J. G. UPTON
University of Essex

1. Introduction

Have you ever used the yellow pages in a telephone directory? Perhaps you needed to find a local steeplejack to mend your roof? You probably found the steeplejack section of the directory, and starting at its beginning, looked through the list of names until you arrived at one living in your area. Now, if everyone requiring a steeplejack were to use this method of selection, then, since names are ordered alphabetically, steeplejacks with names at the beginning of the alphabet would be likely to get more trade than those with names towards the end of the alphabet. That the yellow pages *are* used in this way has not passed unnoticed by tradesmen, who occasionally go to great lengths to devise a name near the beginning of the alphabet for their businesses. In my directory, the Taxis section starts with

A1 Taxis,
AAB Minicabs,
ABC Taxis, . . .

Although this alphabetic bias is clearly recognised by commercial firms, I do not believe any attempt has ever been made to remove the effect of such bias in British elections. Presumably, the effect has been judged to be too small to justify the expense involved in removing it; however, the results given later suggest that the importance of the bias may have been under-rated.

An interesting red herring: Harold Wilson was the first Prime Minister since 1902 to have a surname starting with a letter from the second half of the alphabet![†]

The remainder of this article will outline some results concerning various elections; these have been obtained by my colleague David Brook and myself. For the most part the analysis consists of simply adding up numbers and comparing totals. The difficulty lies in choosing sensible sets of numbers to add up; the resolution of this difficulty demands the application of common sense as much as statistics.

2. Positional bias in British General Elections

I shall start with General Elections because these are the best known. Unfortunately, no definite results have been obtained for them.

There are 635 seats in the British House of Commons, and each seat is contested in the General Election by candidates from a number of different political parties. On the ballot paper for a particular constituency the surnames of each candidate are placed in alphabetical order. Current practice is to include the political party to which the candidate belongs in brackets after his/her name; since most of those who vote are voting for a party rather than for an individual, there would seem to be little chance that the alphabetic effect will bias the results. The General Election which we studied (1964) was held before the introduction of the party affiliation on the ballot paper.

How can we test the hypothesis that the ballot position does not affect the results? A possible approach would be to total the votes gained by all candidates whose names appeared in first (second, third, etc.) position on the ballot paper over all 635 constituencies. However, this does not work, since there may only be one constituency with, say, eleven competing candidates, and none with more, in which case the votes for the 11th position on the ballot will be due to this constituency alone. To remove this difficulty we looked at constituencies in which the same number of candidates stood. Since a Labour candidate would probably get more votes if his opponent were an Independent rather than a Conservative, it is sensible

† This is not a fair statistic since surnames are not evenly distributed across the letters of the alphabet.

to restrict attention to particular combinations of party representation. For example, in the 1964 General Election, there were 187 constituencies in which only Labour and Conservative candidates stood (Table 1).

TABLE 1. Total votes in 187 two-candidate constituencies subdivided by position on ballot paper

First position	3,840,335	Second position	3,881,453

This table contradicts the previous impression that the first position would be the more favoured. The explanation is that we have totalled the wrong numbers! In 1964 Labour won the election: in the 187 constituencies studied above, Labour candidates, by chance, occupied first position on only 80 occasions. Thus Table 1 confuses the effects of position and party. To remove the confusion we calculate the average votes for each party/position combination separately (Table 2).

TABLE 2. Average votes in two-candidate constituencies subdivided by position on ballot paper and by party

	Conservative	Labour
First position	17,843	24,140
Second position	17,390	23,274

The positional effect now appears to favour the top of the ballot. However, the figures in Table 2 are averages only. The individual numbers of votes vary so greatly that (reluctantly) we can ascribe no statistical significance to the results.

To avoid the enormous variations which arise in the study of votes, we can simply look at the relative success of the candidates—did they get elected more often when they occupied the first position on the ballot? Table 3 gives the results for the 306 constituencies in the 1964 election in which the candidates belonged to the Conservative, Labour and Liberal parties.

TABLE 3. Success of candidates in three-candidate constituencies subdivided by position on ballot paper and by party

	Position		
	1	2	3
Conservative:			
Elected	74	77	58
Not elected	40	25	32
Total	114	102	90
Labour:			
Elected	22	38	32
Not elected	65	69	80
Total	87	107	112

The number of elected candidates in each position can be expressed as a percentage of the total number of party candidates in that position. The figures for the Conservatives are 64 per cent in first position, 75 per cent in second position and 64 per cent in third position. The comparable figures for the Labour party are 34 per cent, 36 per cent and 29 per cent. Although, for both parties, the second position is the most favoured, a simple χ^2 test reveals that the differences may be attributed to random variation.

3. Positional bias in British Local Government Elections

My wife sometimes claims that I do nothing at work except drink coffee; I maintain that the coffee stimulates learned discussion. It was while reading the local paper at one such coffee-drinking session in May 1973 that I first encountered the phenomenon of alphabetic positional bias. The paper contained the results of the elections to the membership of the then newly formed Local District Councils. In these elections, local areas were subdivided into small wards, each with an electorate of not more than 5000 people. Because the Councils were totally new, their whole membership had to be elected at once. Thus, in most wards, there were three seats (in some as many as eight), for which each political party put forward a corresponding number of candidates, and each voter had a corresponding number of votes to distribute amongst the rather long, alphabetically ordered, list of candidates.

Naturally, most voters voted *en bloc* for their own party candidates. Some voters, with no strong party affiliation, voted for candidates they knew. Others, unused to having more than one vote, did not use all their votes. A few accidentally voted for the wrong candidates. The combination of all these different patterns of voting behaviour produced the final election results. Typically, all the candidates from a particular party obtained roughly the same number of votes, but in none of the 1000 party groups studied did we find the 3 party members obtaining exactly the same number of votes (in about 5 per cent, 2 out of the 3 tied), and the differences that occurred provided significant evidence of alphabetic positional bias.

To illustrate the effect of the bias, we need a system of referencing the individual candidates. Consider a particular party fielding three candidates: labelled *a*, *b*, *c* in the order in which they appear on the ballot paper. Then a particular outcome, in terms of number of votes received, can be expressed by a permutation such as *bac* meaning that *b* received most votes and *c* received least.

There are 6 possible permutations ranging from *abc*, in which the outcome was in exact alphabetic order, to *cba* in which the reverse occurred. The results of the survey are given in Table 4.

TABLE 4. Observed frequencies of the six outcomes in three-seat wards

Permutation	*abc*	*acb*	*bac*	*bca*	*cab*	*cba*	Total
Observed frequency	232	136	174	151	114	141	948
Expected frequency	158	158	158	158	158	158	948

If there were no bias we would expect each permutation to occur about $948/6 = 158$ times, with the individual frequencies all lying in a range between 130 and 190. This clearly has not occurred and a simple χ^2 test shows that the probability of the observed frequencies having arisen by chance is less than 1 in 200.

The results in Table 4 prove conclusively† that the positional bias existed and affected the election results, but they do not explain exactly how, mathematically, the bias occurred.

Before developing the analysis it would be as well to make clear the political importance of these results. In the May 1973 elections, the candidate obtaining the most votes was elected to serve on the Council for 4 years, the next two most successful candidates serving for 2 years and 1 year respectively. Studying individual constituencies we found that it was likely that in as many as *one ward in five* the order of election, and consequently the possible composition of future councils, had been affected by the positional bias. For example, it is easy to see from Table 4 that *a* got most votes on 368 occasions, *b* on 325 and *c* on 255, compared to the expected figure of 316 for each, in the absence of bias. Furthermore, in about 2 per cent of the wards it appeared likely that there were persons elected to the council who might not have been elected had their surname begun with a letter lower down in the alphabet.

4. Which position is most preferred?

We originally thought that the bias worked in the way that the taxi companies evidently believe, and which we suggested in the introduction. Thus if the candidates of a particular party occupied positions 1, 2 and 5, say, on a 6-candidate ballot paper, then we expected that 1 would be preferred to 2, and 2 to 5. However, it turns out that this is not the case, as we shall see.

The puzzling aspect of Table 4 is not the enormous frequency for *abc* but the relatively large frequency of *cba*, which, if the original theory were correct, would be by far the least common of all the permutations. We were tempted to attribute it to the perverseness of voters who habitually read lists of names from bottom to top, and thereby reversed the usual positional bias. There may indeed be some voters of this kind about—in which case I apologise for the adjective perverse. But there is a simpler explanation, namely that on a 6-position ballot paper, the order of preference of the positions is not 1, 2, 3, 4, 5, 6 but 2, 3, 4, 1, 5, 6.

To be able to detect this, we need to reference the candidates in some other way and not only by their positions on the ballot paper. We have a two-way table with the rows referring to the subset of ballot positions held by the party group, and the columns referring to the within party permutation determined by the votes they received. We restrict our attention to the 149 pairs of party groups contesting seats in which just six names appeared on the ballot. The results appear in Table 5.

† For those unconvinced by Table 4, I should add that subdivision by political party or by area of the country produces tables showing precisely the same pattern of frequencies.

TABLE 5. Observed frequencies of within-party orderings for ballot papers of length 6

Within-party positions on ballot paper	Within-party ordering *abc*	*acb*	*bac*	*bca*	*cab*	*cba*
123	4	1	1	3	2	1
124	2	0	1	0	3	4
125	4	2	7	4	1	3
126	1	3	5	3	0	5
134	1	2	2	1	3	1
135	4	1	4	2	3	2
136	1	5	2	1	2	4
145	1	2	4	4	0	4
146	5	3	2	4	0	3
156	8	1	0	3	0	4
234	3	3	3	1	2	4
235	7	2	4	0	2	2
236	3	3	3	3	2	1
245	6	3	3	1	2	0
246	4	5	4	0	2	1
256	1	3	1	3	1	1
345	5	3	2	3	3	1
346	14	1	1	2	1	2
356	4	1	3	0	2	0
456	4	1	1	1	3	2
Total	82	45	53	39	34	45

The following example may help to clarify the table entries. Suppose in one ward the Labour candidates are named Mr A, Mr B and Mr C, occupying positions 1, 2, 3 respectively on the ballot paper, while the Conservatives are Mr R, Mr S and Mr T in positions 4, 5, 6 respectively. If, in respect of votes, $A > B > C$, and $T > S > R$, then these contribute 1 to the top left and bottom right cells in the table.

How can we pick out some meaningful figures from the 120 entries in the body of the table? We have claimed that each position is preferred to the one below it, except for position 1, so a quick check could be provided by collapsing Table 5 into two rows, one involving party groups of which position 1 was a member, the other involving groups which do not include 1 (Table 6).

TABLE 6. Collapsed version of Table 5

	Within-party ordering *abc*	*acb*	*bac*	*bca*	*cab*	*cba*
Groups including position 1	31	20	28	25	14	31
Groups not including position 1	51	25	25	14	20	14

The asymmetric split of the frequencies of the permutations *abc* and *cba* is apparent from the table, but no clear idea of the overall ordering can be gained.

Elsewhere (reference 1), David Brook and myself have published details of the results obtained by using a rather complicated statistical model. The results that arise from our model suggested the order of preference 2, 3, 4, 1, 5, 6; the same order is also suggested by the following simple technique. Assign a score of 1 to a person who does best within his party, and scores of 2 and 3 to the next two positions. Then for each ballot paper position compute the average score. The results for the data of Table 5 are given in Table 7.

TABLE 7. Average score of ballot paper positions

Position	1	2	3	4	5	6
Average score	2·03	1·74	1·91	1·92	2·17	2·21

Thus it is Basil Brush in position 2 who is greatly advantaged by the positional bias, whereas Aaron Aardvark in position 1 does rather worse than might be expected by chance.

Much the same pattern occurs with ballot papers of greater length; position 1 is usually found to be slightly worse than position 4, with the remainder favoured in their natural order. We have insufficient data to draw any firm conclusions for ballot papers of other lengths.

5. Voting bias in other countries

In Australian General Elections, voting is compulsory: this means that those who would otherwise not vote cast their votes more or less at random. It is well known there that there is an advantage in having a name towards the top of the ballot paper. Indeed Australians call this type of vote a 'donkey' vote, and occasionally candidates change their surnames by deed poll in order to gain some advantage from it.

Changing surnames to gain such an advantage is not a pastime confined to Australia. In 1970 a little-known candidate for Edward Heath's Bexley seat changed his name to Heath, doubtless hoping that he might be advantaged by voters accidentally voting for the wrong Heath. The phenomenon is not of recent origin: as long ago as 1910, Woodrow Wilson mentioned that, in the United States, election candidates changed their names for similar reasons.

In the United States, local elections occur on a grand scale. Sometimes the number of candidates is very large indeed: Mueller (reference 2) has reported an analysis of the voting patterns in an election involving 133 candidates where each voter had 7 votes. In this election the ballot 'paper' consisted of 7 sheets with a maximum of 20 names per sheet. The voting pattern revealed that the name at the bottom of each sheet did rather better than the names immediately above it: there was a simple explanation for this—voters who had assigned 6 of their 7 votes, would sometimes give their remaining vote to the last name on the sheet, rather than turning over to the next! In that election, there was a considerable advantage in appearing on the first sheet.

6. Counteracting the positional bias

In the United States, voting machines are sometimes used in place of ballot papers. The voter pulls one from several rows of levers corresponding to his particular choice of candidate. Bain and Hecock (reference 3) have conducted extensive studies on these machines and have found that much the same sort of bias occurs. Each row is favoured above the rows below it, and, on each row the left-hand end is favoured over the right. Thus voting machines do not seem to be the answer.

One possibility would be to arrange that every possible permutation of the names of the candidates should appear equally often on the ballot papers. However the number of permutations increases very quickly: with 12 candidates there would be more than 479 *million* different permutations.

A simpler way is to cycle the names on the ballot paper so that each name occurs in each position once. This would remove the direct positional bias and would require only the same number of different types of ballot papers as there are candidates. Such a remedy would be readily workable: it remains to be seen if we care enough about our elections to implement it.

References

1. G. J. G. Upton and D. Brook, The determination of the optimum position on a ballot paper. *Appl. Statist.* **24** (1975), 279–287.
2. J. E. Mueller, Choosing among 133 candidates. *Public Opinion Quart.* **34** (1970), 395–402.
3. H. M. Bain and D. S. Hecock, *Ballot Position and Voter's Choice* (Wayne State University Studies, 1957).

Collisions of Atomic Particles

D. S. SCHONLAND
University of Southampton

1. Introduction

In most mechanics textbooks the treatment of elastic collisions is restricted to showing how the principles of conservation of momentum and energy are used in the case of a head-on collision between two solid bodies. These principles can, however, be applied to discuss a more general type of collision process of the kind illustrated in Figure 1.

Here an incident particle of mass M and velocity V is projected so as to pass close by a target particle of mass m that is initially at rest. We suppose that there is a repulsive force between the particles which is zero or very small when the particles are far apart, but which becomes large when the particles are close together. This force sets the target particle into motion and changes the original motion of the incident particle. After the 'collision' when the particles are again far apart, their velocities will tend to constant values because the force between them becomes zero or very small. In Figure 1 the final velocity of the incident particle is shown as $\boldsymbol{v}$ in a direction making an angle ϕ (the angle of scattering) with its initial velocity V and the final velocity of the target particle is $\boldsymbol{u}$ in a direction specified by the angle θ (the angle of recoil). An example of the kind of force being considered is the electrostatic inverse square repulsion between two positively-charged particles.

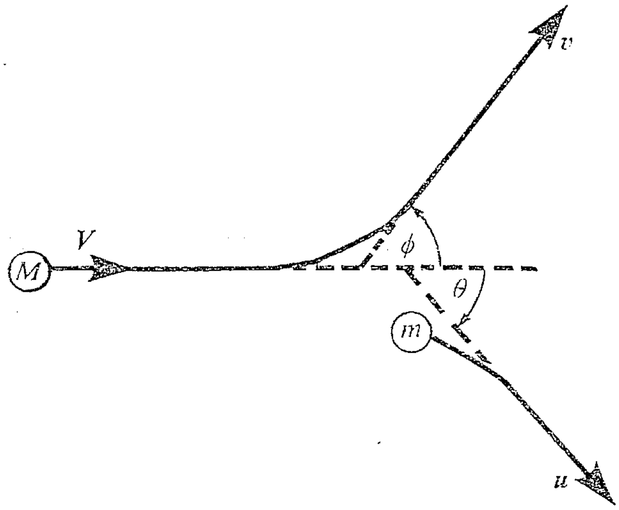


Figure 1. A collision between an incident particle of mass M and a target particle of mass m.

The actual values of the speeds $\boldsymbol{u}$ and $\boldsymbol{v}$ and the angles θ and ϕ will depend on the details of the force, but if we assume the general dynamical principles of conservation of momentum and of energy we can obtain certain relations between $\boldsymbol{u}$, $\boldsymbol{v}$, θ and ϕ that must hold independently of the precise nature of the force. These relations were first discussed by Rutherford and his colleagues in 1911 when they were investigating atomic structure by detailed measurements of the scattering of α-particles (the positively-charged nuclei of helium atoms) by heavy atoms.

2. Conservation of momentum and energy

The final total momentum vector of the two particles must be the same in magnitude and direction as the initial total momentum vector. Thus the vectors representing the various momenta involved must be related by the triangle shown

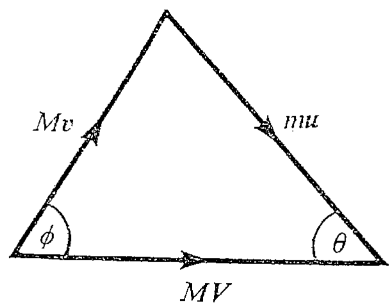


Figure 2. The relation between the momentum vectors.

in Figure 2. Resolving the vectors in directions respectively parallel and perpendicular to the initial velocity V gives the relations

$$Mv\cos\phi + mu\cos\theta = MV, \tag{1}$$

$$Mv\sin\phi = mu\sin\theta. \tag{2}$$

The sum of the kinetic energies of the particles and the potential energy due to the force between them remains constant. The potential energy will have the same value at the beginning and the end of the motion when the particles are far apart and the force between them tends to zero. Consequently, the total kinetic energy has the same value at the beginning and the end of the motion and

$$Mv^2 + mu^2 = MV^2. \tag{3}$$

This is the condition that defines an elastic collision.

3. Consequences of the conservation laws

For given values of M, m and V, the above equations give three relations that must hold between the four quantities u, v, θ and ϕ. We shall use the equations to express the final speeds u and v and the scattering angle ϕ in terms of the recoil angle θ, since this gives results whose significance can be readily appreciated.

Directly from the triangle of Figure 2, or by manipulation of the equivalent equations (1) and (2), we have

$$M^2v^2 = M^2V^2 + m^2u^2 - 2mMuV\cos\theta.$$

This gives an expression for v^2 that can be substituted into (3). Simplification of the resultant equation gives

$$u = \frac{2M}{M+m}V\cos\theta. \tag{4}$$

We now use (3) to obtain the expression for v in terms of θ:

$$v^2 = V^2 - \frac{m}{M}u^2 = V^2\left(1 - \frac{4mM}{(M+m)^2}\cos^2\theta\right). \tag{5}$$

To find the relation between ϕ and θ we return to the triangle of Figure 2, or equations (1) and (2), to obtain

$$\tan\phi = \frac{mu\sin\theta}{MV - mu\cos\theta}.$$

Substitution of the expression (4) for u then gives

$$\tan\phi = \frac{2m\sin\theta\cos\theta}{M + m - 2m\cos^2\theta}$$

which, by use of the relations

$$2\sin\theta\cos\theta = \sin 2\theta, \qquad 2\cos^2\theta - 1 = \cos 2\theta,$$

can be written as

$$\tan\phi = \frac{m\sin 2\theta}{M - m\cos 2\theta}. \tag{6}$$

4. Discussion

We begin the discussion by considering two extreme cases. The first case is that of a 'head-on' collision in which the incident particle is projected directly towards the target particle. Then the target particle will be propelled directly forwards with a recoil angle $\theta = 0$, while the incident particle may either continue its forward motion (scattering angle $\phi = 0$) or have its direction of motion completely reversed (scattering angle $\phi = \pi$).

The other extreme case is when the incident particle is projected so as to pass the target particle at a great distance. Then the interaction between the particles will have minimal effect; the final speed v of the incident particle will be nearly as great as its initial speed V and the final speed u of the target particle will be small. From (4) we see that this requires $\cos\theta$ to be small so that θ must be close to $\pi/2$. (This can also be seen from Figure 2: if the scattering angle ϕ is small and if v is only slightly smaller than V, the recoil angle θ must be nearly a right angle.)

In general, the conditions of the collision will lie between these two extremes. Thus the target particle will recoil at an angle lying in the range $0 \leqslant \theta \leqslant \pi/2$, while its final speed as given by (4) will have a value between $2MV/(M + m)$, corresponding to a head-on collision with $\theta = 0$, and zero, corresponding to a very distant collision where $\theta = \pi/2$.

Now let us examine the behaviour of the incident particle as θ increases from zero to $\pi/2$. This behaviour depends crucially on the masses of the particles and we can distinguish three different cases.

(i) $M < m$

We turn (6) upside down to obtain

$$\cot\phi = \frac{M - m\cos 2\theta}{m\sin 2\theta}.$$

Since $M < m$, the numerator varies from the negative value $M - m$ when $\theta = 0$ to the positive value $M + m$ when $\theta = \pi/2$. The denominator vanishes when $\theta = 0$ or $\pi/2$ and is positive for intermediate values of θ. It follows that, as θ increases from zero to $\pi/2$, the value of $\cot\phi$ increases from $-\infty$ to $+\infty$ and so ϕ itself must decrease from π to zero.

Thus a head-on collision with $\theta = 0$ gives a scattering angle $\phi = \pi$ which corresponds to a direct reversal of the direction of motion of the incident particle. Putting $\theta = 0$ and $\phi = \pi$ into (1) and (4) gives

$$Mv = mu - MV \quad\text{and}\quad u = \frac{2M}{M + m}V$$

so that

$$v = \frac{m - M}{m + M}V.$$

Thus for a head-on collision we may note that the final relative speed of the two particles, which in this case is $u + v$, is the same as the initial relative speed V (i.e. the collision is 'elastic').

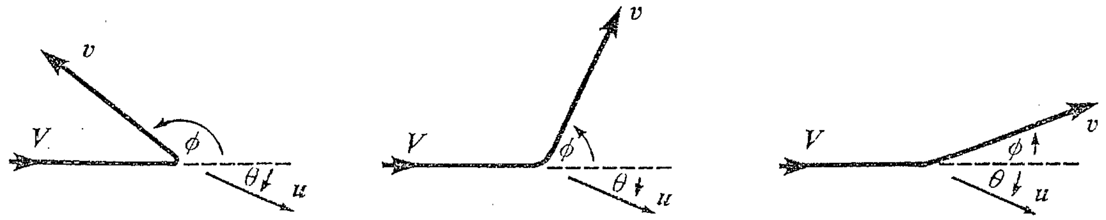


Figure 3. Typical collision paths for the cases: (a) $M < m$, (b) $M = m$, (c) $M > m$.

For a nearly head-on collision, when θ is small, the scattering angle will be fairly close to π and the paths of the particles will look like those shown in Figure 3(a).

(ii) $M = m$
Here (6) simplifies to

$$\tan \phi = \frac{\sin 2\theta}{1 - \cos 2\theta} = \cot \theta.$$

Hence in a collision between two particles of equal mass we have $\theta + \phi = \pi/2$ and the paths of the scattered and recoil particles are always at right angles to each other, as shown in Figure 3(b).

When $M = m$, equations (4) and (5) become simply

$$u = V \cos \theta, \qquad v = V \sin \theta,$$

and we note that in a head-on collision, with $\theta = 0$, the target particle acquires a final speed V, while the incident particle is brought to rest. It is interesting that this result which is well known for the case of a direct impact between two solid bodies of equal mass (illustrated, for example, by Newton's cradle) in fact applies quite generally for any kind of repulsive force between the particles.

(iii) $M > m$
For this case the denominator of (6) is always positive and never vanishes, while the numerator vanishes for $\theta = 0$ and $\pi/2$ and has a positive maximum at $\theta = \pi/4$. Thus, as θ increases from zero to $\pi/2$, the value of $\tan \phi$ increases from zero to a maximum positive value and then decreases again to zero. The scattering angle ϕ correspondingly increases from zero (for a head-on collision with $\theta = 0$) to some maximum value that is less than $\pi/2$ and then decreases again to zero (for a very distant collision with $\theta = \pi/2$). Typical collision paths for this case are shown in Figure 3(c).

In a head-on collision we have $\theta = 0$ and $\phi = 0$. In this case the incident particle continues to move in its original direction. Putting $\theta = 0$ and $\phi = 0$ in (1) and (4) shows that the final speeds of the two particles will be

$$v = \frac{M - m}{M + m} V, \qquad u = \frac{2M}{M + m} V,$$

so that their relative speed $\boldsymbol{u} - \boldsymbol{v}$ after the collision is the same as their initial relative speed V, just as in the cases previously considered.

As the conditions are changed from those of a head-on collision, we note that the same scattering angle ϕ can be obtained with two distinct values of θ. The smaller value of θ corresponds to a relatively close collision in which a fair amount of energy is transferred from the incident to the target particle; the larger value of θ corresponds to a more distant collision in which the incident particle loses a smaller amount of energy.

The maximum value of the scattering angle that can occur in any collision for which $M > m$ is found by differentiating (6) with respect to θ. The differentiation gives

$$\sec^2 \phi \frac{d\phi}{d\theta} = \frac{2m(M \cos 2\theta - m)}{(M - m \cos 2\theta)^2}$$

and we see that $d\phi/d\theta = 0$ and ϕ has its maximum value when $\cos 2\theta = m/M$. Then $\sin 2\theta = (1 - m^2/M^2)^{\frac{1}{2}}$ so that, from (6), the maximum value of ϕ is Φ where

$$\tan \Phi = \frac{m}{(M^2 - m^2)^{\frac{1}{2}}}$$

Since $0 < \Phi < \pi/2$, so that $\sin \Phi$ and $\cos \Phi$ are both positive, this relation can be written in the simpler form

$$\sin \Phi = m/M.$$

For example, when an α-particle of mass $M = 4$ atomic units collides with a proton of mass $m = 1$ atomic unit, the maximum possible value of the angle through which the α-particle can be scattered is

$$\Phi = \sin^{-1}(1/4) = 14° \, 29'.$$

5. Summary

In any collision event the total momentum of the two interacting particles remains constant and, if the collision is elastic, the total energy also remains constant. These two conservation laws enable one to obtain relations between the initial and final velocities of the two particles that will hold whatever the force between them may be. We have considered in particular the way in which the scattering and recoil angles depend on the masses of the particles.

These results are of historic importance since it was the experimental observation that α-particles fired at a thin gold foil could be scattered through very large angles, as in Figure 3(a), that led Rutherford to the hypothesis that the mass of an atom is largely concentrated in a small positively charged nucleus; a hypothesis that he was able to prove quantitatively by a detailed examination of the statistical distribution of the scattering angles.

Problems

1. Use (6) to show that

$$\frac{M}{m} = \frac{\sin(2\theta + \phi)}{\sin \phi}.$$

2. In a collision event observed in a photographic emulsion exposed to a beam of deuterons, the tracks of the particles emerging from the collision made angles of 23·5° and 52·5° with the track of the incident deuteron, but it was not possible to decide from the nature of the tracks which of them was made by the scattered deuteron and which by the recoil particle. Given that a deuteron has mass $M = 2$ atomic units and that the only particle of smaller mass that can leave a track in a photographic emulsion is a proton of mass $m = 1$ atomic units, use the result of Problem 1 to show that the recoil particle was a proton and that the recoil angle was 52·5°. (A deuteron is a particle formed by the combination of a proton and a neutron. It is the nucleus of deuterium, a naturally occurring isotope of hydrogen often referred to as 'heavy hydrogen'.)

3. Use (1) and (2), or the triangle of Figure 2, to express $\boldsymbol{u}^2$ in terms of $\boldsymbol{v}$ and ϕ, regarding M, m and V as given constants. Then combine this result with (3) to deduce that, for any given value of ϕ, the final speed of the scattered particle is given by

$$v = \frac{V}{M + m} \{M \cos \phi \pm (m^2 - M^2 \sin^2 \phi)^{\frac{1}{2}}\}.$$

Discuss the ambiguity in this solution, using the condition that v is real and positive, for the cases $M < m$ and $M > m$, showing that the following situations occur.

(i) If $M < m$, all values of ϕ are possible but we must choose the positive square root in the solution for v. In this case there is a unique relation between v and ϕ such that v decreases from V to $(m - M)V/(M + m)$ as ϕ increases from zero to π.

(ii) If $M > m$, the scattering angle must lie in the range $0 \leqslant \phi \leqslant \sin^{-1}(m/M)$ and, for any value of ϕ in this range, both solutions for v are possible. In this case there are two distinct collisions that produce the same scattering angle.

(Note. This problem shows an alternative approach to the results we have derived.)

Reference

C. F. Powell and G. P. S. Occhialini, *Nuclear physics in photographs: tracks of charged particles in photographic emulsions* (Oxford University Press, 1947).

This non-technical book reproduces interesting photographs of collision events between nuclear particles, and analyses some of them by the methods we have described.

String Figures

R. N. MADDISON
The Open University

Introduction

Between 1890 and 1910 several groups of anthropologists were interested in the patterns of string figures that various tribes around the world knew. Some of the anthropologists saw people making string figures and recorded the hand movements, but others recorded only the final picture, so that there are now perhaps 100 illustrations of figures for which the method of construction is not known. Since then little work on string figures has been done, but a natural question to ask is, 'Given a picture of a string figure, can one work out a likely method by which that string figure could originally have been made?' I have been interested in devising methods, using a computer, to help find the movements by which a particular given string figure may have been made.

Representation of a string figure

The first problem is to devise the method by which a picture, such as Figure 1, can be represented in a digital computer. For the moment I shall restrict the problem to representing figures where the string goes round the fingers and thumbs and the hands are extended fingers upwards as in the figure. Starting from a particular point on the string, such as the middle of the front string nearest the body, and taking an arbitrary direction to go round the figure, such as initially clockwise when looking down from above, we can see that the string first goes round the *left* thumb from

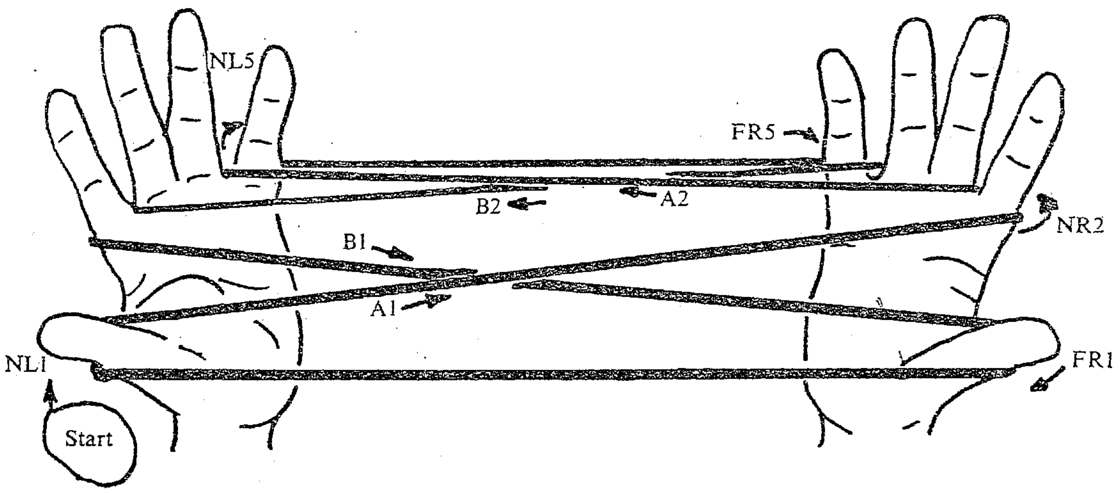


Figure 1. Opening *A*.

the near side, then crosses *above* another string, and so on. A simple code is used to represent this sequence. We write *N* if the string goes from the *near* side to the far side round the back of the finger, and *F* if the direction is from the *far* side to the near side. We use *L* and *R* for the hands, with the thumb as 1, the index finger as 2, the middle finger as 3, the ring finger as 4 and the little finger as 5. Where strings cross, when looking down from above, the string that is *above* is labelled *A* and given a number associated with the crossing point, and the string that is *below* is labelled *B* followed by the same number. Thus the string figure in Figure 1 can be represented by:

$$NL1\ A1\ NR2\ A2\ NL5\ FR5\ B2\ FL2\ B1\ FR1.$$

In this sequence the *order* of the groups of characters corresponds to going systematically along the string.

Making a string figure

A string figure is made by a sequence of hand movements. In each movement either a loop is taken up onto a finger, or a loop is dropped from a finger, or the figure is altered in some other way. For each possible movement we need a method of operating on the sequences of characters to generate a new sequence of characters corresponding to the new figure. The rules for transforming the characters are a little complicated, since they have to cover every possible case, but an example is given in Table 1 showing the movements and corresponding character strings for making Figure 1. That figure is well known as Opening *A*.

TABLE 1. Opening *A*

Hand movements	Representation of figure produced
Put string round little fingers.	*NL5 FR5*
Pick up near string onto thumbs. Extend, so string forms a rectangle with its short sides across the palms.	*NL1 NL5 FR5 FR1*
Pick up left palmar string onto right index finger. Extend.	*NL1 A1 NR2 A2 NL5 FR5 B2 B1 FR1*†
Similarly pick up right palmar string onto left index finger. This completes Opening *A*.	*NL1 A1 NR2 A2 NL5 FR5 B2 FL2 B1 FR1*

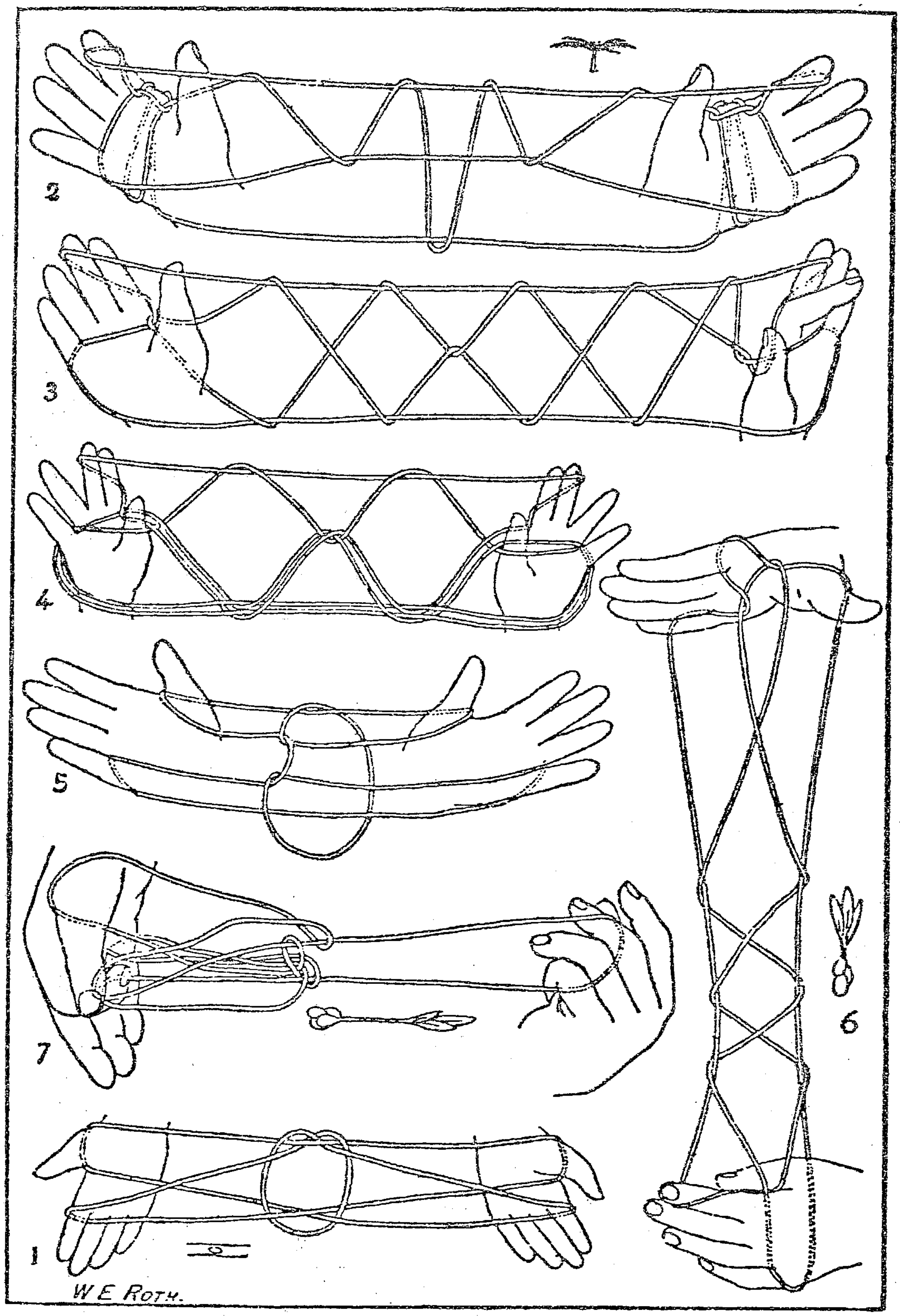
† Since the string goes from the left thumb *above* the right palmar string and round the *near* side of the right index finger, we have *A1 NR2*.

The sequence of characters, *NL1 A1 . . . FR1*, describes a string position. But it does *not* tell us anything about how to *construct* the string figure.

Deducing the sequence of hand movements

Figure 2 is reproduced from Jayne's book on string figures. It does not give the method of construction of the figures illustrated, originally collected by W. E. Roth from tribes in New Zealand. The string figure 5, which comes from Cape Bedford and represents a coconut, looked interesting to me, so I chose this as an example.

I expected that the method of making it would be in some vague sense typical of the sorts of movements used in making other string figures, and in particular would include roughly symmetric actions by the hands even though the final picture is not symmetric.



ROTH'S PLATE IX.—PLANTS

1. Hole in limb of tree: opossum, honey, etc., inside it. Princess Charlotte Bay.
2. Zamia (Cycas) tree. Atherton.
3. Zamia: nuts. Atherton.
4. Two coconuts. Cape Grafton.
5. Coconut. Cape Bedford.
6. Yams. Night Island.
7. Yams. Princess Charlotte Bay. Edible lily root: (Lower) Palmer River.

Figure 2. W. E. Roth's plate IX. The text discusses a method of making 5, Coconut.

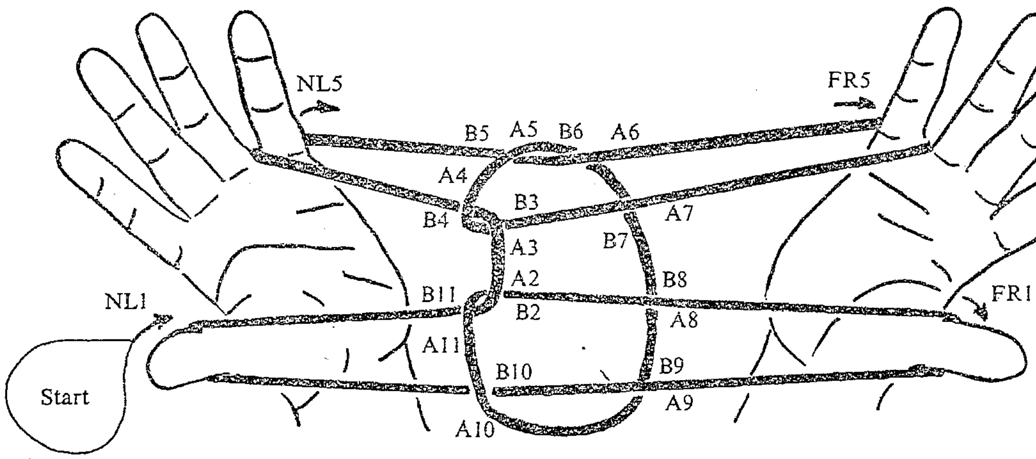


Figure 3. The Coconut redrawn and labelled.

I redrew the picture and labelled it as in Figure 3. This gives the following description:

*NL*1 *B*11 *A2 A3 B4 NL5 B5 A6 FR5 A7 B3 A4*
*A5 B6 B7 B8 B9 A*10 *A*11 *B2 A8 FR*1 *A9 B*10 (1)

One technique is to try to work backwards from this long sequence of characters to produce shorter sequences stage by stage, but in such a way that it would be easy with the hands to work forward from each stage to the next, in the hope of eventually getting back to a figure whose construction is known.

Sometimes when a figure is extended by drawing the hands as far apart as possible, a loop is pulled out straight. If this loop was previously above some other straight string, as in Figure 4, then two crossing points are removed. A sequence such as

· · · *A*53 *A*64 – – – – *B*53 *B*64·–·–·

becomes · · · – – – – – ·–·–·.

In other words, the crossing points 53 and 64 cancel out and the character string is shortened.

So one of the techniques is to look for longish sequences of *A*'s (or *B*'s) and see if they can be made shorter, or even longer. Such a sequence may correspond to a loop above (or below) others. Because the representation of certain types of knots

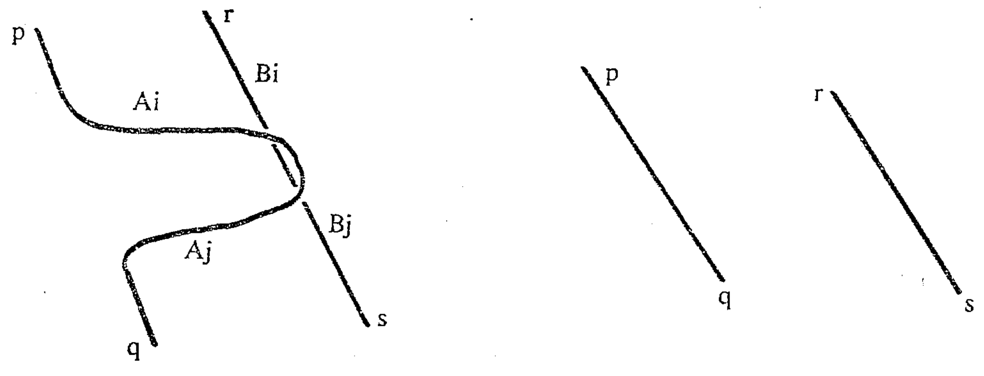


Figure 4. An example of strings sliding when extended; it corresponds to changes to the sequence of characters.

is not unique it may be a hint of a loop sliding over another or the dropping of a loop from a finger at a preceding stage. In sequence (1) we have *B*6 *B*7 *B*8 *B*9 with two *A*'s on each side. If, for the moment, the below loop *B*6 *B*7 *B*8 and *B*9 had been reversed into an above loop, say *A*61 *A*71, *A*81, *A*91, then we could have had an even longer group of *A*'s altogether, namely

*NL*1 *B*11 *A*2 *A*3 *B*4 *NL*5 *B*5 *B*61 *FR*5 *B*71 *B*3 *A*4
*A*5 *A*61 *A*71 *A*81 *A*91 *A*10 *A*11 *B*2 *B*81 *FR*1 *B*91 *B*10 (2)

Such a reversal corresponds to an easy hand action at the right stage. If that above loop originally had been on the right index finger as *FR*2 and dropped from it, at a previous stage we might have had

*NL*1 *B*11 *A*2 *A*3 *B*4 *NL*5 *B*5 *B*61 *FR*5 *B*71 *B*3 *A*4
*A*5 *A*61 *A*71 *FR*2 *A*81 *A*91 *A*10 *A*11 *B*2 *B*81 *FR*1 *B*91 *B*10 (3)

We can compare this with Opening *A* which we know how to construct. In (4) the lower line represents Opening *A* with the crossing points renumbered.

*NL*1 *B*11 *A*2 *A*3 *B*4 *NL*5 *B*5 *B*61 *FR*5 *B*71 *B*3 *A*4
*NL*1 *A*21 *NR*2 *A*22 *NL*5 *FR*5 *B*22
*A*5 *A*61 *A*71 *FR*2 *A*81 *A*91 *A*10 *A*11 *B*2 *B*81 *FR*1 *B*91 *B*10 (4)
*FL*2 *B*21 *FR*1

At each stage of our method we check for a number of types of short sequences which can be simplified; one of these types is shown in Figure 4. Among the characters in the upper line of (4) the check corresponding to Figure 4 finds

. . . *B*5 *B*61 . . . *A*5 *A*61 . . .

so we remove or cancel these to give

*NL*1 *B*11 *A*2 *A*3 *B*4 *FR*5 *B*71 *B*3 *A*4 *A*71
*FR*2 *A*81 *A*91 *A*10 *A*11 *B*2 *B*81 *FR*1 *B*91 *B*10 (5)

In (5) we have another example

. . . *A*91 *A*10 . . . *B*91 *B*10 . . .

so we cancel to give

*NL*1 *B*11 *A*2 *A*3 *B*4 *NL*5 *FR*5 *B*71 *B*3 *A*4 *A*71 *FR*2 *A*81 *A*11 *B*2 *B*81 *FR*1
*NL*1 *A*21 *NR*2 *A*22 *NL*5 *FR*5 *B*22 *FL*2 *B*21 *FR*1
(6)

This seems a little nearer to Opening *A*. In particular, crossing points 2 and 3 correspond to 21 and 22 of Opening *A*, but *NR*2 and *FL*2 for *FR*2 do not seem to fit yet. The early part of sequence (6) implies that, if the start was Opening *A*, then the next step may have been to 'drop the loop on the right index finger' to get

. . . *A*2 *A*3 . . .

from

. . . *A*21 *NR*2 *A*22 . . .

One of the other checks done at each stage is to simplify twisted loops as in Figure 5. This applies twice in the present case.

*B*11 *A*2 and *A*81 *A*11 *B*2 *B*81 become *A*2 *B*11 and *A*11 *B*2,
*B*4 *A*3 and *A*71 *A*4 *B*3 *B*71 become *A*3 *B*4 and *A*4 *B*3 (reversed).

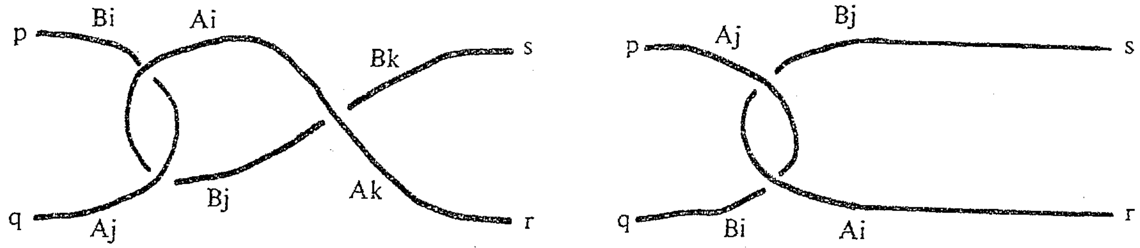


Figure 5. Simplification of a twisted loop.

Thus the original picture can be reduced to

*NL*1	*A*2	*B*11	*B*4	*A*3	*NL*5	*FR*5	*B*3	*A*4	*FR*2	*A*11	*B*2	*FR*1
*NL*1	*A*2			*A*3	*NL*5	*FR*5	*B*3		*FL*2		*B*2	*FR*1 (7)

The main discrepancy is *FR*2 for *FL*2. So the Cape Bedford native must have picked up onto his right index the loop on his left index, appropriately going above or below other strings to get the right effect. From the presence of *A*4 and *A*11, we can deduce that he took up the loop by going *above* the other crossing strings, and since *B*11 comes after *NL*1 in the sequence he must have gone above that string which leads from the left thumb.

We may conclude that the figure was probably made by the sequence in Table 2, which is derived by collecting together all the deductions above and describing the corresponding hand movements in the correct order.

TABLE 2. Coconut

Hand movements
Opening *A* (as Table 1). Drop the loop from the right index finger. Extend. Transfer the left index loop to the right index, holding index fingers tip to tip. Extend. With left thumb and index finger tips going down through the right index loop, temporarily hold the far string from the right thumb and the near string from the right little finger; remove right hand, letting the old right index loop drop below; replace the held loops onto the right thumb and right little finger. Extend, completing the Coconut.

The manipulation of the character sequences could always be done by hand, as the above example shows. However, for a complicated figure it can become quite tedious and difficult and therefore we have written a computer program on the Open University's computer. I am grateful to the University for allowing me the use of these facilities.

Reference

Caroline F. Jayne, *String Figures and How to Make Them.* Constable & Co, London, 1906: Dover edition 1962, SBN 486 20152X.

Letters to the Editor

Dear Editor,

Partitions of sets of integers

Problem 8.4 of *Mathematical Spectrum* (Volume 8, Number 2, p. 64) asked: Is it possible to partition the integers $1, 2, \ldots, 13$ into two subsets such that neither subset possesses three integers in arithmetic progression? The answer is 'no'. However, we can prove a stronger result, namely that it is not possible to partition the integers $1, 2, \ldots, 9$ in this way. For suppose we have partitioned these integers into two subsets A and B in this way, so labelled that $5 \in A$ (5 belongs to A). Suppose $4 \in A$. Then $6 \in B$, $3 \in B$, $9 \in A$, $7 \in B$, $8 \in A$, $2 \in B$, $1 \in A$, and now A contains 1, 5, 9 in arithmetic progression, a contradiction. Hence $4 \in B$. Suppose $6 \in B$. Then $2 \in A$ and $8 \in A$, and now A contains 2, 5, 8, again a contradiction. Hence $6 \in A$. Then $7 \in B$, $1 \in A$, $3 \in B$, $2 \in A$, $8 \in B$, $9 \in A$, and now A contains 1, 5, 9. This is the final contradiction which establishes the impossibility of the partition.

Further, if we consider the integers $1, 2, \ldots, 8$, then three such partitions are possible, namely

$$\begin{array}{lcl} \{1, 2, 5, 6\} & \text{and} & \{3, 4, 7, 8\}, \\ \{1, 3, 6, 8\} & \text{and} & \{2, 4, 5, 7\}, \\ \{1, 4, 5, 8\} & \text{and} & \{2, 3, 6, 7\}, \end{array}$$

so that 9 is the smallest number for which such a partition is impossible.

I then wondered what would happen if we partitioned the integers into three subsets. A little research by hand showed that $1, 2, \ldots, 13$ can be partitioned into three subsets, none of which possesses three integers in arithmetic progression. I developed a computer program which proceeded to partition the integers $1, 2, \ldots, 26$ into three subsets in the manner indicated; one solution is

$$\begin{array}{l} \{1, 2, 5, 6, 12, 14, 15, 17, 21\}, \\ \{3, 4, 7, 9, 16, 18, 19, 24, 26\}, \\ \{8, 10, 11, 13, 20, 22, 23, 25\}. \end{array}$$

Even after considering all solutions for $1, 2, \ldots, 26$, the computer was unable to include the integer 27 without introducing an arithmetic progression of three terms into one of the subsets.

A number of questions now arise. Given a positive integer n, we may ask what is the largest number m such that $1, 2, \ldots, m$ can be partitioned into n subsets in such a way that no subset contains three integers in arithmetic progression. When $n = 1$, we have $m = 2$; when $n = 2$, then $m = 8$; and when $n = 3$, then $m = 26$. We may surmise that, in general, $m = 3^n - 1$.

A colleague, Dr R. J. Firth, proceeded to write a much neater program for solving this more general problem. The only indication that we have of the validity or otherwise of the conjecture that $m = 3^n - 1$ is that, when $n = 4$, the computer was having difficulty in assigning 56 and 57. The partition of the numbers $1, 2, \ldots, 55$ is

$$\begin{array}{l} \{1, 2, 4, 5, 10, 11, 13, 14, 28, 29, 31, 32, 37, 38, 40, 41\}, \\ \{3, 6, 7, 12, 15, 16, 22, 30, 33, 34, 39, 42, 43, 49\}, \\ \{8, 9, 18, 21, 23, 26, 35, 36, 45, 48, 50, 53\}, \\ \{17, 19, 20, 24, 25, 27, 44, 46, 47, 51, 52, 54, 55\}. \end{array}$$

A further unanswered question is the following: if, as above, m is the maximum number for n partitions, how many different partitions can be found? When $n = 1$ (and $m = 2$), there is only one partition; when $n = 2$ (and $m = 8$), there are three, listed above; I do not know the number when $n = 3$.

I wonder whether these problems can be solved by an analytical approach, rather than with the use of a computer.

Yours sincerely,
M. D. SANDFORD
(Royal Military College of Science, Shrivenham)

Dear Editor,

Magic squares

Tut-tut, tut-tut!! In all the standard magic squares, not only are the sums of the verticals and horizontals equal to the magic number, but also so are the sums of the main diagonals. This is not so in the magic squares constructed by A. D. Misra (Volume 8, Number 2, pp. 53–60).†

However there is a remedy. In the case of magic squares of odd order n, it is only necessary to take the top $(n+1)/2$ rows of Misra's arrays and place them underneath the remaining $(n-1)/2$ rows (so that the top rows become the middle rows). This operation clearly leaves the row and column sums unchanged, and it is only necessary to show that the elements in each of the main diagonals have the right sum, namely $\frac{1}{2}n(n^2+1)$. To do this we start with the matrices $A_{n,p}$ and $B_{n,p}$ employed by Misra. It will be recalled that each element a_{ij} of $A_{n,p}$ is one of the numbers $0, 1, \ldots, n-1$ and that the (i, j)th element b_{ij} of $B_{n,p}$ is $a_{ij}+1$. By equation (2) in Misra's article, a_{ij} is the remainder after division by n of the integer

$$(j-1)+(i-1)(p-1).$$

Using the notation

$$a \equiv b \pmod{n},$$

which means that $a-b$ is divisible by n, and remembering that $0 \leqslant a_{ij} \leqslant n-1$, we see that a_{ij} is defined by the statement

$$a_{ij} \equiv (j-1)+(i-1)(p-1) \pmod{n}. \tag{1}$$

We now modify $A_{n,p}$ and $B_{n,p}$ by placing the top $(n+1)/2$ rows underneath the remaining ones, and we denote the resulting matrices by $A'_{n,p}$ and $B'_{n,p}$ respectively, so that the middle rows are $0, 1, 2, \ldots, n-1$ in $A'_{n,p}$ and $1, 2, 3, \ldots, n$ in $B'_{n,p}$. If a'_{ij} and b'_{ij} are the (i, j)th elements of these matrices, we have $b'_{ij} = a'_{ij}+1$ and, from (1),

$$a'_{ij} \equiv (j-1)+[(n+1)/2+i-1](p-1) \pmod{n}. \tag{2}$$

It is not difficult to see that any left-to-right diagonal of $A'_{n,p}$ contains the same elements as some left-to-right diagonal of $A_{n,p}$. But, when $2 \leqslant p \leqslant n-1$ and the pair (n, p) is coprime, then every left-to-right diagonal to $A_{n,p}$ contains all the integers $0, 1, \ldots, n-1$ (Lemma 2 in Misra's article). Hence, in particular,

I. If $2 \leqslant p \leqslant n-1$ and n, p are coprime, the principal left-to-right diagonal of $A'_{n,p}$ contains all the integers $0, 1, \ldots, n-1$; and the corresponding diagonal of $B'_{n,p}$ contains all the integers $1, 2, \ldots, n$.

† Traditionally the main diagonals enter into the definition of a magic square. However the situation is now less clear cut and in two of the articles quoted by Mr Misra (references 2 and 3) the diagonals are disregarded.—Editor.

To obtain the elements of the principal left-to-right diagonal of $A'_{n,n}$ we put $p = n$ and $j = i$ in (2). Then, when n is odd,

$$\begin{aligned} a'_{ii} &\equiv (i-1) + [(n+1)/2 + i - 1](n-1) && \pmod n \\ &\equiv (i-1)n + \tfrac{1}{2}(n+1)(n-1) && \pmod n \\ &\equiv \tfrac{1}{2}(n-1)n + \tfrac{1}{2}(n-1) && \pmod n, \end{aligned}$$

and therefore $a'_{ii} = \frac{1}{2}(n-1)$. Thus

II. If n is odd, every element of the principal left-to-right diagonal of $A'_{n,n}$ is $\frac{1}{2}(n-1)$.

The right-to-left diagonals of $A'_{n,p}$ are derived from the right-to-left diagonals of $A_{n,p}$. Therefore, by Misra's Lemma 3,

III. If $3 \leqslant p \leqslant n$ and $n, p-2$ are coprime, the principal right-to-left diagonal of $A'_{n,p}$ contains all the integers $0, 1, \ldots, n-1$.

Finally we require the elements b'_{ij} of the principal right-to-left diagonal of $B'_{n,2}$ when n is odd. Now $i + j = n + 1$ and, putting $p = 2$ in (2), we have

$$\begin{aligned} a'_{ij} &\equiv (j-1) + [(n+1)/2 + i - 1] && \pmod n \\ &\equiv n + \tfrac{1}{2}(n-1) && \pmod n, \end{aligned}$$

so that $a'_{ij} = \frac{1}{2}(n-1)$. Hence

IV. If n is odd, every element of the principal right-to-left diagonal of $B'_{n,2}$ is $\frac{1}{2}(n+1)$.

It is now easy to prove the analogues of Theorems 1 and 2 in Misra's article.

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

$$M'_n = B'_{n,2} + nA'_{n,n}$$

is a magic square in which the main diagonals have sums equal to the magic number.

Theorem 2′. If n is an odd integer $\geqslant 5$, and p is an integer such that $3 \leqslant p \leqslant n-1$ and $(n, p-1)$, $(n, p-2)$ are coprime pairs, then

$$M'_{n,p} = B'_{n,2} + nA'_{n,p}$$

is a magic square whose main diagonals have sums equal to the magic number.

Consider M'_n. By I and II, the principal left-to-right diagonal has sum

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

and by III and IV the principal right-to-left diagonal has sum

$$n \,.\, \tfrac{1}{2}(n+1) + n(0 + 1 + \cdots + n - 1) = \tfrac{1}{2}n(n+1) + \tfrac{1}{2}n^2(n-1) = \tfrac{1}{2}n(n^2+1).$$

Theorem 1′ is therefore proved. The proof of the second theorem is similar (but does not use II). Examples of magic squares with the diagonal property are

$$M'_3 = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} + 3 \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 7 & 2 \\ 1 & 5 & 9 \\ 8 & 3 & 4 \end{bmatrix},$$

and

$$M'_{5,4} = \begin{bmatrix} 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \end{bmatrix} + 5 \begin{bmatrix} 4 & 0 & 1 & 2 & 3 \\ 2 & 3 & 4 & 0 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 3 & 4 & 0 & 1 & 2 \\ 1 & 2 & 3 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 24 & 5 & 6 & 12 & 18 \\ 15 & 16 & 22 & 3 & 9 \\ 1 & 7 & 13 & 19 & 25 \\ 17 & 23 & 4 & 10 & 11 \\ 8 & 14 & 20 & 21 & 2 \end{bmatrix}.$$

It will be noticed that M'_3 and $M'_{5,4}$ are *associative*, i.e. that the sum of two elements symmetrical about the centre is constant for each magic square. *Spectrum* readers may like to prove that this is true for all the magic squares of Theorems 1′ and 2′.

To construct magic squares of order $4n$, which are magic in the main diagonals also, I start with the vector

$$(2, 4, 6, \ldots, 4n-4, 4n-2, 4n; 1, 3, 5, \ldots, 4n-5, 4n-3, 4n-1)$$

and interchange the first and $(2n)$th element as well as the $(2n+1)$th and last element to obtain

$$(4n, 4, 6, \ldots, 4n-4, 4n-2, 2; 4n-1, 3, 5, \ldots, 4n-5, 4n-3, 1). \qquad (3)$$

The $4n \times 4n$ matrix X'_{4n} then has (3) for its first n and last n rows; the $2n$ middle rows consist of (3) written backwards. The matrix Y'_{4n} has the vector

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

for its first n and last n columns, while the $2n$ middle columns are

$$(0, 1, \ldots, 4n-2, 4n-1).$$

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

$$M'_{4n} = X'_{4n} + 4nY'_{4n}$$

is a magic square in which the main diagonals have sums equal to the magic number.

The proof that M'_{4n} has elements $1, 2, \ldots, (4n)^2$ and is magic in its rows and columns is similar to the proof of Misra's Theorem 3. The last clause follows from the easily proved observation that the principal diagonals of X'_{4n} both consist of the numbers $1, 2, \ldots, 4n$, and the principal diagonals of Y'_{4n} consist of $0, 1, \ldots, 4n-1$.

Incidentally, it can be shown that the magic squares of Theorem 3′ are also associative. The squares M'_4 and M'_8 are

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

Yours sincerely,

T. MARSDEN
(2 Chalfont Avenue, Christchurch, Dorset)

Dear Editor,

Framed magic squares

For me a magic square is a square matrix, formed of distinct numbers, which is such that the elements in every row, in every column and in each of the two main diagonals add up to the same sum. This definition differs somewhat from the one given by A. D. Misra in his article in *Mathematical Spectrum*, Volume 8, Number 2, pp. 53–60. There the diagonals were not restricted, but on the other hand the elements of an nth order magic square were required to be the numbers $1, 2, \ldots, n^2$. When n is an odd integer, I call an $n \times n$ matrix a *framed* magic square if the matrix itself is a magic square, and the matrices (of orders $n-2, n-4, \ldots, 3, 1$) obtained by successively removing the outer frames of numbers are all magic squares as well. As far as I know there is no general method of constructing framed magic squares, but I have devised a system that reduces the amount of trial and error required. The method is adequately illustrated by the construction of a seventh-order framed magic square consisting of the numbers $1, 2, \ldots, 49$.

First the elements of the square, other than 25 (the arithmetic mean of $1, 2, \ldots, 49$) are grouped in pairs $(25 - r, 25 + r)$, where r is called the serial number of the pair. The arrangement is shown in Figure 1, with the serial numbers in the bottom line.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
49	48	47	46	45	44	43	42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	
7th-order frame												5th-order frame								3rd-order				Core
24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	

Figure 1

The number 25 is now placed at the centre of the square and the third-, fifth- and seventh-order frames are filled with the numbers assigned to them in Figure 1. The position of a particular pair of numbers in a frame is determined by the serial number of the pair.

The serial numbers for every frame are divided into four groups which correspond to the top and bottom rows and to the left- and right-hand columns of the frame. Half the serial numbers have to be allotted to the rows and half to the columns. Moreover, two numbers have to appear twice in the table, once in a row and once in a column; these will lead to the corner numbers of the square. Finally, the distribution must be such that the totals for the two groups corresponding to the rows are the same and the totals for the two groups corresponding to the columns are also the same. A possible disposition is shown in Figure 2.

3rd-order frame				5th-order frame				7th-order frame			
Top row	Bottom row	Left column	Right column	Top row	Bottom row	Left column	Right column	Top row	Bottom row	Left column	Right column
1	4	3	1	6	10	9	5	20	13	22	15
3			2	7	12	11	7	21	14	23	16
				9			8	22	17	24	18
									19		20

Figure 2

To use these tables, find from Figure 1 the pair of numbers associated with a given serial number and place the larger number of the pair in the same row or column of the frame as that in which the serial number appears in Figure 2. If the larger number is a corner one in a frame, then the smaller number of the pair is placed in the opposite corner of the frame. If the larger number is in a row, but not in a corner, the smaller number is placed in the corresponding position of the opposite row in the frame. The columns are filled in the same way.

For example, in the table for the third-order frame, the serial number 3 is allotted to both the top row and the left-hand column, and the associated pair is (22, 28). Hence 28 goes in the top left-hand corner of the frame and 22 in the bottom right-hand one. Or consider the serial number 10 in the fifth-order frame. It is in the bottom row of the frame, but not in a corner position. The corresponding pair is (15, 35); and if we place 35 in the fourth position of the bottom row, 15 must go in the fourth position of the top row in the frame. Once the tables of Figure 2 have been constructed, the corner numbers of each frame are fixed. However there is still some latitude in filling the interior positions of the rows and columns. Figure 3 gives a completed square.

47	46	6	12	8	11	45
7	34	31	13	15	32	43
49	17	28	21	26	33	1
48	36	23	25	27	14	2
9	20	24	29	22	30	41
10	18	19	37	35	16	40
5	4	44	38	42	39	3

Figure 3

It is apparent from the construction that a square obtained in this way is indeed a framed magic square. The advantage of the method outlined above is that the serial numbers are easier to handle than the actual numbers of the square: they are smaller, and there are only half as many of them. Using this system I have constructed a framed magic square of order 23.

When n is an odd integer, an $n \times n$ matrix with distinct positive elements is called a *subtractive* magic square if in each row, column and diagonal the sum of the 1st, 3rd, . . ., nth elements minus the sum of the 2nd, 4th, . . ., $(n - 1)$th elements is equal to a fixed number. Such squares can easily be obtained from the magic squares whose construction I have outlined. If we start with such a magic square of order n and replace alternate elements by their complements with respect to $n^2 + 1$, then the result is a framed subtractive magic square. To prove this one has to use not only the magic square properties of the original array, but also the symmetrical location in it of every pair of numbers with sum $n^2 + 1$. For instance the magic square of Figure 3 yields two framed subtractive magic squares (Figure 4).

Finally, I might mention that every magic square gives rise to many multiplicative magic squares. For take any positive integer a. If each number m of a magic square is replaced by the number a^m, the resulting matrix is clearly a multiplicative magic square of integers; and if the original magic square was framed, then the multiplicative magic square is also framed.

3	46	44	12	42	11	5
7	16	31	37	15	18	43
1	17	22	21	24	33	49
48	14	23	25	27	36	2
41	20	26	29	28	30	9
10	32	19	13	35	34	40
45	4	6	38	8	39	47

47	4	6	38	8	39	45
43	34	19	13	35	32	7
49	33	28	29	26	17	1
2	36	27	25	23	14	48
9	30	24	21	22	20	41
40	18	31	37	15	16	10
5	46	44	12	42	11	3

Figure 4

In the same way a framed divisive magic square can be made from a framed subtractive magic square.

Yours sincerely,
C. HARGREAVES
(34 Rooley Lane, Bankfoot, Bradford)

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

9.1. Prove that e^x cannot be expressed in the form $f(x)/g(x)$, where $f(x), g(x)$ are polynomials in x with real coefficients.

9.2. (Submitted by B. G. Eke, University of Sheffield.) A changing room has n lockers numbered 1 to n and all are locked. An attendant performs the sequence of operations $T_1, T_2, \ldots, T_n$, where T_k is the operation whereby the condition of being locked or unlocked is altered in the case of those lockers (and only those) whose numbers are divisible by k. Which lockers are unlocked at the end?

9.3. The real numbers $a_1, \ldots, a_n, b_1, \ldots, b_n$ $(n \geqslant 1)$ are such that

$$a_1 \leqslant \frac{1}{2}(a_1 + a_2) \leqslant \frac{1}{3}(a_1 + a_2 + a_3) \leqslant \cdots \leqslant \frac{1}{n}(a_1 + a_2 + \cdots + a_n),$$

$$b_1 \leqslant \frac{1}{2}(b_1 + b_2) \leqslant \frac{1}{3}(b_1 + b_2 + b_3) \leqslant \cdots \leqslant \frac{1}{n}(b_1 + b_2 + \cdots + b_n).$$

Show that

$$\left(\sum_{k=1}^{n} a_k\right)\left(\sum_{k=1}^{n} b_k\right) \leqslant n \sum_{k=1}^{n} a_k b_k.$$

(See the article by L. Mirsky in this issue.)

Solutions to Problems in Volume 8, Number 2

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

Solution by Patrick Brooke (Winchester College)

Suppose that the given integers have been partitioned into two subsets such that neither subset possesses three integers in arithmetic progression. We first establish the following: *if n, a are positive integers such that $3a < n < 14 - 3a$, then $n - a$, $n + a$ cannot lie in the same subset.* For suppose the contrary. Now $n - 3a$, $n + 3a$ lie between 1 and 13, and they cannot belong to the same subset as $n - a$ and $n + a$. They must therefore both belong to the complementary subset. But now n cannot belong to either subset, a contradiction.

If we take $n = 6$, $a = 1$, we see that 5, 7 cannot belong to the same subset. If we take $n = 7$, $a = 2$, we see that 5, 9 cannot belong to the same subset. But, with $n = 8$, $a = 1$, we have that 7, 9 cannot belong to the same subset. But two of 5, 7, 9 must belong to the same subset. This gives a contradiction, so that it is not possible to partition the given integers in the manner asserted.

For a further discussion of this problem, see the letter by M. D. Sandford in this issue.

8.5. 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

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

Solution by Patrick Brooke

Let i be a positive integer less than r. Then

$$\frac{1}{r-i} + \frac{1}{r+i} = \frac{2r}{r^2 - i^2} > \frac{2}{r}.$$

Hence

$$\sum_{i=1}^{s} \left(\frac{1}{r-i} + \frac{1}{r+i}\right) > \frac{2s}{r}.$$

If we add $1/r$ to both sides we obtain the first inequality.

Now, for each non-negative integer a, put $r_a = 3^a$, $s_a = (3^a - 1)/2$, so that $r_a = 2s_a + 1$.

Then

$$r_a + s_a + 1 = 3^a + \tfrac{1}{2}(3^a - 1) + 1 = \tfrac{3}{2}3^a + \tfrac{1}{2} = 3^{a+1} - \tfrac{1}{2}(3^{a+1} - 1) = r_{a+1} - s_{a+1}.$$

Also, $r_0 - s_0 = 1$ and

$$r_{n-1} + s_{n-1} = 3^{n-1} + \tfrac{1}{2}(3^{n-1} - 1) = \tfrac{1}{2}(3^n - 1).$$

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

$$\begin{aligned} 1 + \frac{1}{2} + \cdots + \frac{1}{m} &= \sum_{a=0}^{n-1} \left(\frac{1}{r_a - s_a} + \frac{1}{r_a - s_a + 1} + \cdots + \frac{1}{r_a + s_a}\right) \\ &> \sum_{a=0}^{n-1} \left(\frac{2s_a + 1}{r_a}\right) \\ &= n. \end{aligned}$$

8.6. Show that the product of four consecutive positive integers cannot be a perfect cube.

Solution

Suppose that the positive integer n is such that the product of n, $n + 1$, $n + 2$, $n + 3$ is a perfect cube. One of $n + 1$, $n + 2$ is odd, and so is relatively prime to the other three. Thus it must itself be a perfect cube. Hence the product of the remaining three must be a perfect cube. Suppose $n(n + 1)(n + 3)$ is a perfect cube. Then

$$\text{either} \quad n(n + 1)(n + 3) = (n + 1)^3 \quad \text{or} \quad n(n + 1)(n + 3) = (n + 2)^3.$$

In the former case we obtain $n = 1$, yet $1 . 2 . 3 . 4$ is not a perfect cube. In the latter case we obtain $2n^2 + 9n + 8 = 0$, which is impossible.

Suppose now that $n(n + 2)(n + 3)$ is a perfect cube. Then

$$\text{either} \quad n(n + 2)(n + 3) = (n + 1)^3 \quad \text{or} \quad n(n + 2)(n + 3) = (n + 2)^3.$$

In the former case we obtain $2n^2 + 3n = 1$, and in the latter $n = -4$, both of which are impossible.

Thus all cases lead to a contradiction.

Also solved by Patrick Brooke.

Book Reviews

Mathematics at the University. By the Universities and Schools Committee of the Mathematical Association, edited by M. L. Cornelius and H. Neill. G. Bell and Sons, London, 1975. Pp. 7. £0·25 singly, £1·25 for 10 copies, £2·00 for 20 copies (including postage) from The Mathematical Association, 259 London Road, Leicester LE2 3BE; or from the publisher.

Sixth formers are often told that reading mathematics at a university or polytechnic is very different from learning the subject at school. This pamphlet (which is a reprint of an article that originally appeared in *The Mathematical Gazette*) is intended to tell the future mathematics student what to expect, and it also offers some excellent advice on how to cope with novel experiences such as lectures and tutorials. There is a description of the likely content of a first-year mathematics course, but this meets an inevitable obstacle: trying to convey to the uninitiated the flavour of analysis or of abstract algebra is very much like an attempt to describe the taste of meat to a life-long vegetarian; in both cases nothing can take the place of personal acquaintance. The article includes a short list of general mathematical books which can usefully be read or dipped into during the period between school and university. Advice on more technical preparatory reading is, very sensibly, left to individual universities.

This pamphlet is genuinely useful, and if sixth formers won't buy it for themselves, they should, at least, make sure that their schools obtain an adequate supply of copies.

University of Sheffield H. BURKILL

Probability, Statistics and Time. By M. S. BARTLETT. Chapman and Hall, London, 1975. Pp. viii+148. £4·75.

This collection of two inaugural lectures, six invited addresses and one invited article by Professor Bartlett span the years 1961–1973. Their titles give an idea of the author's wide range of interests:

1. Probability, statistics and time (inaugural lecture at University College London, 1961).
2. R. A. Fisher and the last fifty years of statistical methodology.
3. The paradox of probability in physics.
4. Inference and stochastic processes.
5. Biomathematics (inaugural lecture in the University of Oxford, 1968).
6. When is inference *statistical* inference?
7. Epidemics.
8. Equations and models of population change.
9. Some historical remarks and recollections on multivariate analysis.

Viewed as a whole, these provide a coherent picture of the author's recent work in probability and statistics, and his views on current developments in the field. While Sections 3 and 8 are possibly too technical for the reader interested mainly in ideas rather than mathematical details, most of the remaining sections can be read with profit by the interested layman and by students of mathematics.

The book can be recommended as an introduction to some important ideas in probability and statistics expounded by one of the fundamental thinkers in the field.

C.S.I.R.O., Canberra J. GANI

Modern Introduction to Classical Mechanics and Control. By D. N. BURGHES and A. M. DOWNS. Ellis Horwood, Chichester, 1975. Pp. 320. £7·50, cloth; £3·50, paperback.

In recent years there has been a deplorable swing of student interest away from physics and Newtonian mechanics; deplorable because of the danger of an acute shortage of qualified physicists, engineers and applied mathematicians if the swing continues much longer.

This book, aimed at sixth-formers and first-year undergraduates, should go some way to redress the balance. It is concerned very largely with particle dynamics, but far from being a dreary repetition of traditional material, the book is an interesting introduction to the basic ideas of the subject, which should help to convince students and others that there are many important problems in modern science and technology which can be approached from the standpoint of Newton's laws of motion.

The authors provide a certain amount of historical background, but their choice of examples and exercises emphasises the contemporary relevance of the subject. For example, space dynamics and optimal control theory are discussed and topics like the solar wind, economic models and vibrations in crystal lattices receive a mention. In fact, a most attractive feature of the book is the way in which the student's attention is directed towards more advanced developments not only in mechanics, but also in computing and pure mathematics.

The book is very readable and has clear explanations and a profusion of worked examples, which serve both to illuminate the principles of the subject and to extend the student's skill in manipulation.

A few chapters on the fundamentals of rigid body dynamics would have been welcome, but perhaps we can hope that the authors will provide a sequel to this book.

University of Exeter G. A. DULLER

Computers and Commonsense. By ROGER HUNT and JOHN SHELLY. Prentice-Hall International, Hemel Hempstead, 1975. Pp. 150. £2·50 limp covers.

This is a most attractively produced book whose appearance is not let down by its contents. Almost all aspects of computers and computing, for instance history, development, hardware and social issues are covered in clear, unambiguous, non-technical terms. (Technical jargon is always explained.)

Line diagrams are used well and so add clarity to the text. The book can be read from cover to cover or by taking isolated chapters for specific topics, without loss of understanding.

Since the book is so comprehensive, topics are not explored to great depth. This must be done in further specialist reading, for which purpose suitable bibliographies are given under each chapter heading.

This book provides an excellent general background for teachers and senior students (i.e. A-level or above) who require a broad appreciation of computers. It would also be a useful reference book for the informed O-level student.

St Leonard's Comprehensive School, Durham T. HOWARD

Notes on Contributors

L. Mirsky is a Professor of Pure Mathematics in the University of Sheffield. Though he is not a professional analyst, he has always been fascinated by inequalities. His principal mathematical interest lay originally in the theory of numbers, then shifted to linear algebra and finally came to rest in combinatorics (where it may or may not remain).

Graham Upton lectures in statistics at the University of Essex. Previously he was a Lecturer at the University of Newcastle-upon-Tyne, where his positional bias research took place. His current research involves creating a mathematical model to explain voting changes and the uniformity of 'swing'.

David Schonland took the Mathematical Tripos in Cambridge and obtained a Ph.D. in Birmingham for work in mathematical physics. He is now a Senior Lecturer in Mathematics at the University of Southampton. Although his research interest lies in atomic physics, he has written a book on group theory and its uses in chemistry.

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.

Contents

ISSN 0025–5653

PRICES (*postage included*)

Prices for Volume 9 (*Issues Nos.* 1, 2 *and* 3):

Subscribers in Britain and Europe: £1·00

Subscribers overseas: £2·00 (US$4.00; $A.3.20)

(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):
£1·30 (US$2.60; $A.2.10) per volume.

Volumes 7 and 8 (3 issues each volume):
£2·00 (US$4.00; $A.3.20).

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