Mathematical Spectrum

Volume 5 1972/73 *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. It is published by the Applied Probability Trust, a non-profit making organization 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 5 of *Mathematical Spectrum* (1972/73) will consist of two issues, the second of which will be published in the Spring of 1973.

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

Football Statistics

PAUL R. JACKSON
University of Sheffield

1. Introduction

Since the inauguration of the Football League Cup in 1960, it has twice been won by third division clubs, Queens Park Rangers in 1967 and Swindon Town in 1969. Aston Villa were beaten finalists in 1971, yet a third division club has never reached the final of the F.A. (Football Association) Cup. The F.A. Cup has always been regarded as the more important competition, and indeed in the earlier years of the League Cup, some of the first division clubs did not even enter this competition. Since the 1966–67 season, the League Cup final has been played at Wembley; victory in it guarantees entry to European competition for first division clubs, and all clubs now enter the competition (reference 1).

It therefore seems of interest to compare the performances of third division clubs in the League Cup before and after 1967, and then to compare their performances in the League and F.A. Cups. We shall attempt to answer the question 'Can a third division club win the F.A. Cup?'

2. League Cup distributions

We shall consider progress of third division clubs from Round 2 onwards, as the first round is played by third and fourth division clubs only. First and second division clubs join the competition in Round 2. The League Cup is a knock-out competition with 5 rounds, a semi-final and a final, or 7 games altogether. Each club continues in the competition until it is first defeated. This suggests that the progress of a third division club in the competition could be described by a geometric distribution.

We shall define the random variable X as the number of successive wins (including one in Round 1) by a third division club competing in the League Cup in any given year. Under this definition, for example, $X = 7$ means that the club wins every one of the 7 games in the League Cup, since X cannot exceed 7. Figure 1 is a histogram of the observed values of X for the 11 years, 1960–71, in question.

Before investigating the distribution of X for the 11 year span, we shall first test for any difference in third division performance before and after the introduction of the Wembley finals. To do this, we shall assume that the distribution is geometric, this assumption being justified later. We then compare the parameters for the distribution of X in Sample 1 (before 1967) with those for X in Sample 2 (after

1967). We denote the values of X in the two samples by X_1, X_2 respectively. The respective sample sizes are $n_1 = 105$ and $n_2 = 66$. Figure 1 breaks down as shown in Table 1.

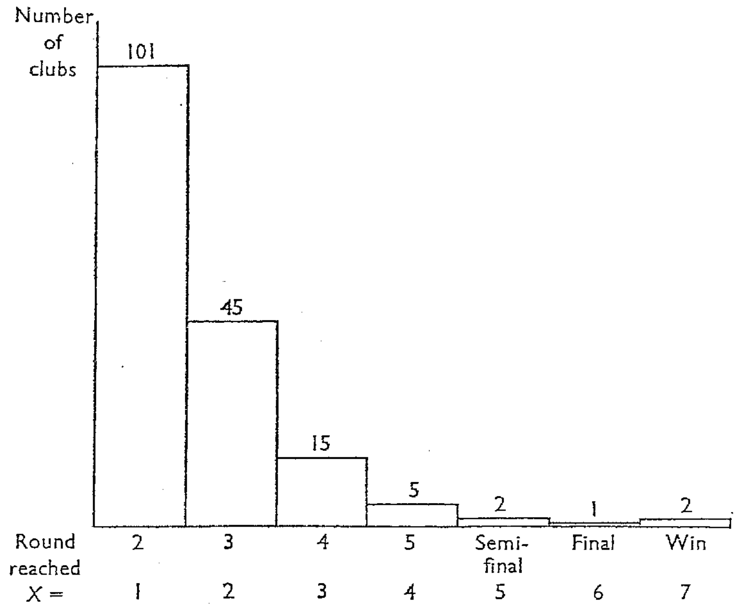


Figure 1. Histogram for the number of successive wins, X, of third division clubs in the League Cup, 1960–71, with corresponding rounds reached.

TABLE 1

Values of X	1	2	3	4	5	6	7
Frequencies of X_1	62	28	10	3	2	0	0
Frequencies of X_2	39	17	5	2	0	1	2

We wish to test the null hypothesis H_0 that X_1 and X_2 have identical geometric distributions of the form $\Pr(X = x) = p^{x-1}q$; $x = 1, 2, 3, \ldots$, with mean $1/q$ and variance p/q^2. Here p, $0 < p < 1$, denotes the probability of 'success' in winning a match, and q is the probability of 'failure' or defeat, so that $p+q = 1$. We assume that p remains constant for all matches. Thus, if q_1, q_2 denote the probabilities of failure for Samples 1 and 2 respectively our hypothesis is H_0: $q_1 = q_2$.

We thus require estimates for q_1 and q_2 based on the data. An optimal estimate for q_1 given by the maximum likelihood method is:

$$\hat{q}_1 = \frac{1}{\bar{X}_1}, \quad \text{where } \bar{X}_1 = \sum_{x=1}^{7} x_1 f_1(x),$$

and where $f_1(x)$ is the relative frequency of x in Sample 1. Hence,

$$\begin{aligned}\bar{X}_1 &= \tfrac{1}{105}(1.62+2.28+3.10+4.3+5.2)\\ &= 1{\cdot}62\end{aligned}$$

so that

$$\hat{q}_1 = 0{\cdot}6177.$$

Similarly, for Sample 2 we find that $\bar{X}_2 = 1{\cdot}76$ and $\hat{q}_2 = 0{\cdot}5689$. Under H_0, $q_1 = q_2 = q$ (say), and a pooled estimate of q is $\hat{q} = (n_1\hat{q}_1 + n_2\hat{q}_2)/(n_1+n_2) = 0{\cdot}6$ approximately.

Now, as the sample sizes are large, the central limit theorem applies (reference 2). Thus under H_0, we have that $\bar{X}_1 - \bar{X}_2$ is normally distributed, with zero mean; the variance can be calculated by substituting $\hat{q}_1$ and $\hat{q}_2$ in the above formula for the variance, p/q^2. It turns out that the observed difference $\bar{X}_1 - \bar{X}_2$ corresponds to only $0{\cdot}14s$, where s is the estimated standard deviation; this is less than $1{\cdot}96s$, the critical 5% significance level. Thus, we may accept the hypothesis that no difference occurs in the performances of third division clubs in the League Cup before and after the introduction of Wembley finals.

We can now, therefore, consider the League Cup data in Figure 1 as a whole and attempt to justify the use of the geometric distribution, with constant parameter $q = 0{\cdot}6$. We shall make use of the χ^2 goodness of fit test.

Our total sample size is 171, and we shall test the hypothesis H_0 that the performance in the League Cup of third division clubs follows this distribution. To validate the χ^2 test, we pool the data for $X \geqslant 5$ to obtain expected frequencies, from Figure 1, of size at least 4.

Denoting by O_x the observed, and E_x the expected frequencies, where $E_x = 171\,\Pr(X = x)$, we have the figures in Table 2.

TABLE 2

x	$\Pr(X = x)$	E_x	O_x	$(O_x - E_x)^2$	$(O_x - E_x)^2/E_x$
1	$q = 0{\cdot}600$	103	101	4	0·038
2	$pq = 0{\cdot}240$	41	45	16	0·390
3	$p^2q = 0{\cdot}096$	16	15	1	0·063
4	$p^3q = 0{\cdot}038$	7	5	4	0·572
$\geqslant 5$	$\sum_{x=5}^{\infty} p^{x-1}q = 0{\cdot}026$	4	5	1	0·250

This gives, $u = \sum_x (O_x - E_x)^2/E_x = 1{\cdot}313$. Under H_0, u has a χ^2_3 distribution with 3 degrees of freedom (since q was estimated from the data). At the 5% significance level, the critical value of χ^2_3 is 7·81. Our value is 1·313 which is well within the acceptance region, so we may accept H_0.

We conclude that the progress of third division clubs beyond the second round of the League Cup could well be governed by a geometric distribution with parameter $q = 0{\cdot}6$.

It follows that the probability of a third division club winning the League Cup, given that it has reached Round 2, is $\Pr(X = 7) = 0{\cdot}0025$ approximately. Since about 16 clubs reach Round 2 each year, the probability of the League Cup being won by any third division club in any year is $16(0{\cdot}0025) = 0{\cdot}04$. Thus, on average, the League Cup will be won by a third division club once in every 25 years. In fact this has happened twice in 11 years. How can we reconcile these two statements?

We shall use the binomial distribution to find the probability of at least two successes (i.e., third division League Cup winners) in 11 trials given a probability of success $p = 1/25$. Now,

$$\begin{aligned} &\Pr(\text{at least 2 successes in } 11 \mid p = 1/25) \\ &\quad = 1 - \Pr(0 \text{ or } 1 \text{ success}) \\ &\quad = 1 - \binom{11}{0}\left(\frac{1}{25}\right)^0\left(\frac{24}{25}\right)^{11} - \binom{11}{1}\left(\frac{1}{25}\right)^1\left(\frac{24}{25}\right)^{10} \\ &\quad = 0{\cdot}0685. \end{aligned}$$

This probability is greater than 0·05; thus at the 5% significance level, we could accept that the probability of the League Cup being won by a third division club in any one year is 1/25, and that this is not inconsistent with observed events.

3. F.A. Cup distributions

The F.A. Cup is a knock-out competition with 6 rounds, a semi-final and a final, with first and second division clubs joining in Round 3. We thus consider third division clubs only beyond this round.

As for the League Cup, a club continues in this competition until it is first defeated, so we define the random variable Y as the number of successive wins after Round 2 by a third division club, in any given year. Again, the maximum value of Y is 7, corresponding to a club winning the F.A. Cup.

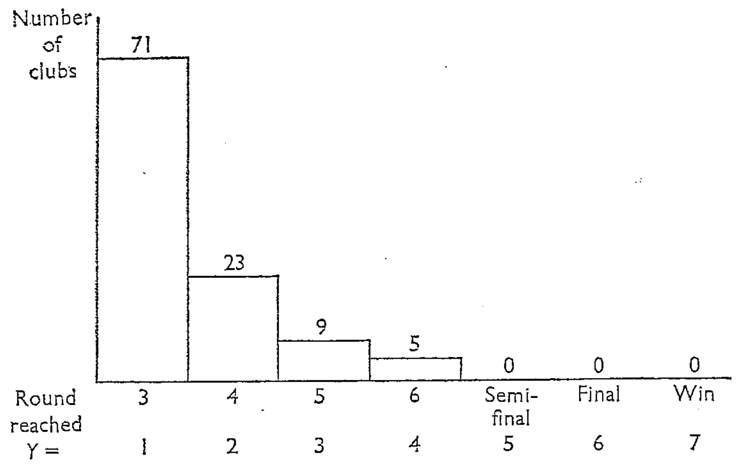


Figure 2. Histogram for the number of successive wins after Round 2, Y, of third division clubs in the F.A. Cup, 1960–71, with corresponding rounds reached.

We shall use the χ^2 goodness of fit test to show that the data in Figure 2 are compatible with a geometric distribution given by $\Pr(Y = y) = p^{y-1}q$ $(y = 1, 2, \ldots)$.

As before, we estimate q by:

$$\bar{Y} = \frac{1}{\hat{q}} = \sum_{y=1}^{7} yf(y)$$
$$= \tfrac{1}{108}(1.71+2.23+3.9+4.5+0)$$
$$= 1{\cdot}517.$$

Hence $\hat{q} = 0{\cdot}66$ approximately, so we test the hypothesis H_0 that Y has a geometric distribution with parameter $q = 0{\cdot}66$.

As before we compare, in Table 3, the observed frequencies, O_y, with the frequencies expected, E_y, under our hypothesis with sample size 108.

TABLE 3

Y	O_y	E_y	$(O_y-E_y)^2$	$(O_y-E_y)^2/E_y$
1	71	72	1	0·014
2	23	24	1	0·042
3	9	8	1	0·125
$\geqslant 4$	5	4	1	0·25

Our observed value of $u = \sum_y (O_y - E_y)^2/E_y$ for the value of χ_2^2 with 2 degrees of freedom is 0·431 which is clearly not significant. Hence, we conclude that progress of third division clubs beyond the third round of the F.A. Cup could well be governed by a geometric distribution with parameter $q = 0{\cdot}66$. Using this, we obtain the probability for such a club to win the F.A. Cup as $\Pr(Y \geqslant 7) = 0{\cdot}0009$ approximately. On average about 10 clubs reach Round 3 each year, which means that approximately once in every 110 years a third division club will win the F.A. Cup! This is certainly consistent with the fact that a third division club has never won the F.A. Cup since the third division was established 44 years ago.

4. Comments

Basically, we have fitted a mathematical model (the geometric distribution) to the observed data; and in so doing have made one or two assumptions which we must now examine. Firstly, a geometric distribution takes all positive integral values, but our random variables X and Y take values 1,..., 7 only. This appears to have no serious effect on the results, however, since the probabilities are so small at the tail end of the distribution that, for example, $\Pr(X = 7)$ and $\Pr(X \geqslant 7)$ are almost exactly the same. Secondly, in fitting the geometric distribution, we must assume that the rounds in which clubs are knocked out are independent. This is not always so, for if two third division clubs meet in a certain round, they cannot both proceed to the next round. This occurred on only a small number of occasions in the rounds in question, and so has only had a minimal effect.

Also, in fitting the geometric distribution, we assumed the parameters p and q were constant for all matches. Obviously a club is more likely to win certain matches than others; but we have assumed that this averages out when a large number of matches is considered. As, in both cases, we had such a good fit to a geometric distribution, this assumption seems justified.

We know from the Football Yearbook (reference 1) that two third division clubs have won the League Cup, but a third division club has never even reached the F.A. Cup final. Let us consider the probabilities of these events. The probability of winning the League Cup is $p^6q = 0{\cdot}0025$, and of winning the F.A. Cup only 0·0009. Thus we see that a third division club is nearly three times as likely to win the League Cup as against the F.A. Cup. A further explanation of why third division clubs have never reached the F.A. Cup final lies in the fact that fewer third division clubs reach the third round of the F.A. Cup (corresponding to $Y = 1$) than reach the second round ($X = 1$) of the League Cup, as they must play one more match in reaching Round 3 of the F.A. Cup. Thus the 'sample size' is smaller for the F.A. Cup and so we can expect fewer clubs to reach the latter stages of the competition.

We have considered the third division here because of the special interest in this division outlined in the introduction. Obviously, a similar analysis could be carried out with any other division or indeed with any other cup competition. Reference 1 would be indispensable if such an exercise were undertaken. The statistical concepts used are fairly elementary, and the reader will find answers to his difficulties explained in reference 2.

Acknowledgment

The author would like to thank Mr R. M. Clark for assistance in preparing this article for publication.

References

1. *Rothmans Football Yearbook* (Queen Anne Press), published annually.
2. Paul L. Meyer, *Introductory Probability and Statistical Applications* (Addison-Wesley, 1965).

Plangers

V. W. BRYANT
University of Sheffield

When I was six years old I read the following problem in one of the learned journals (see reference 1).

Three houses have to be connected by pipes or wires to each of the local gas-works, water-works and electricity power-station. Show how this can be done with no pipes or wires crossing each other.

Even at that early age I saw that the terms were ill-defined, but with fervour I set to work on the exercise. No matter how cunningly I twisted the pipes and wires I was unable to construct all nine connections, as Figure 1 shows, and so I looked forward eagerly to the solution in the next edition. When the 'solution' appeared it consisted of tunnelling the pipe between C and W under the power-station! Cheated, I decided to devote my life to solving the problem. Luckily this phase did not last long and I soon wanted to be an engine-driver again.

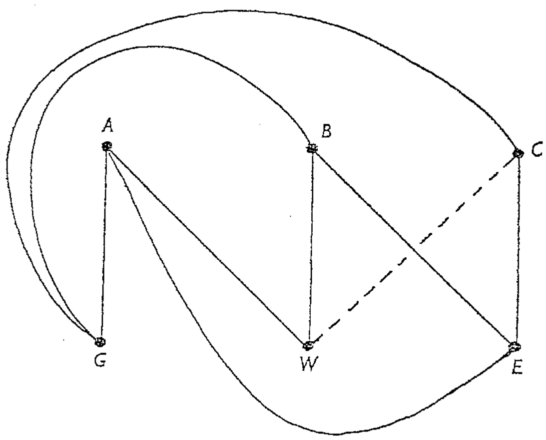


Figure 1

Now, some twenty years later, the decline of the railways has forced me back to my original pursuit, but not quite in the same terms. Put your pencil-point on a piece of paper, trace a line crossing itself only a finite number of times, and call the resulting diagram a 'Planger'. Let I be the collection of intersection points or 'nodes', and call two nodes 'joined' if there exists a stretch of pencil-line between them uninterrupted by other nodes. Thus in Figure 2, I consists of five points, labelled as shown, with 1 joined to 2, 3 and 4 only.

Our original problem is to construct a planger with six nodes labelled A, B, C, G, W, E and with each of A, B, C joined to each of G, W, E. It turns out, in fact, that this is impossible (or that this situation is non-planger.) To have discovered this at the age of six would have been heart-breaking: to discover it now is joy itself. Let me illustrate that it is non-planger: if it were a planger, then in the resulting diagram there would have to be a contorted form of Figure 3 (since A is joined to G and E, etc.) We are then still left to fill in joins $A-W$, $B-E$ and $C-G$, and you can easily see that only one of these can be drawn inside the loop, and only one can be drawn outside.

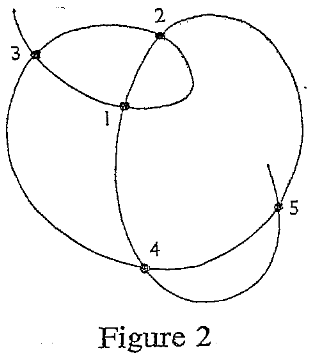


Figure 2

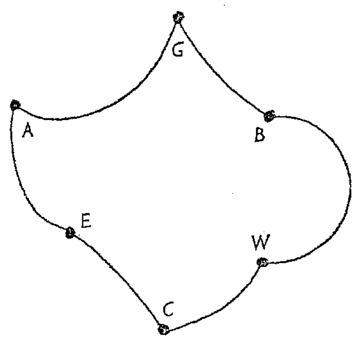


Figure 3

Another non-planger situation is one where there are five nodes each of which is joined to the other four. The proof of this is left as an exercise (see the 'Problems' section). So the '3-services' and 'full-5' situations are non-planger, and the amazing thing is that, in effect, these are the only non-planger situations. More precisely, if you tell me that you want a set of, say, 317 nodes (i.e., points joined to three or more others) and you stipulate which ones are to be joined to which, then it will be a planger provided there are no six nodes with three of them joined to each of the three others, and provided there are no five nodes each joined to each other. A moment's thought will show that it takes a lot more than a moment's thought to prove this. However the following illustration will show that it is true in an easy case when there are only six nodes given. We will assume these six do not contain either a 3-services or full-5 situation and show that it is in fact a planger. Since it contains no full-5, two of the nodes (labelled 1 and 2, say) are not joined and we can form a planger of five of the nodes as shown in Figure 4, (assuming

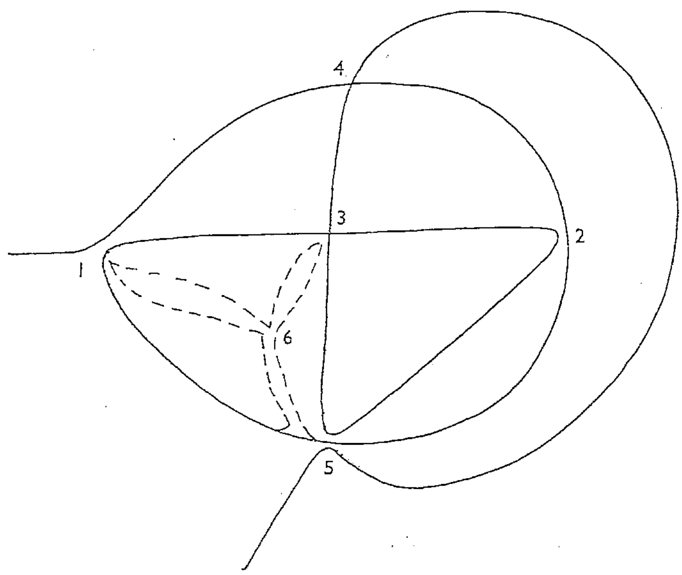


Figure 4

the worst case—that all others of the five are joined). Here 1 and 2 are already joined to 3, 4 and 5, and so since the situation does not contain a 3-services the last node, 6, cannot be joined to all of 3, 4 and 5. Similar arguments show that 6 can therefore only be joined to 1, 3, 4 or 1, 3, 5 or 1, 4, 5 or 2, 3, 4 or 2, 3, 5 or 2, 4, 5. But in each case it is easy to add 6 to the above planger and get a new planger. For example, if 6 is joined to 1, 3, 5, then one such new planger would be as shown. Hence the situation is a planger, as required. The difficulty in this case of six nodes will show you that the case of, say, 317 nodes is non-trivial!

Let us now pass on to a different problem. In between wanting to be an engine-driver and constructing non-intersecting pipes, I was trying to draw Figure 5 without going over the same stretch of lines twice. In other words I was asking

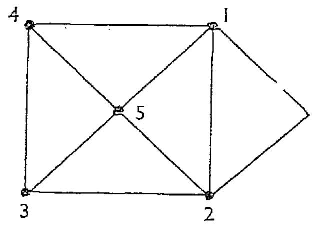


Figure 5

whether this diagram constitutes a planger. A child of five would be able to show you that it does (if you do not know any children of five, then reference to this type of situation will be found in reference 2). In Figure 5, 1, 2 and 5 have an even number of lines leading away from them, but 3 and 4 have an odd number (call them 'odd nodes'). Similarly in the earlier diagram (Figure 2) of a planger if we remove the two 'tails' we obtain a planger in which the starting point and finishing point are nodes (called a 'proper planger') in which 3 and 5 are the only odd nodes. In fact in any proper planger no nodes except perhaps the starting-point and finishing-point will be odd nodes. This is because whenever your pencil-point approaches such a node, it also has to leave it. If the starting-point and finishing-point are different, then they will be odd nodes, but if they coincide it will be an even node. Thus a proper planger has exactly two or no odd nodes.

The world of mathematics is full of lovely surprises, and amongst the loveliest is the fact that the opposite (or converse) of this result is true. In other words if you give me any situation like Figure 5, then to decide whether it is a planger I must simply count the odd nodes: if there are two or none of them, then the figure is a planger. Furthermore, in the case where there are no odd nodes the planger can be chosen with the same starting-point and finishing-point (called a 'closed planger'). I shall sketch a proof of this last result. Take the given situation with no odd nodes and draw a closed planger within it (an example of the construction is shown in Figure 6). Now remove this planger and repeat the process in the remaining situation. Continuing in this way will eventually exhaust the whole situation, and the collection can be rejoined to give one closed planger as shown.

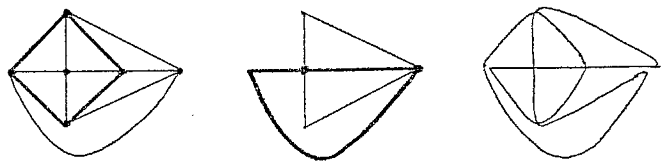


Figure 6a. Given situation with no odd nodes, construct closed planger.

Figure 6b. Remove first planger and construct a second. The remainder is now also a closed planger.

Figure 6c. Put the three closed plangers together to give one as shown, which is the original situation.

In each of the two above problems we have noticed that whenever a certain situation X occurred some fairly trivial consequence Y also occurred. But then we discovered in each case, to our surprise, that the property Y was enough to guarantee X happening. Not all mathematics is as neat. For example, you will notice that, by the definition, whenever a planger is constructed a pencil and paper is used. However, the use of pencil and paper does not guarantee the construction of a planger (if you cannot see why, think of handwriting).

Another point about the structure of the two given problems is worth mentioning: with one mathematical study we have solved two rather different problems from our childhood. This is a basic philosophy of modern mathematics: to find underlying structures, perhaps simple in themselves, and yet covering many different situations. One day someone may define a 'Super-planger' which covers all the mathematical structures in which one is ever interested. Perhaps then I shall have to become an engine-driver again.

References

1. *Beano* (July, 1950).
2. O. Ore, *Graphs and their Uses* (Random House, New York, 1963).

Infinities

CATHERINE V. SMALLWOOD
Polytechnic of North London

Most readers of *Mathematical Spectrum* will have an intuitive notion of finite and infinite sets and the distinction between them. We shall see that it is not difficult to make these ideas precise. The means used for this also enable one to show that some infinite sets are, so to speak, more infinite than others.

First let us consider methods of comparing finite sets. Given any two such sets, we can always decide whether they are of equal size or whether one set is larger than the other. However, how do we come to such a decision?

Take, for example, the set of people in a cinema audience and the set of seats provided for them. One method of comparing the sizes of the two sets is to count the number of elements in each and then compare the two numbers. An alternative but equally effective method is to observe whether there are empty seats or whether there are people with no place to sit. If everyone can be seated, leaving no empty seats, then the set of people and set of seats are of the same size. In this case we have established a 'one–one correspondence' between the set of people and the set of seats: each person occupies one and only one seat and each seat has one and only one person sitting in it.

Obviously when comparing infinite sets we cannot count the total number of elements in each set; but we can try to set up a one–one correspondence between them. The concept of such a correspondence is the basis of our discussion and we define it formally.

Definition. Suppose that the elements of two sets A, B can be grouped in pairs (a, b) (a in A, b in B) in such a way that all elements of A and B appear once and once only. This pairing is called a *one–one correspondence* between A and B; the sets A, B are said to be *similar*; and we write

$$A \sim B.$$

Clearly similarity has the following properties:

(i) $A \sim A$ (reflexivity),

(ii) if $A \sim B$, then $B \sim A$ (symmetry),

(iii) if $A \sim B$ and $B \sim C$, then $A \sim C$ (transitivity).

We now say that a set A is *finite* and has k elements if there exists a natural number k such that $A \sim N_k = \{1, 2, \ldots, k\}$. Any set which is not finite is called *infinite*.

Every set N_k is finite (since $N_k \sim N_k$) and we can give many examples of finite sets. However, do infinite sets exist? The simplest example of such a set is $N = \{1, 2, 3, \ldots\}$. To prove that N is infinite we argue by contradiction.

Suppose that N is finite, so that there is a one–one correspondence between N and a set N_k. Under this correspondence let $n_1, \ldots, n_k$ be the elements of N which correspond to the elements $1, \ldots, k$ respectively of N_k. Put

$$n = \max(n_1, \ldots, n_k).$$

Then $n+1$ belongs to N, but $n+1$ is not one of $n_1, n_2, \ldots, n_k$. Thus the correspondence

$$(1, n_1), \ldots, (k, n_k)$$

is not a one–one correspondence between N_k and N and we have the required contradiction. Hence N is infinite.

The set N and sets similar to N appear so frequently that it is worth coining a special name for them.

Definition. A set which is similar to N is called *countably infinite.* (A set which is finite or countably infinite is called *countable.*)

In fact a countably infinite set is one whose elements can be arranged in an infinite sequence, so that we have a first element, a second, a third, etc.; and these elements may then be denoted by $a_1, a_2, a_3, \ldots$, respectively. For such an arrangement establishes a one–one correspondence between the set and N, with the element corresponding to the number n being written a_n. It follows that the set $E = \{2, 4, 6, \ldots\}$ of even natural numbers, which appears to be 'half the size' of N is in fact similar to N. (In other words there are 'as many' even natural numbers as there are natural numbers.)

On the other hand there are many sets which at first sight appear to be larger than N, but are nevertheless countable. For instance the set of all integers is countably infinite, since it can be arranged as the sequence

$$0, 1, -1, 2, -2, 3, -3, \ldots.$$

It is even more surprising that the set Q of rational numbers is countable. First consider the positive rational numbers. All those appear in the doubly infinite array

$$\begin{array}{llllll}
\frac{1}{1}, & \frac{2}{1}, & \frac{3}{1}, & \frac{4}{1}, & \frac{5}{1}, & \ldots \\
\frac{1}{2}, & \frac{2}{2}, & \frac{3}{2}, & \frac{4}{2}, & \frac{5}{2}, & \ldots \\
\frac{1}{3}, & \frac{2}{3}, & \frac{3}{3}, & \frac{4}{3}, & \frac{5}{3}, & \ldots \\
\frac{1}{4}, & \frac{2}{4}, & \frac{3}{4}, & \frac{4}{4}, & \frac{5}{4}, & \ldots \\
\frac{1}{5}, & \frac{2}{5}, & \frac{3}{5}, & \frac{4}{5}, & \frac{5}{5}, & \ldots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \ldots
\end{array}$$

If we now write down the numbers in the successive diagonals, leaving out any number (such as 2/2) which has previously occurred in reduced form, we obtain the sequence

$$\tfrac{1}{1}, \tfrac{1}{2}, \tfrac{2}{1}, \tfrac{1}{3}, \tfrac{3}{1}, \tfrac{1}{4}, \tfrac{2}{3}, \tfrac{3}{2}, \tfrac{4}{1}, \ldots. \qquad (1)$$

This sequence consists of all the positive rational numbers, and the set of these numbers is therefore countably infinite. Finally we note that, if the sequence (1) is denoted by $r_1, r_2, \ldots$, then the set Q of *all* rational numbers (positive, negative and zero) can be arranged as the sequence

$$0, r_1, -r_1, r_2, -r_2, \ldots.$$

Let us now return to the question of the relative size of infinite sets. We shall show that the set R of real numbers is not countable, i.e., that it is not similar to N. Since N is similar to a subset of R (for instance N itself, or Q) R is, in a natural sense, larger than N. It is convenient first to consider the interval $(0, 1)$.[1]

Theorem. The set of real numbers in the interval $(0, 1)$ *is not countable.*

Proof. Suppose that the real numbers in $(0, 1)$ do form a countable set. This set is infinite and so countably infinite, so that its members can be arranged in an infinite sequence

$$a_1, a_2, a_3, \ldots.$$

Now represent each a_n as an infinite decimal

$$a_n = 0{\cdot}\alpha_{n1}\,\alpha_{n2}\,\alpha_{n3}\ldots,$$

say, in which recurring 9's are not used. Such a representation is unique.

We now form a new number

$$b = 0{\cdot}\beta_1\,\beta_2\,\beta_3\ldots,$$

where $\beta_1, \beta_2, \ldots$ are integers between 1 and 8 chosen so that $\beta_1 \neq \alpha_{11}$, $\beta_2 \neq \alpha_{22}$, etc. Then b is a real number in the interval $(0, 1)$ which differs from a_n in the nth decimal place. Since no β_n is 9, b is not one of the numbers $a_1, a_2, a_3, \ldots$ and this contradicts the original assumption that all real numbers in $(0, 1)$ appear in the sequence $a_1, a_2, a_3, \ldots$.

It is quite easy to show that all open intervals, finite or infinite, are similar. For example the formula

$$y = \frac{x - \frac{1}{2}}{x(1-x)} \quad (0 < x < 1)$$

establishes a one–one correspondence between $(0, 1)$ and R. Thus R is not countable.

[1] The *open* interval (a,b) is the set of x in R such that $a < x < b$, the *closed* interval $[a,b]$ is the set of x such that $a \leqslant x \leqslant b$. The half-open intervals $[a,b)$ and $(a,b]$ are defined by analogous inequalities.

In general

$$(a, b) \sim [a, b) \sim [a, b] \tag{2}$$

and, in fact, all intervals (of any kind) are similar. It is curious that the proof of (2) is by no means obvious. We shall return to the problem after the corollary to the next theorem.

Irrational numbers are met so much less frequently than rational ones that it may come as a surprise to discover that there are actually many more of them. This follows quite easily from the uncountability of R. For suppose that the set I of irrational numbers is countable. Then the members of I can be arranged as a sequence

$$s_1, s_2, s_3, \ldots.$$

Since Q can certainly be represented by a sequence

$$r_1, r_2, r_3, \ldots,$$

we should therefore find that all real numbers occurred in the sequence

$$r_1, s_1, r_2, s_2, r_3, s_3, \ldots.$$

But this is impossible, since R is not countable, and so I cannot be countable.

We have now come across several examples of infinite sets which are similar to 'proper' subsets of themselves (i.e., subsets which do not contain all the elements of the original set). This feature is in fact characteristic of all infinite sets. First we prove that countably infinite sets are the 'smallest' infinite sets.

Theorem. Every infinite set contains a countably infinite subset.

Proof. Let A be an infinite set. Take some element of A and call it a_1. Then the set $A - \{a_1\}$ consisting of all elements of A except a_1 is not empty; denote one of its elements by a_2. This process may be continued indefinitely. For at the nth stage the set

$$A - \{a_1, a_2, \ldots, a_{n-1}\},$$

which contains the members of A other than $a_1, a_2, \ldots, a_{n-1}$, is not empty, since A is not finite. The set $\{a_1, a_2, \ldots\}$ is a countably infinite subset of A.

Corollary. If A is an infinite set and a_1 is any element of A, then A is similar to $A - \{a_1\}$.

Proof. The set A contains a countably infinite subset S. From the proof of the last theorem it is clear that we can take a_1 to be a member of S and that S may be written $\{a_1, a_2, a_3, \ldots\}$. The function f defined by

$$f(x) = \begin{cases} a_{n+1} & \text{if } x = a_n \quad (n = 1, 2, \ldots), \\ x & \text{if } x \text{ belongs to } A - S \end{cases}$$

establishes a one–one correspondence between A and $A - \{a_1\}$. Thus $A \sim A - \{a_1\}$.

The corollary shows that any infinite set is similar to a proper subset of itself. It is a tedious, though not difficult, exercise to show (by induction) that a finite set cannot be similar to a proper subset of itself. Hence a set is infinite if and only if it has a proper subset to which it is similar. This condition is often used as an alternative method of defining infinite sets.

Another immediate consequence of the corollary is the fact, noted earlier, that $[a,b] \sim [a,b) \sim (a,b)$.

It is also worth observing that an argument of the kind used in proving the corollary can be used to show that I is not only uncountable, but is actually similar to R.

So far we have only met two kinds of infinite sets, those similar to N and those similar to R. But the next theorem shows that there are many more. In it we consider the set of all subsets of a given set. For instance the set $\{1,2,3\}$ has the following eight subsets:

$$\text{the empty set, } \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}.$$

It is easy to see that a set with n elements has 2^n subsets and that $2^n > n$ for every natural number n. Thus a finite set has more subsets than elements. Our final result generalises this proposition.

Theorem. If A is any set and Σ is the set of all subsets of A (including the empty set and A itself), then A is not similar to Σ.

Proof. Suppose that $A \sim \Sigma$, so that there is a one–one correspondence between the elements x of A and the elements S of Σ (which are subsets of A). Given x in A, denote the subset of A to which it corresponds (under the above one–one correspondence) by S_x. Now define the subset S^* of A as follows:

$$x \text{ belongs to } S^* \text{ if and only if } x \text{ does not belong to } S_x. \tag{3}$$

Since S^* is an element of Σ, there is an a in A such that

$$S^* = S_a.$$

Then (3), with $x = a$, becomes the contradictory statement a belongs to $S^* = S_a$ if and only if a does not belong to $S_a = S^*$; and the proof is complete.

It follows from the last theorem that there is no largest infinite set and that there are not just one or two infinities, but infinitely many.

The theory, of which we have described the beginning, is the creation of Georg Cantor (1845–1918) and was expounded in a remarkable series of papers starting in 1874. Cantor's ideas, in their novelty, aroused fierce contemporary opposition, but they can now be seen to have been a major influence in modern mathematics.

Time Independent Quantum Mechanics in One Dimension

D. J. ROAF
Exeter College, Oxford

1. Introduction

Quantum mechanics is thought of as a difficult and advanced subject, to be taught as the culmination of a degree course. There are, however, many ideas in quantum mechanics which are quite simple: this article will consider some of them.

To keep the discussion simple, we shall restrict it to one-dimensional time-independent problems. The only mathematical background needed will be the ideas of complex numbers, second order differential equations and elementary probability. Knowledge of the eigenvalues and eigenvectors of matrices will help.

In classical mechanics in one dimension, if we know the total energy E of a particle of mass m and how its potential $V(x)$ varies with its position x, we can calculate its kinetic energy T (from the equation $T+V=E$) and hence its speed (from $T=\frac{1}{2}m(dx/dt)^2$) and its subsequent motion.

In quantum mechanics, things do not appear so precise; we can only talk about probabilities, but we can make exact statements about them. The particle's motion is described by its wave-function $\psi(x)$; the probability of finding the particle between x and $x+\delta x$ is $|\psi(x)|^2\delta x$. The wave-function may be real or complex. Because we are talking about probabilities, $\int_{-\infty}^{\infty}|\psi(x)|^2dx=1$ (the particle must be *somewhere*). The mean value of its position x is, therefore, $\int_{-\infty}^{\infty}\psi^*(x).x\psi(x)\,dx = \int_{-\infty}^{\infty}x|\psi(x)|^2dx$ (with $\psi^*=u-iv$ the complex conjugate of $\psi=u+iv$, where u and v are real). We write the mean value in angular brackets as $\langle x\rangle = \int_{-\infty}^{\infty}\psi^*x\psi\,dx$.

Any *observable* (physical quantity) A, has an *operator* A which enables us to find its mean value. Thus

$$\langle A\rangle = \int_{-\infty}^{\infty}\psi^*A\psi\,dx. \tag{1.1}$$

For the observable x the operator is x itself. But for the momentum p the operator is $\boldsymbol{p}=-i\hbar d/dx$ where $\hbar=h/2\pi$ and h is Planck's Constant (which is $6{\cdot}55\times10^{-34}$ joule-sec in S.I. units). Thus

$$\langle p\rangle = -i\hbar\int_{-\infty}^{\infty}\psi^*\frac{d\psi}{dx}dx. \tag{1.2}$$

For the kinetic energy $T=p^2/2m$, the operator is $\boldsymbol{T}=(1/2m)(-i\hbar d/dx)^2 = (-\hbar^2/2m)\,d^2/dx^2$, and $\langle T\rangle = (-\hbar^2/2m)\int\psi^*(d^2\psi/dx^2)\,dx$. We can find similar expressions for other observables.

Now in classical mechanics we have $T+V=E$ where $T=$ kinetic energy, $V=$ potential energy and $E=$ total energy. The operators corresponding to T, V and E are $(-\hbar^2/2m)(d^2/dx^2)$, $V(x)$ and E respectively.

The wave-function $\psi(x)$ is the solution of the Schrödinger equation

$$\frac{-\hbar^2}{2m}\frac{d^2\psi}{dx^2}+V(x)\psi(x)=E\psi(x). \tag{1.3}$$

We shall see that only for certain values of E (eigenvalues) are there satisfactory wave-functions (eigenfunctions) $\psi(x)$. We require ψ and $d\psi/dx$ to be continuous everywhere (except at points where $V(x)$ has infinite discontinuities where only ψ need be continuous). We also require ψ to tend to zero as x tends to $\pm\infty$, and $\int_{-\infty}^{\infty}|\psi|^2 dx=1$.

The justification for this peculiar theory is that it works.

2. The infinite square well

Let us look at a simple example called the *infinite square well.* Suppose

$$V=0 \quad \text{for } |x|<a, \quad \text{and} \quad V=\infty \quad \text{for } |x|>a.$$

In classical mechanics the particle would oscillate between $x=-a$ and $x=+a$; it could have any energy, and an equal probability of being found anywhere. In quantum mechanics, for $|x|>a$ the Schrödinger equation can only be satisfied by $\psi=0$. For $|x|<a$ we have

$$\frac{d^2\psi}{dx^2}=\frac{-2mE}{\hbar^2}\psi(x) \tag{2.1}$$

which has solution

$$\psi=Ae^{i\alpha x}+Be^{-i\alpha x},$$

where $\alpha^2=2mE/\hbar^2$ and A and B are arbitrary constants. But $\psi(-a)=0=\psi(a)$ (ψ is continuous at $x=-a$ and $x=a$), hence

$$Ae^{i\alpha a}+Be^{-i\alpha a}=0=Ae^{-i\alpha a}+Be^{i\alpha a},$$

so that $A=-Be^{2i\alpha a}=-Be^{-2i\alpha a}$ or $e^{4i\alpha a}=1$. Thus $\alpha=n\pi/2a$ for some integer n and $A=-Be^{in\pi}$. If n is odd, then $A=B$ and $\psi=2A\cos(n\pi x/2a)$. If n is even, then $A=-B$ and $\psi=2iA\sin(n\pi x/2a)$.

The value of $|A|$ can be obtained by integrating $\int_{-\infty}^{\infty}|\psi|^2 dx=1$, so for n odd we have $4|A|^2\int_{-a}^{a}\cos^2(n\pi x/2a)\,dx=4|A|^2a=1$. We will choose a real solution if possible; if n is odd we have for $|x|<a$,

$$\psi=a^{-\frac{1}{2}}\cos(n\pi x/2a) \quad \text{and} \quad E=n^2\pi^2\hbar^2/8ma^2. \tag{2.2}$$

For n even

$$\psi=a^{-\frac{1}{2}}\sin(n\pi x/2a) \quad \text{and} \quad E=n^2\pi^2\hbar^2/8ma^2. \tag{2.3}$$

Notice that the particle cannot have zero total energy since $n=0$ corresponds to $\psi=0$ everywhere so $\int_{-\infty}^{\infty}|\psi|^2 dx\neq 1$.

For the infinite square well ground state (i.e., state with lowest energy, $n = 1$) we have

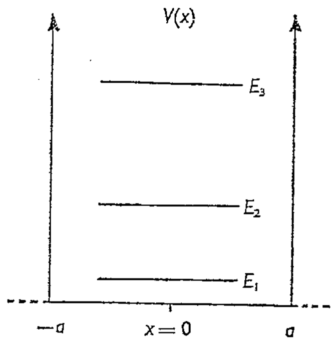
$$\langle x\rangle = (1/a)\int_{-a}^{a} \cos(\pi x/2a)\, x \cos(\pi x/2a)\, dx = 0,$$

$$\langle p\rangle = (1/a)\int_{-a}^{a} \cos(\pi x/2a)\, (i\hbar\pi/2a)\sin(\pi x/2a)\, dx = 0, \tag{2.4}$$

$$\langle p^2\rangle = (1/a)\int_{-a}^{a} \cos(\pi x/2a)\, (\hbar^2\pi^2/4a^2)\cos(\pi x/2a)\, dx = \hbar^2\pi^2/4a^2,$$

so that the mean kinetic energy is

$$\langle T\rangle = \hbar^2\pi^2/8ma^2 = E.$$



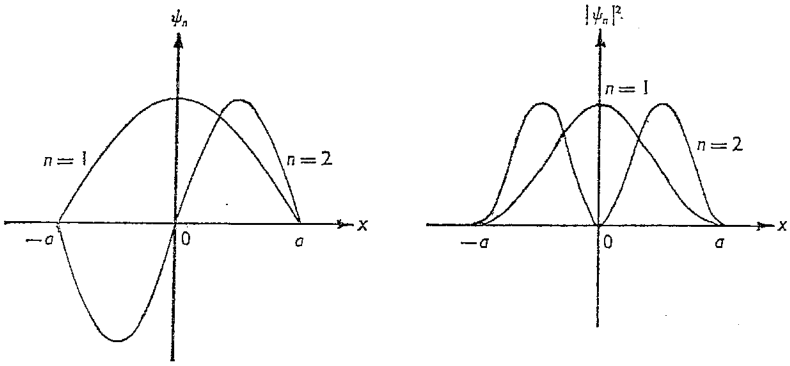


Figure 1. The infinite square well.

3. Eigenvalues and eigenvectors in matrix theory

The eigenvalues and eigenfunctions of Schrödinger's equation are similar to the eigenvalues and eigenvectors of matrices (also called characteristic values and vectors). Those who know nothing about these may ignore the rest of this section; those who know some matrix theory will be reminded of some elementary results so that the similarities with the next section can be better appreciated.

Firstly, if we have a matrix A (for simplicity 2×2) and a real non-zero vector $\boldsymbol{k}$ we can calculate

$$Ak = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} a_{11}k_1 + a_{12}k_2 \\ a_{21}k_1 + a_{22}k_2 \end{pmatrix}. \tag{3.1}$$

This new vector will, in general, be in a different direction from $\boldsymbol{k}$. But if it is in the same (or opposite) direction then $A\boldsymbol{k} = \lambda\boldsymbol{k}$ for some λ. In this equation λ is called an eigenvalue, $\boldsymbol{k}$ an eigenvector. Now suppose A is real and symmetric ($a_{21} = a_{12}$) and suppose that

$$(x \quad y) \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} > 0 \tag{3.2}$$

for all real non-zero x, y. This property makes A *positive-definite*. Then for a non-zero $\boldsymbol{k}$, with transpose (row) vector $\tilde{\boldsymbol{k}}$,

$$\lambda\tilde{k}k = \tilde{k}Ak > 0.$$

But $\tilde{k}k > 0$, so λ is real and positive.

Secondly, if $Ak_1 = \lambda_1 k_1$ and $Ak_2 = \lambda_2 k_2$, where $\lambda_1 \neq \lambda_2$, then $\tilde{k}_1 k_2 = 0$. For, transposing $Ak_1 = \lambda_1 k_1$,

$$\begin{aligned} \lambda_1 \tilde{k}_1 &= \tilde{k}_1 A, \\ \lambda_1 \tilde{k}_1 k_2 &= (\tilde{k}_1 A) k_2 = \tilde{k}_1 (Ak_2) = \tilde{k}_1 \lambda_2 k_2 = \lambda_2 (\tilde{k}_1 k_2). \end{aligned} \tag{3.3}$$

Hence if $\lambda_1 \neq \lambda_2$, $\tilde{k}_1 k_2 = 0$.

Thirdly, if $Ak_1 = \lambda_1 k_1$ and $Ak_2 = \lambda_2 k_2$ where $\lambda_1 = \lambda_2$, then $k = \alpha k_1 + \beta k_2$ is also an eigenvector with the same eigenvalue. For,

$$Ak = A\alpha k_1 + A\beta k_2 = \alpha\lambda_1 k_1 + \beta\lambda_2 k_2 = \lambda_1(\alpha k_1 + \beta k_2) = \lambda_1 k. \tag{3.4}$$

4. Eigenvalues and eigenfunctions of Schrödinger's equation

Let us look at some properties of Schrödinger's equation.

Firstly, if $V(x) > 0$ for all x then $E > 0$ for any non-zero solution $\psi(x)$, where E must be real. For, consider the complex conjugate $\psi^*(x)$ multiplied by the equation (1.3), so that

$$\frac{-\hbar^2}{2m}\psi^* \frac{d^2\psi}{dx^2} + V(x)\psi^*\psi = E\psi^*\psi. \tag{4.1}$$

Integrating over the range $(-\infty, \infty)$, we have

$$-\int_{-\infty}^{\infty} \psi^* \frac{d^2\psi}{dx^2} dx = -\left[\psi^* \frac{d\psi}{dx}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{d\psi^*}{dx} \cdot \frac{d\psi}{dx} dx = \int_{-\infty}^{\infty} \left|\frac{d\psi}{dx}\right|^2 dx, \tag{4.2}$$

which is positive since the integrand is everywhere positive (remember that as $x \to \pm\infty$, $\psi^* \to 0$). To simplify the printing, we shall omit the limits $\pm\infty$ for the rest of this section.

Now $\int V\psi^*\psi\,dx > 0$ since the integrand is positive everywhere, and

$$\int E\psi^*\psi\,dx = E = \frac{\hbar^2}{2m}\int\left|\frac{d\psi}{dx}\right|^2 dx + \int V\psi^*\psi\,dx. \tag{4.3}$$

Hence $E > 0$.

Secondly, if there are two non-zero wave-functions ψ_1, ψ_2 corresponding to different total energies E_1, E_2, then $\int \psi_1^*\psi_2\,dx = 0$. For,

$$E_2\psi_2 = \frac{-\hbar^2}{2m}\frac{d^2\psi_2}{dx^2} + V\psi_2,$$

so that

$$\begin{aligned} E_2\int\psi_1^*\psi_2\,dx &= \frac{-\hbar^2}{2m}\int\psi_1^*\frac{d^2\psi_2}{dx^2}dx + \int\psi_1^* V\psi_2\,dx \\ &= \frac{-\hbar^2}{2m}\left[\psi_1^*\frac{d\psi_2}{dx}\right] + \frac{\hbar^2}{2m}\int\frac{d\psi_1^*}{dx}\frac{d\psi_2}{dx}dx + \int\psi_1^* V\psi_2\,dx \\ &= 0 + \frac{\hbar^2}{2m}\left[\frac{d\psi_1^*}{dx}\psi_2\right] - \frac{\hbar^2}{2m}\int\frac{d^2\psi_1^*}{dx^2}\psi_2\,dx + \int\psi_1^* V\psi_2\,dx \qquad (4.4) \\ &= 0 + \int\left\{\frac{-\hbar^2}{2m}\frac{d^2\psi_1^*}{dx^2} + V\psi_1^*\right\}\psi_2\,dx \\ &= \int E_1^*\psi_1^*\psi_2\,dx = E_1^*\int\psi_1^*\psi_2\,dx. \end{aligned}$$

But $E_1^* = E_1$, so either $E_1 = E_2$ or $\int\psi_1^*\psi_2\,dx = 0$.

Thirdly, as the equation is linear, if $E_1 = E_2$ then $\psi = \alpha\psi_1 + \beta\psi_2$ is a solution of the equation with the same eigenvalue. (Substitute to confirm this.)

5. The harmonic oscillator

Another simple example is the *harmonic oscillator*

$$V = \tfrac{1}{2}kx^2 \quad (k > 0).$$

This is the physical potential for a spring, or for a vibrating molecule. Classically we have $m\,d^2x/dt^2 + kx = 0$, i.e., simple harmonic motion of angular frequency $(k/m)^{\frac{1}{2}}$ and *any* total energy E; but the particle can only be found in the region $|x| \leqslant (2E/k)^{\frac{1}{2}}$ and it will be more likely to be found near the extreme points than near the centre ($x = 0$) since it moves faster at the centre.

The Schrödinger equation (1.3) becomes

$$\frac{-\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \tfrac{1}{2}kx^2\psi = E\psi, \tag{5.1}$$

and can be solved by numerical techniques for any arbitrary value of E (starting from, say, $\psi = 0$, $d\psi/dx = 1$ at $x = 0$ or from $\psi = 1$, $d\psi/dx = 0$ at $x = 0$). If one does this it will turn out that for most values of E, $\psi \to \infty$ as $x \to \infty$. But for some values of E (the eigenvalues) we will find that $\psi \to 0$ as $x \to \infty$ and $\int |\psi|^2 dx$ will be finite. We can then rescale ψ to make $\int |\psi|^2 dx = 1$.

Another approach is to notice that for large values of x the equation (5.1) becomes approximately

$$\frac{d^2\psi}{dx^2} = \frac{mkx^2\psi}{\hbar^2}, \tag{5.2}$$

and that this has the approximate solution $\psi = e^{-\frac{1}{2}\alpha x^2}$ (where $\alpha^2 = mk/\hbar^2$). If we then try a solution

$$\psi = u(x)\, e^{-\frac{1}{2}\alpha x^2}$$

we can obtain for $u(x)$ the differential equation

$$\frac{d^2u}{dx^2} - 2\alpha x \frac{du}{dx} - \alpha u = -\frac{2mEu}{\hbar^2}. \tag{5.3}$$

This has a series solution

$$u(x) = x^c(a_0 + a_1 x + a_2 x^2 + \dots),$$

and by requiring that $u(x)\, e^{-\frac{1}{2}\alpha x^2} \to 0$ as $x \to \infty$ we can restrict the values of E to the eigenvalues.

In a large number of quantum mechanical problems we are more interested in the energy eigenvalues than in the wave-functions. For example a spectroscopist who measures the frequencies of the light emitted by an atom is really measuring the changes in energy of an electron as it moves from one state to another; for if the change in energy is ΔE, we know that $\Delta E = \hbar\nu$, where ν is the angular frequency of the light. Sometimes it is necessary to solve for the wave-functions in order to find the energy eigenvalues, but there are some problems where this can be avoided. There is an interesting technique for the harmonic oscillator which is also applicable in some other problems.

Suppose we make the substitution $y = (km/\hbar^2)^{\frac{1}{4}} x$ so that the equation (5.1) becomes

$$\frac{-\hbar^2}{2m}\left(\frac{km}{\hbar^2}\right)^{\frac{1}{2}} \frac{d^2\psi}{dy^2} + \tfrac{1}{2}k\left(\frac{\hbar^2}{km}\right)^{\frac{1}{2}} y^2\psi = E\psi. \tag{5.4}$$

Then writing $(2E/\hbar)(m/k)^{\frac{1}{2}} = F$ we have to solve

$$-\frac{d^2\psi}{dy^2} + y^2\psi = F\psi.$$

Now

$$\begin{aligned} y\left(+\frac{d}{dy}\right)\left(y - \frac{d}{dy}\right)\psi &= y^2\psi - y\frac{d\psi}{dy} + \frac{d(y\psi)}{dy} - \frac{d^2\psi}{dy^2} \\ &= \left(y^2 - \frac{d^2}{dy^2} + 1\right)\psi = (F+1)\psi \end{aligned} \tag{5.5}$$

if ψ is an eigenfunction with eigenvalue F. And similarly $(y-d/dy)(y+d/dy)\psi = (F-1)\psi$. Now these operators are associative, so if $u = (y+d/dy)(y-d/dy)(y+d/dy)\psi$ we can, multiplying the last two brackets first, write this as

$$u = \left(y+\frac{d}{dy}\right)(F-1)\psi = (F-1)\left(y+\frac{d}{dy}\right)\psi$$

or, multiplying the first two brackets first, as

$$u = \left(y^2-\frac{d^2}{dy^2}+1\right)\left(y+\frac{d}{dy}\right)\psi.$$

Hence

$$\left(y^2-\frac{d^2}{dy^2}\right)\left\{\left(y+\frac{d}{dy}\right)\psi\right\} = (F-2)\left\{\left(y+\frac{d}{dy}\right)\psi\right\},$$

so if ψ is an eigenfunction with eigenvalue F, $\{(y+d/dy)\psi\}$ is an eigenfunction with eigenvalue $F-2$.

Now we come to the interesting point. We know that all eigenvalues are positive, since $V(x) > 0$, as we saw in Section 4. We know that if ψ is the eigenfunction corresponding to F, $(y+d/dy)\psi$ is the eigenfunction corresponding to $(F-2)$, and similarly $(y-d/dy)\psi$ is the eigenfunction corresponding to $(F+2)$.

So given any particular F which is an eigenvalue, we find that $(F-2)$ is an eigenvalue, and so are $(F-4)$, $(F-6)$, etc. But this will lead to negative eigenvalues!

We can avoid this only if $(y+d/dy)\psi \equiv 0$ for some wave-function ψ_0 with eigenvalue F_0, so that $(y+d/dy)\psi_0$ is a trivial (zero) eigenfunction. Then

$$\left(y-\frac{d}{dy}\right)\left(y+\frac{d}{dy}\right)\psi_0 = 0,$$

but

$$\left(y-\frac{d}{dy}\right)\left(y+\frac{d}{dy}\right)\psi_0 = (F_0-1)\psi_0,$$

so we must have $F_0 = 1$, so that $E_0 = \frac{1}{2}\hbar(k/m)^{\frac{1}{2}}$. The other eigenvalues are then 3, 5, 7, etc., so the energy-levels are

$$E_n = \tfrac{1}{2}(2n+1)\hbar(k/m)^{\frac{1}{2}} \quad (n \text{ integral} \geqslant 0).$$

If we want to, we can find ψ_0, for, since $(y+d/dy)\psi_0 = 0$,

$$x\psi_0+(km/\hbar^2)^{-\frac{1}{2}}\frac{d\psi_0}{dx} = 0 \tag{5.6}$$

which has solution $\psi_0 = Ae^{-\frac{1}{2}(km)^{\frac{1}{2}}x^2/\hbar}$ or $\psi_0 = Ae^{-\frac{1}{2}\alpha x^2}$, where A is a constant, and $\alpha = (km)^{\frac{1}{2}}/\hbar$. The other solutions are then

$$\psi_1 \propto \left(y-\frac{d}{dy}\right)\psi_0 \propto \left(\frac{d}{dx}-\alpha x\right)Ae^{-\frac{1}{2}\alpha x^2}$$

so $\psi_1 = Bxe^{-\frac{1}{2}\alpha x^2}$, where B is a constant, with $E_1 = \frac{3}{2}\hbar(k/m)^{\frac{1}{2}}$ and

$$\psi_n \propto \left(y - \frac{d}{dy}\right)^n \psi_0 \quad \text{with} \quad E_n = \tfrac{1}{2}(2n+1)\,\hbar(k/m)^{\frac{1}{2}}.$$

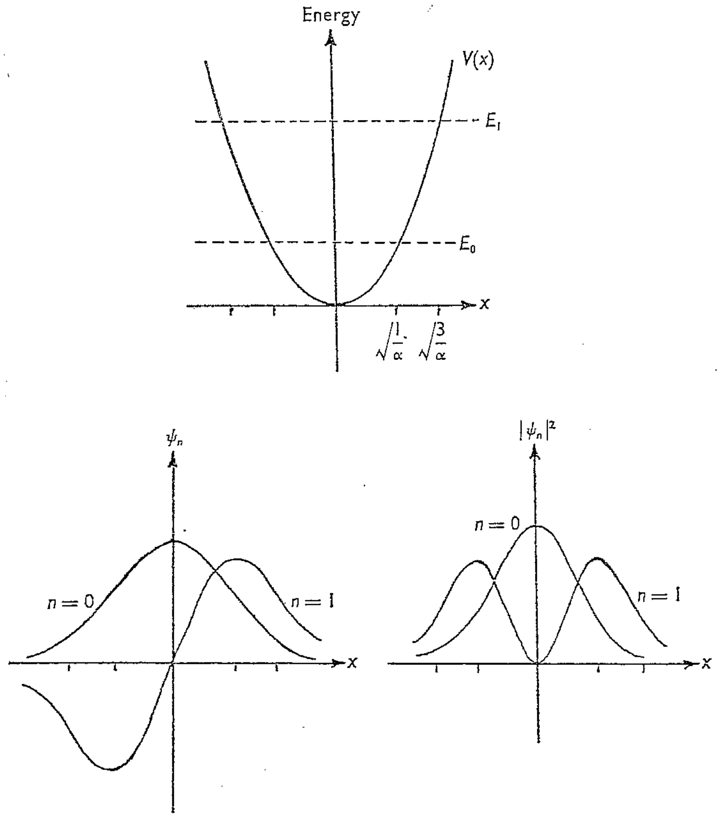


Figure 2. The harmonic oscillator.

It will be noticed that these solutions give a finite probability of the particle being found outside the classical limits, $x = \pm(2E/k)^{\frac{1}{2}} = \pm[(2n+1)/\alpha]^{\frac{1}{2}}$.

6. The uncertainty principle

Any particular measurement must give a definite result. But if a set of measurements is taken on systems with a particular wave-function, then, because of the probability implications of the wave-function, we may obtain differing results from our measurements. A measure of the scatter or *uncertainty* of these measurements is the *standard deviation*; this is the square-root of the mean of the squares of the differences between the measurements and their mean.

Given a wave-function ψ, the mean value of the observable A whose operator is A is $\langle A\rangle = \int\psi^* A\psi\, dx$. All integrals in this section are over the range $(-\infty, \infty)$. The difference between A and $\langle A\rangle$ will have operator $A - \langle A\rangle$ so that its square will be $(A - \langle A\rangle)^2$. Thus the mean of the squares of the differences will be

$$(\Delta A)^2 = \int \psi^*[A - \langle A\rangle]^2\psi\, dx. \tag{6.1}$$

Now operators corresponding to real observables are always *Hermitian*; that is $\int \chi^* A\phi\,dx = \int (A\chi)^*\phi\,dx$ for all functions χ and ϕ. For example, for the observable x we have $\int \chi^* x\phi\,dx = \int (x\chi)^*\phi\,dx$; and for p we have

$$\begin{aligned}\int \chi^*\left(-i\hbar\frac{d\phi}{dx}\right)dx &= -i\hbar\chi^*\phi + i\hbar\int \frac{d\chi^*}{dx}\phi\,dx \\ &= \int\left(-i\hbar\frac{d\chi}{dx}\right)^*\phi\,dx,\end{aligned} \tag{6.2}$$

provided ϕ and χ tend to zero as x tends to $\pm\infty$.

Now $(A-\langle A\rangle)$ is Hermitian if A is, so

$$\begin{aligned}(\Delta A)^2 &= \int \psi^*[A-\langle A\rangle]^2\psi\,dx \\ &= \int [(A-\langle A\rangle)\psi]^*[(A-\langle A\rangle)\psi]\,dx \\ &= \int |(A-\langle A\rangle)\psi|^2 dx.\end{aligned} \tag{6.3}$$

Thus $(\Delta A)^2 = 0$ if and only if $(A-\langle A\rangle)\psi = 0$, or $A\psi = \langle A\rangle\psi$. So unless ψ is an eigenfunction of the operator A there will be uncertainty in the measurement of the observable A.

If we have the two functions χ and ϕ again and if $\int \chi^*\phi\,dx = u+iv$, where u and v are real, then $\int \phi^*\chi\,dx = u-iv$ so that

$$\int (\chi^*\phi - \phi^*\chi)\,dx = 2iv$$

and

$$\left|\int (\chi^*\phi - \phi^*\chi)\,dx\right| = 2v.$$

Suppose $I = \int (\chi+i\lambda\phi)^*(\chi+i\lambda\phi)\,dx$ for a real constant λ. Then

$$\begin{aligned}I &= \int \chi^*\chi\,dx + i\lambda\left[\int (\chi^*\phi - \phi^*\chi)\,dx\right] + \lambda^2\int \phi^*\phi\,dx \\ &= \int \chi^*\chi\,dx - 2\lambda v + \lambda^2\int \phi^*\phi\,dx.\end{aligned} \tag{6.4}$$

But $I \geqslant 0$ for all real λ, since the integrand is non-negative, so the roots (for λ) of the real quadratic equation $I = 0$ are complex or coincident. Hence

$$4\int \chi^*\chi\,dx \int \phi^*\phi\,dx \geqslant 4v^2 = \left|\int (\chi^*\phi - \phi^*\chi)\,dx\right|^2. \tag{6.5}$$

Now if

$$\boldsymbol{\alpha} = x - \langle x\rangle \quad \text{and} \quad \boldsymbol{\beta} = -i\hbar\frac{d}{dx} - \langle p\rangle,$$

$\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are Hermitian operators, and $(\boldsymbol{\alpha\beta} - \boldsymbol{\beta\alpha})\psi = i\hbar\psi$. Then from (6.3)

$$4(\Delta x)^2(\Delta p)^2 = 4\int |\boldsymbol{\alpha}\psi|^2 dx \int |\boldsymbol{\beta}\psi|^2 dx.$$

So writing $\chi = \alpha\psi$, $\phi = \beta\psi$, and using (6.5),

$$4(\Delta x)^2(\Delta p)^2 \geqslant \left| \int (\alpha\psi)^*(\beta\psi) - (\beta\psi)^*(\alpha\psi)\, dx \right|^2.$$

Thus

$$2(\Delta x)(\Delta p) \geqslant \left| \int \psi^*\alpha\beta\psi - \psi^*\beta\alpha\psi\, dx \right|,$$

$$2(\Delta x)(\Delta p) \geqslant \left| \int \psi^*(\alpha\beta - \beta\alpha)\psi\, dx \right|.$$

Hence

$$2(\Delta x)(\Delta p) \geqslant \left| \int \psi^* i\hbar\psi\, dx \right| = \hbar.$$

Thus we cannot specify exactly position and momentum simultaneously: the more precise the position the less precise the momentum.

Let us look again at the infinite square well. We found that in the ground-state, $\psi = a^{-\frac{1}{2}}\cos(\pi x/2a)$ for $|x| < a$ and $\psi = 0$ for $|x| > a$, $\langle x \rangle = \langle p \rangle = 0$ and $\langle p \rangle = \hbar^2\pi^2/4a^2$. Now

$$\langle x^2 \rangle = (1/a) \int_{-a}^{a} \cos(\pi x/2a)\, x^2 \cos(\pi x/2a)\, dx$$

$$= (1/a) \int_{-a}^{a} x^2[\tfrac{1}{2} + \tfrac{1}{2}\cos(\pi x/a)]\, dx = (\pi^2 - 6)\, a^2/3\pi^2. \qquad (6.6)$$

It follows that

$$(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2 = \hbar^2\pi^2/4a^2,$$

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = (\pi^2 - 6)\, a^2/3\pi^2,$$

$$(\Delta p)^2(\Delta x)^2 = \tfrac{1}{12}(\pi^2 - 6)\,\hbar^2 > \tfrac{1}{4}\hbar^2.$$

If we want Δx to be smaller we must choose a smaller, but then Δp is larger. But note that since $\hbar$ is very small the effect of the *uncertainty principle* will only be noticeable for very small particles.

7. Conclusion

This has been a survey of a simple part of quantum mechanics, but the ideas of wave-functions, energy levels and uncertainty are the same in three dimensions and nearly the same in time-dependent systems.

Suggestions for further reading

Some physical background can be found in G. Gamow's *Mr Tompkins in Paperback* (Cambridge, 1965), Chapter 7 and 8, where Gamow discusses the complications we should suffer if h were very large.

Further mathematics can be found in *Introduction to Quantum Mechanics* by P. T. Matthews (McGraw-Hill, New York, 1963).

Seventh British Mathematical Olympiad

Mathematics students may be interested in the problems set for the British Mathematical Olympiad in 1971. The questions were set and marked by Mrs Margaret Hayman and Mr Robert Lyness; this paper has previously been published in the *Science Teacher*, Volume 14, No. 6, and is reproduced here with the permission of the Editor, Mr M. Goldsmith.

1. (a) Factorise $(x+y)^7-(x^7+y^7)$.

   (b) Prove that there is no integer n for which $2n^3+2n^2+2n+1$ is a multiple of 3.
2. $x_1 = 9$ and $x_{r+1} = 9^{x_r}$ for all positive integral r. Prove that the last two digits of x_3, written to base 10, are the same as the last two digits of x_4 written to base 10. What are these two digits?
3. Of a regular polygon of $2n$ sides there are n diagonals which pass through the centre of the inscribed circle. The angles which these diagonals subtend at two given points A and B on the circumference are $a_1, a_2, a_3, \ldots, a_n$ and $b_1, b_2, b_3, \ldots, b_n$. Prove

$$\sum_{i=1}^{n} \tan^2 a_i = \sum_{i=1}^{n} \tan^2 b_i.$$

4. Given a set of $(n+1)$ positive integers, none of which exceeds $2n$, prove by induction or otherwise that at least one member of the set must divide another member of the set.
5. ABC is a triangle whose angles A, B and C are in descending order of magnitude. Circles are drawn such that each circle cuts each side of the triangle internally in two real distinct points. The lower limit to the radii of such circles is the radius of the in-circle of the triangle ABC. Show that the upper limit is not R, the radius of the circumcircle, and find this upper limit in terms of R, A and B.
6. (a) $I(x) = \int_c^x f(x,u)\,du$, where c is a constant. Show why

$$I'(x) = f(x,x) + \int_c^x \frac{\partial f}{\partial x}\,du.$$

   (b) Find $\lim\limits_{\theta=0} \cot\theta \sin(t\sin\theta)$.

   (c) $G(t) = \int_0^t \cot\theta\sin(t\sin\theta)\,d\theta$. Prove $G'(\pi/2) = 2/\pi$.
7. h and k are two real numbers such that $h > k > 0$. Find the probability that two points chosen at random on a straight line of length h should be at a distance of less than k apart.

EITHER

8. A is a 3×2 matrix and B is a 2×3 matrix. The elements of each matrix are numbers. $AB = M$ and $BA = N$. $\det M = 0$ and $\det N \neq 0$. Also $M^2 = kM$ where k is a number.

Determine $\det N$ in terms of k, proving your result.

(You may assume the usual rules for combining matrices but note that k is not a matrix. kC is defined as the matrix whose elements are each k times the corresponding elements of C.)

OR

9. Two uniform solid spheres of equal radii are so placed that one is directly above the other. The bottom sphere is fixed and the top sphere, initially at rest, rolls off. If the coefficient of friction between the two spherical surfaces is μ, show that slipping occurs when

$$2\sin\theta = \mu(17\cos\theta - 10),$$

where θ is the angle the line of centres makes with the vertical. (Moment of inertia of a solid sphere about a diameter is $\frac{2}{5}Mr^2$.)

Letter to the Editor

Dear Editor,

The following problem was set in Volume 3, No. 2:

Let $z_1, z_2, \ldots, z_n$ be any complex numbers. Show that there exists a complex number z which minimizes the sum

$$\sum_{k=1}^{n} |z - z_k|^2.$$

This may be considered as a problem concerning moments of inertia in applied mathematics, as follows.

In an Argand diagram, suppose that Z_k is the point associated with z_k, Z the point associated with z, and let G be the centroid of unit masses at each point Z_k. The problem is to minimize

$$\sum_{k=1}^{n} (ZZ_k)^2.$$

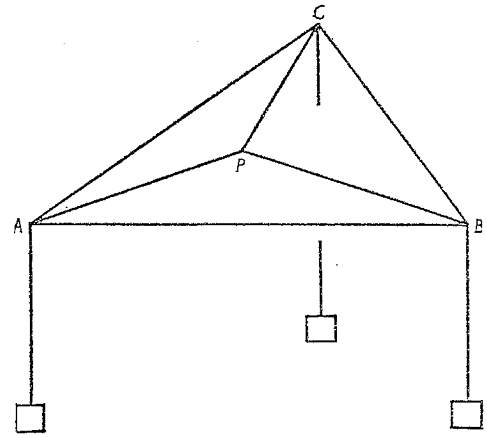
This expression is I_Z, the moment of inertia about Z of the n masses, and by the theorem of parallel axes

$$I_Z = I_G + n(ZG)^2$$

and its minimum value is seen to occur when Z is at G, so that the minimum value of Z is given by

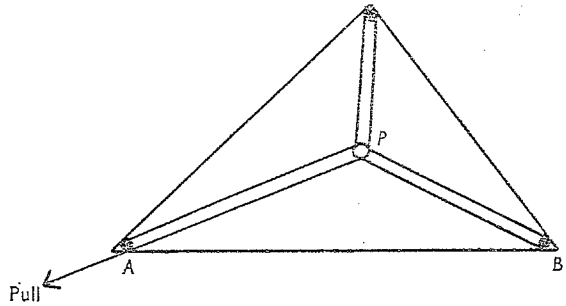
$$z = \frac{1}{n}\sum_{k=1}^{n} z_k.$$

Readers may be interested in a similar problem:
ABC is a triangle in which no angle exceeds 120°. *Find the position of the point P which is such that* $PA+PB+PC$ *is a minimum.*



With the plane of the triangle horizontal, fix smooth pulleys at A, B and C. Pass lengths of string over the pulleys, the strings carrying equal masses at one end and being tied together at the other. The equilibrium position is that for which the potential energy of the system is a minimum so that a maximum length of string is vertical and $PA+PB+PC$ is a minimum. Since the three tensions have equal magnitude and produce equilibrium the strings must make angles of 120° with each other and P is the point at which the sides of the triangle each subtend an angle of 120°. (If the triangle has an angle in excess of 120° no such point exists and the minimum value of the sum occurs when P is at the vertex where the greatest angle lies.)

After I had shown this method to a class, a pupil, C. W. Vout, suggested a pleasant alternative which, unlike the foregoing, I think to be original.



Locate small smooth pegs at A, B and C. Tie one end of a length of inextensible string to A, passing the string through a smooth ring and round the other pegs as shown in the diagram. Pull on the free end of the string, keeping the string in contact with the peg at A. As the string is pulled, $PA+PB+PC$ decreases until a position is reached in which further pulling increases the tension in the string without altering the position of the ring. $PA+PB+PC$ has then attained its minimum value, the ring is in equilibrium, and each of the sides of the triangle subtends an angle of 120° at P.

Yours sincerely,

R. W. Payne
(Dulwich College)

Problems and Solutions

Readers who have not yet reached the age of 20 on 1 October 1972 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

5.1. (i) Show that the situation consisting of five nodes each joined to the other four is not a 'planger'. (ii) Show that if a planger consisting of m nodes and a total of n joins is constructed with a pen-knife instead of a pencil, then the piece of paper will fall into $n-m+2$ pieces. (See the article in this issue by V. W. Bryant. In Figure 2 there are 5 nodes, 9 joins and 6 pieces.)

5.2. (Submitted by A. J. Douglas, University of Sheffield.) In negotiating a sale, the seller S first suggests a price s which the prospective buyer B counters with an offer b ($< s$). S then suggests $\frac{1}{2}(s+b)$, and so they proceed, each offer being the average of the two previous. Find the nth bids made by both S and B, and show that these tend to the same limit (the selling price) as $n \to \infty$. Does it make any difference to the selling price if B takes the initiative by making the first offer, and if so what is the difference?

If S begins by suggesting a price s and B cannot afford to pay more than p, what is the highest offer with which B can counter?

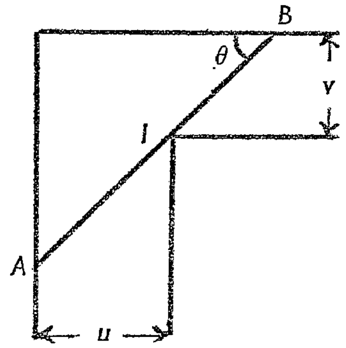
5.3. (Submitted by A. K. Austin, University of Sheffield.) A debating society decides to choose the best speech delivered in a given year, so a sequence of scores is allocated to each speech. The first score of a speech is the number of other speeches in the year which have referred back to that speech. The second score of each speech is the sum of the first scores of the speeches which have referred back to that speech. The third score of a speech is the sum of the second scores of the speeches which have referred back to it. And so on. What eventually happens to the scores?

5.4. There are n identical uniform planks each of length l stacked on top of one another. What is the maximum possible horizontal displacement of the top plank from the bottom one?

Solutions to Problems in Volume 4, Number 2

4.5. A thin rigid rod is carried along a straight corridor of width u. At the end of the corridor is one at right angles to it of width v, and it is required to carry the rod round the corner. If the rod must always be held in a horizontal position, what is the maximum possible length of the rod?

Solution by J. A. Blake (*Rugby School*)



The maximum length of the rod will be the minimum value of the length $AB = l$ as the angle θ varies between $\pi/2$ and 0. Now

$$l = u \sec\theta + v \operatorname{cosec}\theta$$

and

$$\frac{dl}{d\theta} = u \sec\theta \tan\theta - v \operatorname{cosec}\theta \cot\theta.$$

This is zero when $\tan\theta = (v/u)^{1/3}$, and this gives a minimum for l. Thus the maximum possible length of the rod is

$$(u^{2/3}+v^{2/3})^{3/2}.$$

Also solved by Philip Taylor, Ken Parell, Colin Gillespie (Sale Boys Grammar School), S. D. Barber (Dulwich College), Angus Rodgers (Gonville and Caius College, Cambridge), David Mell (King George V School, Southport), Susan Street (Barton Peveril Grammar School, Eastleigh).

4.6. If x, y, z denote the lengths of the sides of a triangle, show that $3(yz+zx+xy) \leqslant (x+y+z)^2 < 4(yz+zx+xy)$.

Solution by N. S. Penny, L. Miles, R. J. Ball (*The Grammar School, Ebbw Vale*)
Since $(x-y)^2 \geqslant 0$, we have

$$x^2+y^2 \geqslant 2xy.$$

Similarly, $y^2+z^2 \geqslant 2yz$, $z^2+x^2 \geqslant 2zx$. If we add these inequalities, we obtain

$$2(x^2+y^2+z^2) \geqslant 2(yz+zx+xy)$$

which gives that

$$(x+y+z)^2 \geqslant 3(yz+zx+xy).$$

Since x, y, z denote the lengths of the sides of a triangle, we have

$$x+y > z, \quad y+z > x, \quad z+x > y.$$

These give

$$xz+yz > z^2, \quad xy+xz > x^2, \quad yz+yx > y^2.$$

If we add these inequalities, we obtain

$$2(yz+zx+xy) > x^2+y^2+z^2,$$

which gives that

$$4(yz+zx+xy) > (x+y+z)^2.$$

Also solved by Angus Rodgers (Gonville and Caius College, Cambridge), David Mell (King George V School, Southport), H. Rabinowitz (Leeds Grammar School).

4.7. Let A be a square matrix with real elements. Let s be a real number greater than or equal to all the row and column sums. Show that each element of the matrix can be replaced by one at least as large so that all the row and column sums are s.

(*Note*: The 2×2 matrix $\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$ has row sums 3, 7 and column sums 4, 6.)

Solution by Angus Rodgers (Gonville and Caius College, Cambridge)

Suppose that A has n rows and columns, denote by r_i (respectively c_j) the sum of the elements in its ith row (respectively jth column), and let t be the sum of all its elements. Then $s-r_i \geqslant 0$, $s-c_j \geqslant 0$ for every i, j, and

$$\sum_{i=1}^{n}(s-r_i) = \sum_{j=1}^{n}(s-c_j) = ns-t.$$

Hence $ns-t \geqslant 0$, and if $ns-t = 0$ then all the row and column sums are already s. If $ns-t > 0$, we can make each row and solumn sum s if, for, $1 \leqslant i, j \leqslant n$, we add

$$\frac{(s-r_i)(s-c_j)}{ns-t}$$

to the element in the ith row and jth column.

4.8. Let n be a positive integer and let $a_1, a_2, \ldots, a_n$ be any real numbers greater than or equal to 1. Show that

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

Solution by Angus Rodgers (Gonville and Caius College, Cambridge)

$$\begin{aligned}(1+a_1)(1+a_2)\ldots(1+a_n) &= 2^n\left(1+\frac{a_1-1}{2}\right)\left(1+\frac{a_2-1}{2}\right)\ldots\left(1+\frac{a_n-1}{2}\right)\\ &\geqslant 2^n\left(1+\frac{a_1-1}{2}+\frac{a_2-1}{2}+\ldots+\frac{a_n-1}{2}\right)\\ &\geqslant 2^n\left(1+\frac{a_1-1}{n+1}+\frac{a_2-1}{n+1}+\ldots+\frac{a_n-1}{n+1}\right)\\ &= \frac{2^n}{n+1}(1+a_1+a_2+\ldots+a_n).\end{aligned}$$

Also solved by J. A. Blake (Rugby School), David Mell (King George V School, Southport), H. Rabinowitz (Leeds Grammar School).

Problem Solvers Series. Edited by L. MARDER. George Allen & Unwin Ltd, London. £1·75, hardback; £0·80, paperback.

1. **Ordinary Differential Equations.** By J. HEADING. 1971. Pp. 92.
2. **Calculus of Several Variables.** By L. MARDER. 1971. Pp. 84.
3. **Vector Algebra.** By L. MARDER. 1971. Pp. 88.
5. **Calculus of One Variable.** By K. E. HIRST. 1972. Pp. 85.

This new series of mathematics books is based on the idea that students need to see problems solved in order to understand and appreciate the concepts, theorems and techniques presented to them in lectures. The series is aimed at first year undergraduates in the first instance (although more advanced texts are promised in later volumes), but I suspect that students of the physical sciences and engineering will find these books more useful than specialist mathematicians.

It is the policy of the series that the greater part of each book should be taken up with solved problems; as a consequence, the theory is only sketched and proofs omitted. It must therefore be stressed that these books should be used in conjunction with a lecture course or with a standard text, and not as a substitute for these.

A brief summary of the contents is given below:

1. **Ordinary Differential Equations**
   (a) First order equations: separable, exact, Clairaut and Riccati equations.
   (b) Higher order equations: equations with constant coefficients, particular integrals. Homogeneous equations. Techniques involving change of variables.
   (c) Simultaneous equations: matrix methods, phase-plane analysis, normal modes.
   (d) Solutions in series, and Bessel's equations.
   (e) Solutions by Laplace transforms.

2. **Calculus of Several Variables**
   (a) Partial differentiation: chain rule, Laplace's equation in plane polar coordinates, wave equation, Euler's theorem for homogeneous functions.
   (b) Jacobians and transformations: implicit functions, functional dependence, transformations.
   (c) Taylor's theorem: maxima and minima, Lagrange multipliers, envelopes.
   (d) Multiple integrals: double and triple integrals, change of variables.
   (e) Line and surface integrals: Green's theorem.

3. **Vector Algebra**
   This volume covers the manipulation of vectors and their applications in mechanics and geometry. It provides a rather elementary treatment of an elementary topic.
   (a) Scalar, vector and triple products: geometrical interpretation, vector equations.
   (b) Systems of forces: equivalent systems, couples.
   (c) Differentiation of vectors: curvature and torsion of space curves, Serret–Frenet, motion of particles, rotating coordinate systems.

5. **Calculus of One Variable**
   (a) Limits, differentiation, chain rule, Taylor's theorem, extreme values, estimation of errors of approximation.
   (b) Integration: techniques, geometric and physical applications.

On the whole, the worked problems are well chosen. They range from the elementary to the more demanding, and examples which indicate the geometric and physical applications are included. The paperback versions are cheap enough to encourage students to acquire at least some of these books.

King's College, London A. S.-T. LUE

Introductory Real Analysis. By A. N. KOLMOGOROV and S. V. FOMIN. (Translated from the Russian by RICHARD A. SILVERMAN.) Prentice-Hall, Inc, Englewood Cliffs, N.J., 1970. Pp. xii+403. £7·00.

In *Mathematical Spectrum*, Volume 3, pages 17–21, N. Youd (who was at that time a pupil at St Paul's School, London) showed how to solve a Cambridge Scholarship calculus problem by a mixture of geometrical intuition and linear algebra, hopefully extended to infinite dimensional space. That, roughly speaking, is what this book is about, except, of course, that the intuitive ideas are made rigorous and the hopes are fully realized.

In studying differential or integral equations, one often uses transformations of the kind $\psi(x) = (A\phi)(x)$ in which a function ϕ is replaced by ψ. If the functions $\phi(x)$, $\psi(x)$ are thought of as 'vectors' (check that functions behave like vectors, with real numbers as scalars), then the transformation A is like an ordinary linear transformation, or linear operator, on the vectors. For example, if a function $K(x, y)$ of two variables is defined in such a way that $K(x, y) = 0$ for $y > x$ and is continuous for $y < x$, then the transformation A given by

$$\psi(x) = (A\phi)(x) = \int_a^b K(x, y)\, \phi(y)\, dy$$

leads to the notion of a Volterra operator. One can ask whether or not there is a vector (a function) f for which $Af = \lambda f$ where λ is a real number. By analogy with ordinary vectors such a function is called an eigen-vector and the numbers λ and their associated transformations give rise to what is called the spectrum of the operator. For certain spaces, called Hilbert spaces, which arise in quantum theory, that spectrum is related to the colour spectrum of light; this provides a beautiful connection between mathematics and physics, all the more astounding since it was not explicit at first.

This book contains all the pure mathematical background necessary to start on the adventure of 'doing calculus by geometry and linear algebra', which is properly called functional analysis. It begins with the necessary algebraic and geometrical background (very nicely done in Chapters 1 to 4) and the study of linear transformations of the kind alluded to above (Youd's ideas, essentially Bessel's inequality, are to be found in their proper setting on page 150). The book concludes with chapters on integration (beautifully done) and differentiation, in which these basic notions are thoroughly discussed, in some respects from the point of view already established. The whole book is a pleasure to read, though the experienced mathematician will note some infelicities. These do not detract from the value of the book as a whole: the distinguished Russian authors have been well served by their translator.

The book covers the whole of the analysis course taken by a university undergraduate specializing in analysis; so clearly it goes outside the school syllabus. On the other hand, many topics begun in school come to fruition in the sort of material presented here; school teachers, and occasionally their pupils, ought to be aware of the fact and look at a book of this kind from time to time.

A well-stocked school library could usefully spend some of its allocation on a copy, though for many it will probably represent a luxury and a pleasure to be deferred.

University of Nottingham J. V. ARMITAGE

Mathematics Made Difficult. By CARL E. LINDERHOLM. Wolfe Publishing Co, London, 1971. Pp. 207. £2·75.

The dedication of this book ends 'with apologies to St. Thomas Aquinos' (sic). It sums up the book rather well if I say I am not quite sure whether the spelling is deliberate. I am even less sure who the book is aimed at. The text falls into two parts, the funny bits and the mathematical bits, and I found the effort of switching channels rather too much. Consequently I laughed at little of the humour and learned little from the mathematics.

The humour I found rather juvenile, and often too obvious. It is hardly adult to point out that the G. of G. Peano stands for Giuseppe and not Grand; and I found the long sections on M. Boulangiaire the baker (who was obviously educated solely by means of Dienes' multibase blocks) tedious in the extreme. My approach to the irrationals does not benefit from this type of motivation.

The mathematics often seemed pointless. No space is wasted on definitions, either of symbols or of terms. Very early in the book the following appears:

> If $\{x\} \hookrightarrow X \xrightarrow{s} X$ (abbreviated X) and $\{y\} \hookrightarrow Y \xrightarrow{t} Y$ are pointed functions a morphism $X \xrightarrow{f} Y$ is a function $X \to Y$ such that the squares of
>
> $$\begin{tikzcd} \{x\} \arrow[r, hook] \arrow[d] & X \arrow[r, "s"] \arrow[d, "f"] & X \\ \{y\} \arrow[r, hook] & Y \arrow[r, "t"] & Y \end{tikzcd}$$
>
> commute (which means that $x \mapsto y$ and $fs = tf$).

This is not just Mathematics Made Difficult, it is Mathematics Made Ridiculous.

There are some nice touches, however, such as the section where the numbers are written:

1, 2, 3, 4, 5, 1ᔭ, 1Ɛ, 1ᘔ, 1Ɩ, 10, 11... .

Unfortunately there are errors in the examples where some digits are printed the wrong way up. Another interesting section is the one entitled 'Guess the next number'. But even so, I could find better things to spend my library allowance on, let alone my own money.

University of Nottingham — K. E. SELKIRK

Further Mathematical Diversions. By MARTIN GARDNER. George Allen & Unwin Ltd, London, 1971. Pp. 255. £1·75.

There is a remote possibility that there are a few readers of this magazine who have not yet discovered the works of Martin Gardner. For these, if they indeed exist, it is necessary to say that he writes a regular 'Mathematical Games' article for the monthly journal *Scientific American*. This book is the fifth in a series containing items collected from *Scientific American* columns. Each of the articles has been expanded from its original form mainly by the inclusion of material provided by readers.

It is difficult to exaggerate the contribution that Martin Gardner has made to recreational mathematics in its widest sense, and this book is a valuable addition to a distinguished line. The mathematical sophisticate might argue that there is little in it that is truly original, and in a sense this is true. There are, however, many delightful 'twists' of familiar situations and the reader is encouraged to consider a wide range of problems, paradoxes, catches, games, puzzles, etc.

In recommending this book to all who enjoy mathematics and especially to those teachers who are looking for ideas to enliven mathematics in the classroom, perhaps the reviewer might be allowed to show some personal preference by offering the following small selection of items which appeal to him personally.

The Paradox of the Unexpected Hanging. This is a more gruesome version of the examination paradox which arises from the following statement made by a teacher to her class: 'I am going to give you a test on *one* certain day next week but you will not know which day it is until I tell you on the morning of the day of the test'. The interpretation of this apparently harmless remark is shown to be the subject of violent controversy amongst expert logicians.

Tests of Divisibility. In addition to more familiar material, this section contains some little known tests for divisibility by 7.

Rep-Tiles. Four trapezoidal shapes can be fitted together to form a similar trapezium or 'replica' (try it for yourself). This section deals with shapes which 'replicate' in this manner including some which cannot be constructed in a finite number of steps.

There are, in addition, references to chess, solitaire, hexapawn, the fourth dimension and flatland, *e*, knots, cat's cradles and many other topics. The book is well produced with a pleasing type face and clear diagrams. It has a useful bibliography giving substantial further reading references for each chapter. If you have not done so already, go and buy it for yourself, your school/college library, or your favourite girl/boy friend who hates mathematics.

University of Nottingham D. S. HALE

New Mathematical Diversions. By MARTIN GARDNER. George Allen & Unwin Ltd, London, 1970. Pp. 253. £2·40.

This book by Martin Gardner is written and compiled very much in the style of his previous ones. It begins by introducing the reader to the concept of different counting systems, with special reference to the binary, and then continues with punch cards, groups and braids. There is a chapter on Lewis Carroll and some of the puzzles he made up, including his game of doublets and a very interesting maze. Then follows a chapter on paper cutting, introducing many more of the fascinations of plane geometry. There are three chapters spaced through the book which contain some good problems. A history of board games and the mathematics involved is mentioned, including an introduction to the long forgotten game of 'Reversi'. There is also a chapter on packing spheres into a fixed volume, and finding the densest packing arrangement.

Topics of particular interest to me were the chapters on Victor Eigen, the ellipse and some calculus.

I liked this book very much and enjoyed reading it. The topics are clearly presented with many diagrams, and make good reading for the younger (pre-O level) mathematical enthusiast. It is a must for a school library, but until it comes out in paperback, one penny per page is probably too much for an individual to pay.

Try problem number 4, page 221.

5th Form, St. Paul's School, London S.W. 13 D. L. HUNT

Mathematical Formulae for Engineering and Science Students. By S. BARNETT and T. M. CRONIN. Bradford University Press (Crosby Lockwood & Sons Ltd, London), 1971. Pp. 50. £0·75, offset litho.

Designed for 'open book' examinations, this is a concise list, not a textbook, of formulae, definitions and theorems classified under ten headings: Series; Transcendentals; Integrals–Differentials; Integration (reduction) Formulae including Gamma and Beta functions, the usual Ordinary Differential Equations up to Linear with real constant (or reducible) coefficients; Laplace Transform; Vectors up to div, curl and Green's Theorem; Matrices up to inverse and rank; Numerical Methods; Statistics.

The authors believe this to be the only book of its kind, restricted to what is sufficient for student needs. It is not so overwhelming in its content as to alarm a sixth former aiming at an Engineering or Science Diploma or Degree; such a student will already know some part of it, and should therefore be in a position to estimate the depth of background work which intelligent use of its formulae demands.

I would place it on the shelves of a sixth form science library as part of a first reply to the question 'what kind of maths shall I be doing if . . . . ?' A private owner might delight in ticking off what he understands or can prove.

University of Nottingham R. L. LINDSAY

Notes on Contributors

Paul R. Jackson attended King Edward VI Camp Hill School for Boys, Birmingham. He has studied statistics for about four years, and was a second-year student at Sheffield University, reading pure mathematics and statistics, when he wrote the article in this issue. He has been interested in football for a number of years, and is a keen supporter of Aston Villa.

V. W. Bryant, a graduate of Manchester and Sheffield, is a Lecturer in Pure Mathematics in the University of Sheffield. His interests lie primarily in geometry and, more particularly, in the study of convexity. Indeed, convexity is the theme of a forthcoming book by Bryant and R. J. Webster, who is known to the readers of *Mathematical Spectrum* for his contributions to previous issues.

Catherine Smallwood graduated from the University of Sheffield and for the past two years has been a Lecturer in Mathematics at the Polytechnic of North London. Her mathematical interests lie in the domain of classical analysis and, more particularly, in the theories of measure and integration.

D. J. Roaf is a Lecturer in Theoretical Physics at the University of Oxford, and a Fellow in Mathematics at Exeter College, Oxford. He has also taught at Monash University, Melbourne, Australia. His article developed from some lectures given to schoolteachers attending a course in Oxford. His hobbies are change ringing and local politics.

Contents

PRICES (*postage included*)

	Individual annual subscription (1 volume of 2 issues)	Bulk annual subscriptions (5 or more copies of 1 volume of 2 issues)	Single issues
U.K. only	£0·40	£0·35	£0·23
Australia	$A1·45	$A1·25	$A0·80
U.S.A.	$1·60	$1·40	$0·90

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 John Wright & Sons Ltd, Bristol