MATHEMATICAL SPECTRUM

A MAGAZINE FOR STUDENTS AT SCHOOLS COLLEGES AND UNIVERSITIES

Volume 14 1981/82 Number 1

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 14 of *Mathematical Spectrum* will consist of three issues, of which this is the first. The second will be published in January 1982 and the third in May 1982.

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 New Cover for *Mathematical Spectrum*

In Volume 12 Number 3 we invited our readers to suggest a possible new cover design for *Mathematical Spectrum*. A number of readers responded; the new cover, which retains a version of the spectrum symbol, is based on a design submitted by Peter Cheek of the Mathematics Department, Brighton Polytechnic, who has been awarded the £15 prize.

Bonds of Friendship

VICTOR BRYANT, *University of Sheffield*

Victor Bryant is a Lecturer in Pure Mathematics at the University of Sheffield. He is co-editor of two books of *Sunday Times* Brain-teasers (published by Unwin Paperbacks) and is himself a regular setter of Brain-teasers. He is also editor of the *Mathematical Gazette*.

In this article we shall look at ways in which modern geometry can be used as a basis for popular puzzles. The three examples cited are based upon *Sunday Times* Brain-teasers, and we thank the editor of that newspaper for his kind permission to reproduce them here. Consider first the following puzzle, of the handshaking variety.

First puzzle: Handshakes

My wife and I attended a dinner party at which there were eight other people, namely the four couples Mr and Mrs A, Mr and Mrs B, Mr and Mrs C, and Mr and Mrs D. Introductions were made and a certain number of handshakes took place. (But, of course, no one shook hands with him/herself nor with his/her spouse, nor more than once with the same person.) At the end of the evening I asked each other person how many hands he/she had shaken, and I was surprised to find that the answers given were all different. Also, the total of the number of handshakes made by the other four men was the same as the total of handshakes made by all five women. The only woman with whom I shook hands was Mrs A. Whose hands did my wife shake, and how many hands did Mrs A shake?

Of course, this puzzle can be solved by various methods, some much longer than others. Let us represent the people at the party by 0, 1, 2, 3, 4, 5, 6, 7, 8 (their number being equal to the number of hands they shook) and me. Mark these ten as points around a circle, and join two of them by a straight line if they shook hands. Now 0 shook hands with no one, so 8 shook hands with everyone except 0. So 1 shook 8's hand and no one else's which means that 7 shook all the hands except 0's and 1's. Continuing in this way gives the situation illustrated in Figure 1. Husbands do

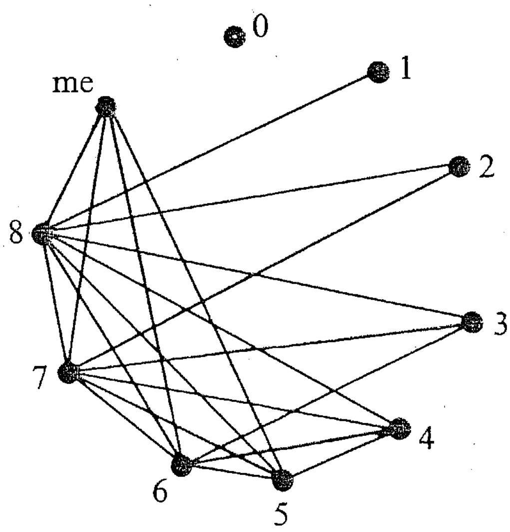


Figure 1

not shake hands with wives, and so 8 and 0 are married to each other. But then 7 must be married to 1, 6 to 2, 5 to 3, and 4 to me. The remainder of the solution is not going to concern us much; we leave the reader to conclude that three of 5, 6, 7 and 8 are men and that their sum is 18 (being half of $0 + 1 + \cdots + 8$). Hence Mrs A is 8 and my wife shakes hands with Mr B, Mr C, Mr D and Mrs A.

What does concern us here about that solution is the use of points to represent people and straight lines joining pairs of points to represent certain bonds between them. Let us now pause to state a couple of results about such illustrations, which we shall then be able to use later.

Mark any number of points around a circle and join some pairs of them by straight lines. The two results which will be relevant to us are:

(i) If you start with six or more points then either three of them will be joined to each other or three of them will be such that none is joined to the other. For example, in Figure 2 both situations actually occur: 136 is a triangle and 125 is an 'invisible' triangle.

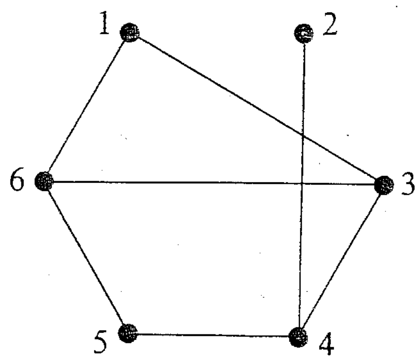


Figure 2

In general, think of the joined pairs as being joined by black lines and the unjoined pairs as being joined by white lines. Then (i) says that there is a triangle of one colour. Its proof is then surprisingly straightforward. Consider any point; it is joined to at least five others and the joins are white or black, so at least three of these joins must be of one colour. So we may assume that the point *a* is joined in black to each of *b*, *c* and *d*. If *b* is joined to *c* in black, then *abc* is a black triangle; similarly, if *b* is joined to *d* in black, or if *c* is joined to *d* in black, then we get a black triangle. On the other hand, if none of *bc*, *bd* and *cd* is black then, of course, *bcd* is a white triangle, and (i) follows.

(ii) If each of the points is at the end of exactly two of the lines, then the points can be re-positioned so that the figure becomes a collection of disjoint polygons. (See Figure 3.)

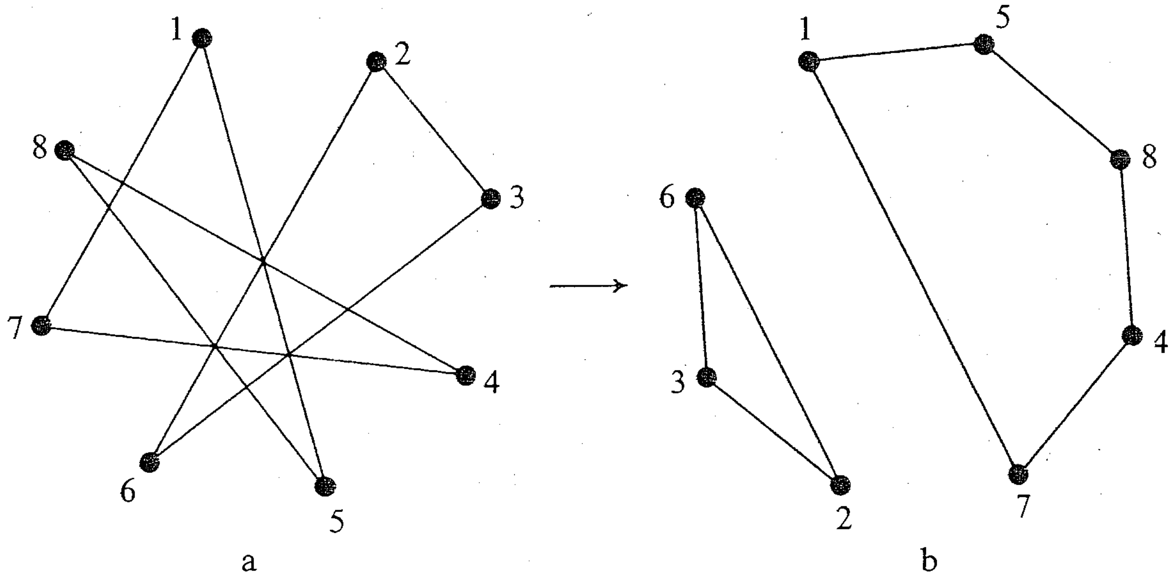


Figure 3

We leave the proof of this one to the interested reader. Basically, the fact that each point is joined to two others means that, if you choose a point and start to 'walk' along one of the lines and never turn back, then your route is determined for you, and you will eventually return to the point at which you started. That tour then forms the first of your polygons, and you repeat the process at some other point.

These two pleasant results from 'graph theory', a branch of modern geometry, will now be relevant to our next two puzzles.

Second puzzle: Silver wedding

. . . After the toasts, a small group of people, consisting of the couple themselves [Ann and John], Ann's mother Freda, my wife Chris, and at least one of Ann and John's three sons, were standing together. In congratulatory mood, each person in the group either exchanged kisses with or shook hands with each other member of the group. For example, Ann's son Stuart kissed someone who kissed someone else who kissed Freda, who shook hands with someone who shook hands with someone else who shook hands with Chris, who shook hands with someone who kissed one of Ann's sons. (And no two men kissed each other.) I noticed that no three of the people in the group all kissed each other, and no three all shook hands. How many of Ann and John's sons were in the group, and whom did my wife kiss?

Again imagine the people in the group as points around a circle, and join two by a straight line if they shook hands. Then the absence of a handshaking trio means that the figure has no triangle, and the absence of a kissing trio means that the figure has no 'invisible' (or 'white') triangle. By (i), therefore, there are fewer than six points. So the five named people form the entire group (and only one son is in the group). Furthermore, our proof of (i) shows that no point can be joined in black to three or more others, and no point can be 'joined in white' (or unjoined) to three or more others. So each of the five points is joined to precisely two others. Hence (ii) tells us that our lines form a collection of polygons, but with only five points there can only be a single polygon, as shown. Since S(tuart) and J(ohn) shook hands, we can place them in Figure 4 as shown. The remaining information (that some x kisses S and F, etc.) means that the above pentagon is S, J, C, A, F, and so C kisses S and F.

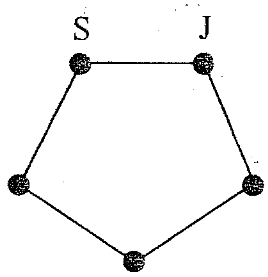


Figure 4

Third puzzle: Baby-sitting club

My wife and six of her married friends form a baby-sitting circle, but all is not as innocent as it seems. Each of the seven women is having an affair with a husband of one of the other members of the circle. Each of the husbands is having an affair with one of the women.

Eileen's husband's lover is not Eileen's lover's wife. Beryl's husband is not Dorothy's lover. The two men in the life of Fay's lover's wife are Christine's husband and Christine's lover's wife's lover.

Recently Anne and Gwen each divorced her own husband and married the husband of the other. No other relationship changed. Surprisingly, everything said in the previous paragraphs was true before and after this rearrangement took place.

Tonight Gwen's husband's lover is going to the theatre with her own husband, and Eileen's husband's lover is going to baby-sit for them. Anne's husband's lover is also going to the show with her own husband, and my wife is going to baby-sit for them. So, to keep me company, my lover is coming round to my place for the evening. Who is my wife and who is my lover?

The fourteen people involved in this highly immoral example can again be represented by points around a circle, with two being joined if they are married *or* if they are lovers. So each is joined to two others and hence, by (ii), the lines form a collection of disjoint polygons. Also, since each join is a man to a woman, each of the polygons has an even number of edges (i.e. quadrilateral, hexagon, octagon, etc.). The information about Fay and Christine tells us that they are part of the same hexagon. So the polygons are either a hexagon and two quadrilaterals or a hexagon and an octagon. In fact, because the information applies before and after the rearrangement, the reader will discover, after some effort, that one of these situations must represent the relationships before the remarriages, and the other situations must represent the relationships afterwards (see Figure 5). The first

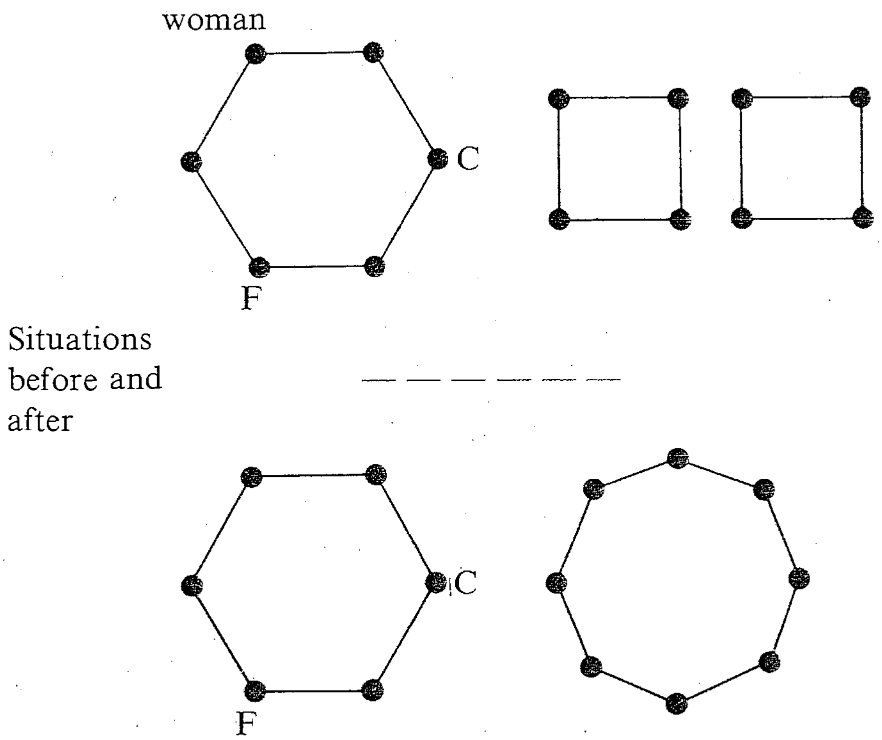


Figure 5

information about Eileen tells us that she is not part of a quadrilateral, so we can soon deduce that the hexagon includes F, C and E, and that A and G are in different quadrilaterals in the upper part of Figure 5. We leave the reader to piece together the remaining information and to work out which of the male points is 'me'.

I hope that you have enjoyed this glimpse of Brain-teasers from a geometer's point of view. There is a footnote to the last puzzle. In case any reader might be offended by its immorality, the puzzle was actually redrafted in the *Sunday Times* to concern wives playing chess with their husbands and also with a regular 'second partner'. So the puzzle concluded with the fact that my wife would be out tonight and my second partner would be coming round for a game of chess. However, the *Sunday Times* readers were not so naive; amongst the post-cards sent in with the answers on them were footnotes such as '*chess*? who are they kidding?!'

Mathematics in India in the Middle Ages

CHANDRA B. SHARMA, *University of Manchester Institute of Science and Technology*

Dr Sharma comes from a small village, Galuapur, in the Kanpur district of India. After gaining a Ph.D. from IIT Kanpur he joined the Transport Technology Department of Loughborough University of Technology where he obtained another Ph.D. He is presently a lecturer in the Applied Mechanics Group of the Mathematics Department in the University of Manchester Institute of Science and Technology (UMIST), where his current research interests are in the field of elastic stability. This article is reprinted, with the editor's permission, from *The Mancunian Indian*, the journal of the Indian Association, Manchester.

In the years from 400 AD to 1200 AD there had been mathematical activities of immense importance in India. Several scholars of mathematics of international repute, like Aryabhatt, Varahmihir, Brahmagupta, Mahaviracharya, Shridharacharya and Bhaskaracharya etc., displayed brilliant mathematical powers and made significant contributions which will remain in use for ever. These glories were achieved as a spin-off to their studies in astronomy, and mathematics to them was merely a handmaiden to astronomy. Some of the basic discoveries are the ones we use everyday, e.g. the decimal system, invention of zero and negative numbers. Their outstanding contribution, however, was to the study of indeterminate equations, and here they went far beyond their times and anticipated some of the discoveries of modern algebra. In the following, a brief glimpse of the works of three of them is given.

Aryabhatt† was born in 476 AD at Pataliputra (Patna, the present-day capital of Bihar Pradesh). In his celebrated work, *Aryabhattiyam*, the third chapter is devoted to mathematics. He gave rules for square roots, summing an arithmetic progression, solution of simple and quadratic equations and simple algebraic identities. The most important feature, however, is his treatment of indeterminate equations by the application of continued fractions which substantially is the same as the method used today. One typical example of his work is given in the rule for finding the value of π 'Add four to one hundred, multiply by eight, and add again sixty-two thousand; the result is the approximate value of the circumference of a circle whose diameter is twenty thousand.' This leads to the value of π as $3{\cdot}1416$, a remarkably correct value to four decimal places.

Brahmagupta was the most prominent of the Hindu mathematicians of the seventh century. He lived and worked in the great astronomical centre of Hindu science, Ujjayini (modern-day Ujjain in Madhya Pradesh). When he was only thirty years old, Brahmagupta wrote an astronomical work in twenty-one chapters, entitled *Brahma-Sputa-Siddhanta*, which includes special chapters on arithmetic and algebra. His work is based largely upon his predecessors, notably Aryabhatt; nevertheless, it contains much original research. Arithmetical operations, including

† The first Indian satellite was aptly named 'ARYABHAT' in his honour.

extraction of roots, problems on interest, arithmetic series and sum of the squares of natural numbers, etc., occur frequently in its pages. He knew the correct formula for the area of a triangle in terms of the sides, and gave a similar one for cyclic quadrilaterals. His treatment of equations was prefaced by the rules for dealing with negative quantities and zero—viz. negative multiplied or divided by negative gives positive, negative subtracted from zero is positive, etc. The nature and extent of his work will be best understood by a couple of examples he poses and solves. In every case a rule appropriate to the problem precedes the solution:

'Five hundred *drammas* were loaned at a rate of interest not known. The interest of that money for four months was lent to another person at the same rate and it accumulated in ten months to 78. Tell me the rate of interest.'

'In a certain lake, swarming with geese and cranes, the tip of a bud of lotus was seen half a cubit above the surface of the water. Forced by the wind it gradually advanced and was submerged at a distance of two cubits. Compute, quickly, mathematician, the depth of water.'

Another outstanding scholar among many Hindu mathematicians was **Bhaskaracharya** who flourished during the twelfth century. He was the most powerful and original thinker of his days and his two most celebrated works, *Lilavati* and *Vij-Ganita* (algebra), form the most complete exposition of Hindu mathematics until modern times. *Lilavati* was translated into Persian by Faizi in 1587 by direction of the emperor Akbar. Faizi states that Lilavati was the name of Bhaskara's daughter and that the astrologers predicted that she should never wed. Bhaskara, however, divined an auspicious moment for her marriage and left an hour cup floating in a vessel of water. This cup had a small hole in the bottom and was so arranged that the water would trickle in and sink it at the end of the hour. Lilavati, however, with a natural curiosity, looked to see the water rising in the cup, when a pearl dropping from her garment chanced to stop the influx. So the hour passed without the sinking of the cup and Lilavati was thus fated never to marry. To console her, Bhaskara wrote a book in her honour, saying: 'I will write a book of your name which shall remain for ever, for a good name is a second life and the groundwork of eternal existence.'

Among the many rules which he dealt with exhaustively he made one of the very important observations; 'the fraction, whose denominator is zero, is termed an infinite quantity' and gave a very beautiful conception of infinity, saying that 'there is no alteration in this quantity, though it may be inserted or extracted, as no change takes place in the infinite and immutable God at the time of destruction or creation of worlds, although numerous orders of beings are absorbed or put forth.'

The *Vij-Ganita* is a work on algebra. Quadratic equations are solved by completing the square and it is given that the square root of a positive quantity is two-fold, positive and negative, which incidentally in Sanskrit (and Hindi) is designated as 'debt' or 'loss'. The imaginary roots were dismissed.

Observe the ingenuity, elegance and the appeal of the following couple of problems (which, incidentally, can readily be solved by the solution of quadratic equations):

'The son of Pritha (Arjuna), exasperated in combat, shot a quiver of arrows to slay Karna. With half his arrows he parried those of his antagonist; with four times the square root of the quiverful he killed his horses; with six arrows he slew Shalya (the charioteer); with three he demolished the umbrella, standard and bow; and with one he cut off the head of the foe. How many were the arrows which Arjuna let fly?'

'One pair out of a flock of geese remained sporting on water, and saw seven times half the square root of the flock proceeding to the shore, tired of the diversion. Tell me, dear girl, the number of the flock.'

Much more can be said about these past prodigies, but I would like to end this piece by quoting the phrase of eulogy of Hermann Hankel, one of the greatest historians of mathematics (who not only knew mathematics well, being a professor of the subject (1867) at the University of Leipzig, but was well versed in various oriental languages):

'If one understands by algebra the application of arithmetical operations to complex magnitudes of all sorts, whether rational or irrational numbers or space-magnitudes, then the learned scholars of Hindustan are the real inventors of algebra.'

A Paradox

The following paradox appeared in the journal *Pi Mu Epsilon* Volume 6, Number 4 (1976), and is reproduced here with permission:

$$i(\sqrt{i}+\sqrt{(-i)}) = i\sqrt{i}+i\sqrt{(-i)} = \sqrt{(-i)}+\sqrt{i} = \sqrt{i}+\sqrt{(-i)}.$$

How Infectious Diseases Spread in Households

NIELS BECKER, *La Trobe University*

The author is a Senior Lecturer in the Department of Mathematical Statistics at La Trobe University in Victoria, Australia. His main research interests lie in the statistical analysis of epidemic data and survival data; he is also known to enjoy a good argument about the foundations of statistical inference. For relaxation he jogs, plays golf and makes strange noises on a guitar.

1. Epidemic chains

When studying the spread of infectious diseases it is not usually possible to conduct experiments, and so, we must rely instead on observations made during the course of naturally occurring epidemics. Natural epidemics consist of chains of infections, which arise, at least partially, by chance. Epidemics are thus best described by probability models and analysed by a corresponding statistical analysis. We illustrate here how such a statistical analysis can help to provide answers to important epidemiological questions.

For simplicity we consider only outbreaks of the disease in households of size three. By a household of size three we mean one which contains only three individuals who are susceptible to the disease just prior to the outbreak. Thus a household might actually consist of a family of two parents and three children, where (as in the case of measles) the parents may have had the disease some time ago and thereby acquired lasting immunity. Individuals infected by contacts from outside their own household are called introductory cases. As we are interested in studying the spread of the disease within the household, affected households with three introductory cases give us no useful information. With two introductory cases there are two possible epidemic chains which we denote by 2–0 and 2–1. In a 2–0 chain the remaining susceptible escapes infection and so there is no secondary case, whereas the 2–1 chain indicates that there is a secondary case and so all susceptibles in the household become infected.

With one introductory case there are four possible epidemic chains which are denoted by 1–0, 1–1–0, 1–1–1 and 1–2. The chain 1–0 indicates that the introductory case did not infect anyone else in the household, while the chain 1–2 indicates that the introductory case is directly responsible for infecting each of the other two susceptibles. The chains 1–1–0 and 1–1–1 indicate that the introductory case is directly responsible for infecting one of the other two susceptibles, so that there is one secondary case. In the chain 1–1–0 the final susceptible escapes infection so that there is no tertiary case, while in the chain 1–1–1 the final susceptible is infected by the secondary case and so there is one tertiary case. Under suitable circumstances, it is possible to classify outbreaks in households of size three into the above epidemic chains and we assume this to be the case in practice, without offering more detailed explanations.

2. Some probability models

Let us now derive expressions for the probabilities of the various chains. In households with one introductory case two susceptibles are exposed to this case. Assume that each *independently* escapes infection with probability Q. Then, given Q, the probability distribution of the number of secondary cases is given by

number of secondary cases	0	1	2
probability, given Q	Q^2	$2Q(1-Q)$	$(1-Q)^2$

in accordance with the binomial probability distribution. Readers will hardly need to be reminded that the mean and variance of this distribution are $2(1-Q)$ and $2Q(1-Q)$ respectively. The value of Q depends on how long the introductory case remains infectious and on the intensity of the illness. These properties vary among individuals, so that it is appropriate to consider Q as a random variable. To find the unconditional probability distribution of the number of secondary cases we need to take the expectation, or mean value, of each of the above conditional probabilities. This gives

number of secondary cases	0	1	2
probability	ϕ	$2(\theta-\phi)$	$1-2\theta+\phi$

where $\theta = EQ$, $\phi = EQ^2$ with E denoting the expectation. When the number of secondary cases is 0 or 2 it is not possible to have any tertiary cases. When there is one secondary case there will be no tertiary cases with probability θ and one tertiary case with probability $1-\theta$. This leads to the following chain probabilities for households of size three with one introductory case:

chain	1–0	1–1–0	1–1–1	1–2
probability	ϕ	$2\theta(\theta-\phi)$	$2(1-\theta)(\theta-\phi)$	$1-2\theta+\phi$

The chain probabilities for households of size three with two introductory cases are

chain	2–0	2–1
probability	λ	$1-\lambda$

where λ denotes the probability of the remaining susceptible escaping infection when exposed simultaneously to two cases.

3. Estimating the model parameters

We do not know the specific probabilities, but have expressions for them in terms of the parameters, namely the unknown constants θ, ϕ and λ. These parameters may take different values for different diseases and different communities, but their values must satisfy $0 \leq \phi \leq \theta \leq 1$, $\phi \geq 2\theta - 1$ and $0 \leq \lambda \leq 1$, so that the above chain probabilities all lie between 0 and 1. In any particular application we must estimate the values of θ, ϕ and λ, and statistical methods are used for the efficient estimation and the testing of hypotheses about these parameter values.

Let us suppose, then, that we intend to observe a total of m households with two introductory cases and imagine that in M_0 of these the remaining susceptible escapes infection. That is to say, the chain 2–0 is observed M_0 times. Then M_0 has a binomial distribution given by

$$P(M_0 = x) = \binom{m}{x} \lambda^x (a - \lambda)^{m-x}, \qquad x = 0, 1, \ldots, m. \tag{1}$$

The estimator M_0/m is said to be *unbiased* for λ, because its expectation is $EM_0/m = \lambda$. The standard deviation of M_0/m, namely $[\lambda(1 - \lambda)/m]^{1/2}$, can be used to indicate the accuracy of M_0/m as an estimator for λ. These results follow directly from the binomial distribution (1) for M_0, whose mean and variance are respectively $m\lambda$ and $m\lambda(1 - \lambda)$.

Let us also suppose that we intend to observe a total of n households with one introductory case and imagine that these are classified into the various possible chains to give

chain	1–0	1–1–0	1–1–1	1–2
number of times observed	N_0	N_{10}	N_{11}	N_2

so that $n = N_0 + N_{10} + N_{11} + N_2$. We see that each of N_0, N_{10}, N_{11} and N_2 taken separately has a binomial distribution, while together they have a multinominal distribution. In particular, note that

$$P(N_0 = x) = \binom{n}{x} \phi^x (1 - \phi)^{n-x}, \qquad x = 0, 1, \ldots, n, \tag{2}$$

so that the mean and variance of N_0 are respectively $n\phi$ and $n\phi(1 - \phi)$. Therefore, N_0/n is an unbiased estimator for ϕ with standard deviation $[\phi(1 - \phi)/n]^{1/2}$. To arrive at an estimator for θ note that $N_1 = N_{10} + N_{11}$ has the binomial distribution given by

$$P(N_1 = x) = \binom{n}{x} (2\theta - 2\phi)^x (1 - 2\theta + 2\phi)^{n-x}, \qquad x = 0, 1, \ldots, n. \tag{3}$$

Therefore, from the properties of the binomial distribution (3), $EN_1 = 2n(\theta - \phi)$, and so $N_1/2n + N_0/n$ is an unbiased estimator for θ. The standard deviation of this estimator is seen to be $[(\theta + \phi)/2n - \theta^2/n]^{1/2}$.

Other estimators can be used. Consider, for example, only the households with one secondary case. Each of these leads to the chain 1–1–0 with probability θ, and so

$$P(N_{10} = x; \text{given } N_1) = \binom{N_1}{x} \theta^x (1 - \theta)^{N_1 - x}, \qquad x = 0, 1, \ldots, N_1. \tag{4}$$

It follows that N_{10}/N_1 is an unbiased estimator of θ, but it can be shown that its standard deviation is greater than the standard deviation of the earlier estimator $N_1/2n + N_0/n$. On this basis N_{01}/N_1 is judged to be inferior to $N_1/2n + N_0/n$.

Other estimators can be constructed, and indeed the possibility of combining the estimators N_{01}/N_1 and $N_1/2n + N_0/n$ to produce a better estimator for θ suggests itself. We shall not go into that possibility here, nor shall we explore other, more efficient methods of estimation.

Some actual data will help us to illustrate how the model and the estimators can be applied to particular diseases. In Table 15.8 of reference 1, some data are quoted for the incidence of measles in households of size three. The observed frequencies for 60 households with one introductory case and 11 households with two introductory cases are summarised by

chain	1–0	1–1–0	1–1–1	1–2	2–0	2–1
number of times observed	6	11	6	37	4	7

Hence $n = 60$ and the observations on N_0, N_{10}, N_{11} and N_2 are 6, 11, 6 and 37, respectively. Also $m = 11$ and the observation on M_0 is 4. This leads us to the estimate $4/11 \simeq 0{\cdot}364$ for λ. The accuracy of this estimate is indicated by the standard deviation $[\lambda(1-\lambda)/11]^{1/2}$, which we estimate to be 0·145 by substituting our estimate for λ in this formula. We say that the estimate 0·364 of λ has standard error 0·145. Similarly, the estimate of ϕ is $6/60 \simeq 0{\cdot}100$ and has standard error 0·039, while the estimate of θ is $17/120 + 6/60 \simeq 0{\cdot}242$ and has standard error 0·043. By substituting the values $\lambda = 0{\cdot}364$, $\phi = 0{\cdot}100$ and $\theta = 0{\cdot}242$ into the given expressions for the chain probabilities, the above epidemic model for households of size three becomes completely specified.

So far we have looked only at the problem of parameter estimation. We turn now to the other important application of statistics, namely hypothesis testing.

4. Testing a hypothesis about risks of infection

Let us try to formulate an epidemiologically meaningful hypothesis. Around 1928, the American epidemiologists Reed and Frost formulated similar epidemic chain models. One difference in our model is that we have taken Q to be a random variable and so allowed for variations in the infectiousness of cases. However Reed and Frost made an additional assumption, namely that when a susceptible is exposed to more than one infectious individual, then each of these poses a separate and independent risk of infection to the susceptible. In terms of the above probability model this hypothesis may be stated as $\lambda = \theta^2$. We now suggest one way of testing the equivalent hypothesis $H\colon \lambda - \theta^2 = 0$ by making use of the estimators for λ and θ suggested above.

We base our test on the estimator $M_0/m - (N_1/2n + N_0/n)^2$ for the parameter $\lambda - \theta^2$. This is not an unbiased estimator: while we know that $N_1/2n + N_0/n$ is unbiased for θ, it does not follow that $(N_1/2n + N_0/n)^2$ is unbiased for θ^2. In fact

$$\begin{aligned} E(N_1/2n + N_0/n)^2 &= [E(N_1/2n + N_0/n)]^2 + \operatorname{Var}(N_1/2n + N_0/n) \\ &= \theta^2 + (\theta + \phi)/2n - \theta^2/n \end{aligned} \tag{5}$$

but the bias will be small if n, the number of observed households with one introductory case, is large. If, in an application, the fixed numerical values for m and n, and the observations on M_0, N_0 and N_1 are substituted in our estimator for $\lambda - \theta^2$ and we obtain a negligible value, then we would be content to accept the hypothesis

$H: \lambda - \theta^2 = 0$. If on the other hand the value is not negligible, then we would doubt and possibly reject the hypothesis H. Just how far removed from 0 the observed value of the estimator needs to be before we reject H depends on the alternative hypothesis we have in mind.

A meaningful alternative hypothesis is given by the following argument. If the presence of one infectious individual causes the household environment to become infected to the same level as the presence of two infectious individuals, and susceptibles become infected by contact with the environment, then $\lambda = \theta$. The true situation is possibly not as extreme as this. The chance of escaping infection by two introductory cases is possibly less than the chance of escaping infection by one introductory case, although it is perhaps not as small as the chance of escaping infection by separate and independent exposure to two infectious individuals. In short we expect the value of λ to lie somewhere between θ^2 and θ. The hypothesis H specifies the extreme $\lambda = \theta^2$ and a meaningful alternative is the hypothesis $K: \lambda - \theta^2 > 0$. This makes our test what statisticians refer to as 'one-sided' and we should consider rejecting H only if the observed value of $M_0/m - (N/2n + N_0/n)^2$ is much larger than 0. The critical value beyond which we should reject H is computed from the probability distribution of our estimator.

The exact distribution of the estimator under the hypothesis H is quite complicated, but fortunately it can be approximated by a normal distribution, when m and n are large. Hence, provided we have enough data, we need to derive only the mean and the standard deviation of the estimator. We have already established that, for large n, the mean of $M_0/m - (N_1/2n + N_0/n)^2$ is approximately $\lambda - \theta^2$. The standard deviation is the positive square root of

$$\begin{aligned} \operatorname{Var}[M_0/m - (N_1/2n + N_0/n)^2] &= \operatorname{Var}(M_0/m) + \operatorname{Var}[(N_1/2n + N_0/n)^2] \\ &\simeq \lambda(1-\lambda)/m + 2\theta^2(\theta + \phi - 2\theta^2)/n, \end{aligned} \tag{6}$$

where, in the last line, we have already deleted terms which are negligible when n is large. Therefore, using the normal distribution as an approximation and the 5% level of significance, we reject the Reed–Frost hypothesis, $H: \lambda = \theta^2$, if $M_0/m - (N_1/2n + N_0/n)^2$ exceeds the critical value

$$1{\cdot}64[\lambda(1-\lambda)/m + 2\theta^2(\theta + \phi - 2\theta^2)/n]^{1/2}. \tag{7}$$

Of course we do not know the values of θ, ϕ and λ, but may substitute our estimates of these into (7) for the critical value.

We now illustrate this test of the hypothesis $H: \lambda = \theta^2$ with reference to the measles data given above. For these data our estimate of $\lambda - \theta^2$ is $0{\cdot}364 - (0{\cdot}242)^2 \simeq 0{\cdot}305$, and substituting our estimates for λ, θ and ϕ into (7) gives $0{\cdot}240$. Therefore, since $0{\cdot}305$ exceeds $0{\cdot}240$, we reject the hypothesis $H: \lambda = \theta^2$ in favour of the hypothesis $K: \lambda > \theta^2$. The epidemiological interpretation is that there is significant evidence to indicate that the probability of escaping measles infection when exposed to two infectious individuals of the same generation is greater than the probability of escaping infection by separate and independent exposure to each of two infectious individuals.

We have given a simple illustration of how statistical analyses of infectious disease data based on probability models can help the study of epidemics. Great progress has been made in the control of infectious diseases, but even today over 1000 million people are affected by infectious diseases such as malaria, schistosomiasis, filariasis and trachoma. Mathematics and statistics have an important role to play when studying these diseases; by helping to measure the infectiousness, determining the fraction of the community that must be vaccinated in order to prevent major epidemics, and answering similar important questions.

Reference

1. N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases* (Griffin, London, 1975).

Blind Student Tests New Computer Equipment

A third-year B.Sc. Mathematics for Business student at Middlesex Polytechnic is the first user of a new computer terminal designed for the blind. Twenty-two-year-old Paul Holliman, who has been blind from birth, is working on the new equipment while on the industrial placement year of his course at the National Physical Laboratory (NPL), at Teddington, Middlesex.

Paul uses two pieces of equipment which enable him to make a full contribution to the work of the linear algebra section of NPL's Division of Numerical Analysis and Computer Science. One is a computer terminal called 'Brailink', designed to Middlesex Polytechnic's order by Clarke and Smith of Wallington, the company which developed the talking book machine. 'Brailink' is portable and can be plugged into any computer network; after Paul's year at NPL, it will be available for use at Middlesex Polytechnic by any blind student. The other is a Sagem hard-copy Braille character printer which is loaned to NPL by the Manpower Services Commission.

Paul's main job is to test computer programs which have been designed to solve industry's mathematical problems. Programs are entered on the Brailink via the keyboard as on a conventional terminal, but instead of a visual display unit there is a perforated metal strip through which pins spelling out Braille code can be felt. The terminal also has its own microprocessor so that programs can be prepared off-line on cassette without having to wait for computer time to be available.

Of course, Paul has to work harder than other students to overcome his handicap. Graphs, for instance, are a problem. Members of staff have to produce special large-sized graphs with raised lines for him. He has a pocket calculator which has been specially adapted for his needs by an electronics expert. He also has an IBM golf-ball typewriter adapted for blind users, with one golf-ball devoted entirely to mathematical symbols. Paul has gained confidence at work, thanks to the loan of the equipment and the helpfulness of his colleagues, and he enjoys an active social life among the young people of Teddington.

As for the future, Paul knows that potential employers will be worried about getting equipment for him but on the whole he is hopeful. 'What's encouraging,' he says, 'is that in spite of the recession we don't know of any Maths for Business graduate on the dole.'

Continuous Transformation Groups

W. D. HALFORD, *Massey University*

Dean Halford is a Senior Lecturer in Mathematics in the Department of Mathematics and Statistics at Massey University, Palmerston North, New Zealand. His main research interests are in general relativity and cosmology, but he holds a special interest in mathematical modelling in general. One of his particular concerns is teaching mathematics at the secondary school–university interface. He is on the Council of the New Zealand Mathematical Society and is Editor of the Society's Newsletter.

1. Introduction

This article is the first of two which together provide a brief introduction to Lie groups; the second will appear in the next issue of *Mathematical Spectrum*. One may approach this area of mathematics in several ways, but the way which is chosen here is based on the concrete example of a transformation group, an object with which you may already have had some contact.

The study of groups is now a part of many school mathematics curricula. One reason for this is that a group is an important and relatively simple algebraic structure which focuses attention on the common underlying parts of seemingly disparate sections of pure mathematics—as a 'unifier' it is worthy of (and has been given) much study. Another reason is that group theory is an important tool in the hands of applied mathematicians, physicists, chemists, biologists and social scientists.

We may broadly classify groups as either discrete or continuous. A *discrete group* is a set G of elements for which a composition rule (or 'multiplication') $\circ$ is given and which obeys the group axioms, namely:

(G1) *Closure.* The product (or resultant) $a \circ b$ of any two elements a, b of the set G is a unique element which also belongs to G.

(G2) *Associative.* When three or more elements are multiplied the order of the compositions is immaterial, i.e. if a, b, c are elements of G, then

$$a \circ (b \circ c) = (a \circ b) \circ c.$$

(G3) *Identity.* The set G contains an element e which is such that the composition of each element a of G with e leaves a unchanged:

$$a \circ e = e \circ a = a.$$

We call e the *identity element*; it is easy to see that there is only one such element in G.

(G4) *Inverse.* For each element a of G there is an element b of G such that

$$a \circ b = b \circ a = e.$$

We call b the *inverse* of a and usually write $b = a^{-1}$; again, it is easy to see that an element has only one inverse.

As an example of a discrete group we mention the large class of *symmetry groups* which are so important in physics and chemistry, for example in crystallography. The symmetry of a body or configuration is described by giving the set of all transformations which preserve the distances between all pairs of points of the body and bring the body into coincidence with itself. This set of transformations forms a group called a symmetry group. Again, the *symmetric group* (of all permutations on n symbols) denoted by S_n is a group which plays a large role in physics.

A *continuous group* is one for which the elements not only satisfy the group axioms (G1)–(G4) above, but also satisfy a continuity requirement. The concept of continuity utilizes the idea of 'nearness', in the sense that corresponding to a small change in one of the group elements involved in a product there is a consequential small change in the product itself. (The formal definition of continuity of a function is met in the calculus.) One may say, in short, that the elements of a continuous group form a continuum, as will be explained later.

An important type of continuous group is a *Lie group* which arises, for example, in the study of differential equations and the special functions of mathematical physics. Lie groups often enter physics through the presence of symmetries in a kinematical or dynamical sense e.g. in relativity theory or quantum field theory.

In the second article we shall sketch a short biography of the creator of Lie groups, the Norwegian mathematician Sophus Lie (1842–1899), and go on to describe how these groups arise in the study of differential equations. The rest of the present article is devoted to transformation groups, a topic which you will have met if you have studied transformation geometry. Because of their importance as continuous groups we give an account detailed enough to lead on to the study of Lie groups.

2. Some transformation groups

2.1. *Group of translations.* In coordinate geometry we may interpret the pair of equations

$$x_1 = x + b_1, \qquad y_1 = y,$$

where b_1 is a real constant, as a *translation* through a distance b_1 parallel to the X-axis from the point P with coordinates (x, y) to the point P_1 with coordinates (x_1, y_1). (See Figure 1.) Let us denote this operation by T_{b_1} and write

$$T_{b_1}: \qquad x_1 = x + b_1, \qquad y_1 = y. \tag{1}$$

Consider a second translation

$$T_{b_2}: \qquad x_2 = x_1 + b_2, \qquad y_2 = y_1, \tag{2}$$

which takes the point P_1 with coordinates (x_1, y_1) to the point P_2 with coordinates

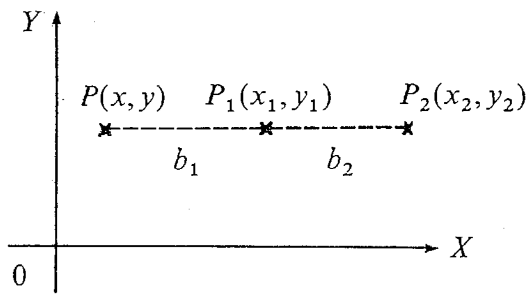


Figure 1

(x_2, y_2) through a distance b_2 parallel to the X-axis. The combined effect (*product*) of these two translations may be worked out analytically:

$$T_{b_2}T_{b_1}: \qquad x_2 = x_1 + b_2 = x + b_1 + b_2, \qquad y_2 = y_1 = y,$$

which is obviously of the form

$$T_b: \qquad x_2 = x + b, \qquad y_2 = y, \tag{3}$$

where $b = b_1 + b_2$. Geometrically, the point $P(x, y)$ is translated through a distance $b_1 + b_2$ parallel to the X-axis by the single translation T_b, which we may write as $T_{b_1+b_2}$. Thus we have established that the product of two translations is a translation, so the group axiom (G1) is satisfied, the elements of G being translations in this case.

It is not difficult to prove that (G2) also holds. For we merely observe that $(b_1 + b_2) + b_3 = b_1 + (b_2 + b_3)$ by the associative law of the algebra of real numbers, so that translating from P to P_2 to P_3 is the same as translating from P to P_1 to P_3. (P_3 has coordinates (x_3, y_3) where $x_3 = x_2 + b_2$ and $y_3 = y_2$.) We write this as

$$T_{b_3}(T_{b_2}T_{b_1}) = (T_{b_3}T_{b_2})T_{b_1}$$

where T_{b_3}: $x_3 = x_2 + b_3$, $y_3 = y_2$. Thus the associative law holds for translations.

Let us put $b_1 = 0$ in equations (1) and denote this transformation by

$$I: \qquad x_1 = x, \qquad y_1 = y.$$

I is called the *identity transformation*. Geometrically, we start at the point P and end at P. If we perform the translation (1) followed by the identity transformation I, we have

$$IT_{b_1}: \qquad x_1 = x + b_1 + 0 = x + b_1, \qquad y_1 = y,$$

which is just T_{b_1}. This is also what results if we perform I then T_{b_1}. Hence

$$IT_{b_1} = T_{b_1}I = T_{b_1}.$$

This is (G3).

Now suppose that in equations (2) we give b_2 the value $-b_1$. Then performing the translations (1) and (2) in that order, we have

$$T_{-b_1}T_{b_1}: \qquad x_2 = x_1 - b_1 = x + b_1 - b_1 = x, \qquad y_2 = y_1 = y.$$

The product (resultant) of these translations is therefore the identity transformation. Geometrically, P_2 is now coincident with P; we have gone from P to P_1 and then back to P. We call T_{-b_1} the *inverse* of T_{b_1} and usually write T_{-b_1} as $T_{b_1}^{-1}$. Thus

$$T_{b_1}^{-1}T_{b_1} = I.$$

One can also show that $T_{b_1}T_{b_1}^{-1} = I$. Hence (G4) is satisfied.

Thus all the group axioms are satisfied and we have a *transformation group*. The real numbers b_1, b_2, b_3 appearing above are called *parameters*. In fact, only one of these occurs in any translation so there is only one *effective* or *essential* group parameter. Hence the group of translations is a 1-*parameter group*. If we restrict the values of the parameter to the integers, say, the group is said to be discrete. On the other hand, if the parameter is allowed to range over the entire set of real numbers, or some interval of real numbers, then the group is said to be continuous since the real numbers form a continuum; in this case the operation T_{b_1} can be converted to the operation T_{b_2} by continuous variation of the parameter from the value b_1 to the value b_2. We normally think of the translation group as a continuous group. If each of the parameters of a continuous group ranges over a bounded domain, say the closed interval of real numbers $[m, n]$, the group is sometimes said to be *closed* and the set of elements of the group, considered as a set of 'points' (see Section 3), is said to be *compact*. If each of the parameters ranges over an unbounded domain, the set of group elements is said to be *non-compact*. (The idea of compactness is a topological one; here we are using the 'closed' and 'bounded' properties of each of the group parameters in the topological sense. Indeed, the transformation groups we are considering are *Lie groups*, as we shall see in the next article, and a Lie group is necessarily a *topological group*.) For the translation group the single parameter ranges over an unbounded domain (all the real numbers) so this group is non-compact.

2.2. *Scale transformations* (*dilatations*). The transformation given by

$$T_{a_1}: \qquad x_1 = a_1 x, \qquad y_1 = a_1 y, \qquad (a_1 \neq 0) \tag{4}$$

is called a *scale transformation* or *dilatation* since it scales up (or down) distances in the plane. We may interpret this as a motion from the point $P(x, y)$ to the point $P_1(x_1 y_1)$ as indicated in Figure 2. Because the triangles OAP, OA_1P_1 are similar,

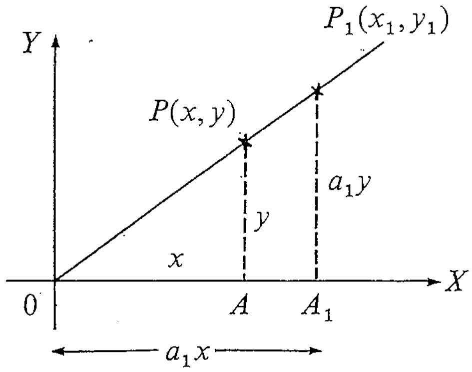


Figure 2

this is often called a *similarity transformation*. The real number a_1 in equations (4) is the one effective parameter of the transformation.

Take a second scale transformation

$$T_{a_2}: \qquad x_2 = a_2 x_1, \qquad y_2 = a_2 y_1, \tag{5}$$

with parameter value a_2. The product (resultant) of the two transformations (4) and (5) is

$$T_{a_2}T_{a_1}: \qquad x_2 = a_2 x_1 = a_2 a_1 x, \qquad y_2 = a_2 y_1 = a_2 a_1 y.$$

This is of the form

$$T_a: \qquad x_2 = ax, \qquad y_2 = ay, \tag{6}$$

where $a = a_2 a_1$, so is itself a scale transformation, and group axiom (G1) is satisfied.

You may like to show that the other group axioms (G2), (G3) and (G4) are satisfied, proving that the scale transformations form a 1-parameter continuous group; it is non-compact, since the essential parameter a takes all positive real values. The parameter value giving the identity transformation is 1, and that giving the inverse transformation is $1/a$.

2.3. *Rotation group in two dimensions*. From coordinate geometry the equations

$$\begin{aligned} x_1 &= x\cos t - y\sin t, \\ y_1 &= x\sin t + y\cos t \end{aligned} \tag{7}$$

may be interpreted as defining a rotation of the point $P(x, y)$ about the origin through the angle t into the position $P_1(x_1, y_1)$. (See Figure 3.) Using the same technique as in Examples 2.1 and 2.2, it can be shown that the set of rotations (7) with real-valued parameter t forms a 1-parameter continuous group called the *rotation group in two dimensions* and is denoted by the symbol $O(2, \mathbb{R})$. (O stands for 'orthogonal', a term which will become clear later; 2 is the number of dimensions; $\mathbb{R}$ indicates that the parameter set is a set of real numbers.)

If t_1, t_2 are the parameter values of two successive rotations about 0, then the product (resultant) transformation has parameter value $t_1 + t_2$, corresponding to a single rotation through the angle $t_1 + t_2$. The inverse transformation is given by the parameter value $-t$ in equations (7), while the identity is given by $t = 0$. Note that, if

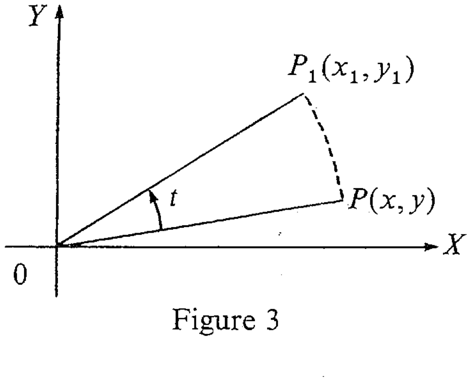


Figure 3

we let t vary continuously from 0 to 2π, then P rotates through a complete circle. Hence the rotation group in two dimensions is compact.

2.4. *Linear group in two dimensions.* Consider the equations

$$x_1 = ax + by, \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0, \tag{8}$$
$$y_1 = cx + dy,$$

where a, b, c, d are real numbers and $|\cdot|$ denotes the determinant. These equations generate a 4-parameter continuous group of transformations called the *real general linear group in two dimensions*, denoted by $\mathrm{GL}(2, \mathbb{R})$. The parameters can vary over the unbounded domain of real numbers, so the group is non-compact. This is the first example of a group with more than one continuously varying essential parameter, so we shall spend a little time on it.

Let the transformation from point $P(x, y)$ to point $P_1(x_1, y_1)$ be given by (8) with parameter values a_1, b_1, c_1, d_1 and let a successive transformation from point P_1 to point $P_2(x_2, y_2)$ be given by (8) with parameter values a_2, b_2, c_2, d_2. Then

$$\begin{aligned} x_2 &= a_2x_1 + b_2y_1 \\ &= a_2(a_1x + b_1y) + b_2(c_1x + d_1y) \\ &= (a_1a_2 + b_2c_1)x + (a_2b_1 + b_2d_1)y \\ &= a'x + b'y, \end{aligned} \qquad \begin{aligned} y_2 &= c_2x_1 + d_2y_1 \\ &= c_2(a_1x + b_1y) + d_2(c_1x + d_1y) \\ &= (a_1c_2 + c_1d_2)x + (b_1c_2 + d_1d_2)y \\ &= c'x + d'y, \end{aligned}$$

where

$$\begin{aligned} a' &= a_1a_2 + b_2c_1, \qquad & b' &= a_2b_1 + b_2d_1, \\ c' &= a_1c_2 + c_1d_2, \qquad & d' &= b_1c_2 + d_1d_2, \end{aligned} \tag{9}$$

showing that the product is a transformation of the same type, and the group axiom (G1) is satisfied. (It should be noted that $\begin{vmatrix} a' & b' \\ c' & d' \end{vmatrix} \neq 0$.)

We can proceed in similar fashion to verify that the axioms (G2), (G3) and (G4) hold. This procedure is cumbersome, particularly as we must check that the determinant of the parameter is non-zero each time. It is rather more elegant and also more enlightening to write the transformation equations (8) in matrix form:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0, \tag{10}$$

or

$$\boldsymbol{x}_1 = A\boldsymbol{x}, \qquad \det A \neq 0, \tag{10a}$$

where the column vectors $\boldsymbol{x}_1, \boldsymbol{x}$ and the non-singular matrix A are given by

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \qquad \boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{and} \qquad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The transformations $P \to P_1$ and $P_1 \to P_2$ are now given by

$$\boldsymbol{x}_1 = A_1\boldsymbol{x}, \qquad \boldsymbol{x}_2 = A_2\boldsymbol{x}_1,$$

where A_1 and A_2 are the matrices with entries a, b, c, d carrying subscripts 1 and 2 respectively. The product transformation $P \to P_1 \to P_2$ is then given by

$$\mathbf{x}_2 = A_2\mathbf{x}_1 = A_2A_1\mathbf{x} = A'\mathbf{x}, \tag{11}$$

where $A' = A_2A_1$ is the ordinary matrix product, A' having entries a', b', c', d'. If you check this matrix multiplication out you will find that you obtain precisely equations (9), with the condition

$$\det A' = \det(A_2A_1) = (\det A_2)(\det A_1) \neq 0.$$

Let

$$A = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

then equation (10a) gives $\mathbf{x}_1 = \mathbf{x}$. Therefore the unit matrix I represents the *identity transformation*. Since $\det A \neq 0$ the inverse matrix A^{-1} exists. Let us pre-multiply both sides of equation (10a) by A^{-1}:

$$A^{-1}\mathbf{x}_1 = A^{-1}A\mathbf{x} = I\mathbf{x} = \mathbf{x}.$$

Geometrically, having gone originally from $P(\mathbf{x})$ to $P_1(\mathbf{x}_1)$ (a transformation represented by the matrix A), we have then returned to P (a transformation represented by the matrix A^{-1}). That is, the *inverse transformation* matrix is A^{-1}.

We can summarize this information in the language of group theory as follows: The real general linear group in two dimensions GL(2, $\mathbb{R}$) is 'isomorphic' to the group of invertible 2×2 matrices, with matrix multiplication as the group composition rule. (It is well known and can be proved independently that the set of invertible $n \times n$ matrices, with matrix multiplication as the composition rule, forms a group.)

2.5. *Orthogonal group in two dimensions.* An important requirement in many physical situations is that the transformation equations (8) preserve (leave invariant) the sum $x^2 + y^2$; that is, we require

$$x_1^2 + y_1^2 = (ax + by)^2 + (cx + dy)^2 = x^2 + y^2.$$

The four parameters a, b, c, d are subjected to the following three relations:

$$a^2 + c^2 = 1, \qquad b^2 + d^2 = 1, \qquad ab + cd = 0. \tag{12}$$

In matrix terms this is equivalent to

$$A^{\mathrm{T}}A = I, \tag{12a}$$

where A^{T} is the transpose of A, as can be seen directly by performing the matrix multiplication. A matrix A which obeys the requirement (12a) is called an *orthogonal matrix*. The three relations (12) reduce the four parameters effectively to only one. The transformation (8) can now be written in terms of this single parameter in the form (7), as you can verify by putting $a = \cos t = d$, $b = -\sin t = -c$. Thus we recover the rotation group O(2, $\mathbb{R}$); it is a subgroup of the group GL(2, $\mathbb{R}$). We could have noticed earlier that preserving the sum $x^2 + y^2$ means preserving the square of the distance from the origin to the point $P(x, y)$; the transformation which does this is a rotation about the origin. The orthogonality of the transformation matrix A is recognized in the alternative name *orthogonal group.*

2.6. *Other groups*. There are several other continuous transformation groups of importance in mathematics and physics. We mention in passing the following:

(a) The *affine group* in two dimensions consisting of all transformations of the form

$$\boldsymbol{x}_1 = A\boldsymbol{x} + \boldsymbol{b},$$

where A is a non-singular 2×2 matrix (i.e. $\det A \neq 0$) and $\boldsymbol{x}_1, \boldsymbol{x}, \boldsymbol{b}$ are 2×1 column vectors. This group has 6 essential parameters. An important subgroup is the group of rigid motions (rotations plus translations) in the plane in which A is restricted to be an orthogonal matrix; this subgroup is called the *Poincaré group* in two dimensions. The affine group in three dimensions is an important extension of the two-dimensional affine group.

(b) The *real general linear group* in n dimensions, $\mathrm{GL}(n, \mathbb{R})$, which is represented by non-singular $n \times n$ matrices with real-number entries. There are n^2 essential parameters.

(c) The *real special linear group* in n dimensions, $\mathrm{SL}(n, \mathbb{R})$, is a subgroup of $\mathrm{GL}(n, \mathbb{R})$ in which the $n \times n$ matrices appearing in the transformation equations have determinant 1. The number of essential parameters is $n^2 - 1$.

(d) The *orthogonal group* in n dimensions, $\mathrm{O}(n, \mathbb{R})$. We restrict the transformation equations of $\mathrm{GL}(n, \mathbb{R})$ to those which are represented by a non-singular $n \times n$ matrix A with real entries which is orthogonal i.e. $A^{\mathrm{T}}A = I$. This imposes $n + \frac{1}{2}n(n-1)$ conditions on the n^2 parameters, leaving $\frac{1}{2}n(n-1)$ essential parameters. $\mathrm{O}(n, \mathbb{R})$ is a compact subgroup of $\mathrm{GL}(n, \mathbb{R})$.

(e) The *special orthogonal group* in n dimensions, $\mathrm{SO}(n, \mathbb{R})$, is a subgroup of $\mathrm{O}(n, \mathbb{R})$ in which the transformation matrix A has $\det A = 1$. The group $\mathrm{SO}(n, \mathbb{R})$ is a compact subgroup of $\mathrm{GL}(n, \mathbb{R})$.

(f) The *unitary group* in n dimensions, $U(n)$. This is the set of transformations in which the $n \times n$ transformation matrix A has complex number entries and also $A^+A = I$. Here A^+ is the Hermitian conjugate of A, i.e. A^+ is the transpose of the matrix whose entries are the complex conjugates of the corresponding entries of A. The number of essential parameters is n^2. This group is compact.

(g) The *special unitary group* in n dimensions, $\mathrm{SU}(n)$. This is the subgroup of $\mathrm{U}(n)$ in which the transformation matrix A is restricted by the condition $\det A = 1$. The group $\mathrm{SU}(n)$ is compact.

(h) The *projective* (or *fractional linear*) *group* in two dimensions. The transformation equations are

$$x_1 = \frac{ax + by + c}{px + qy + r}, \qquad y_1 = \frac{dx + ey + f}{px + qy + r},$$

where the coefficients $a, b, c, d, e, f, p, q, r$ are real numbers with the restriction

$$\begin{vmatrix} a & b & c \\ d & e & f \\ p & q & r \end{vmatrix} \neq 0.$$

There are 8 essential parameters since we can always make $a(\neq 0)$ take the value 1 by suitable division. Note that the above projective transformations reduce to the real linear transformations (8) upon putting $r = 1$, $c = f = p = q = 0$.

(i) The *Lorentz group* in four dimensions, O(3, 1). This is the underlying group of relativity theory. It is the group of Lorentz transformations represented by a 4×4 matrix A with real entries which satisfies

$$AGA^{\mathrm{T}} = G,$$

where G is the matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

There are 6 essential parameters. The group is non-compact. (If we add a constant term to each of the transformation equations there will be a further 4 essential parameters added, and we shall have the 10-parameter *Poincaré group* of rotations and translations in four dimensions.)

3. Groups continuously connected to the identity

The elements of a group are often called 'points'. Extending the use of geometrical language further (notice how useful geometrical ideas are!), this set of 'points' is said to constitute a *group space* or *group manifold.*

Now consider again the translation group G of Example 2.1. The transformation equations are

$$x_1 = x + a, \qquad y_1 = y, \tag{13}$$

with parameter a which ranges continuously over the set of real numbers. If instead of these transformation equations we had

$$x_1 = -x + a, \qquad y_1 = y, \tag{14}$$

then we have not a single continuous motion from $P(x, y)$ to $P_1(x_1, y_1)$ but in fact a transformation in *two parts* (see Figure 4):

(i) An *inversion* (reflection in the Y-axis) $P \to P'$ given by

$$x' = -x, \qquad y' = y;$$

(ii) A translation $P' \to P_1$ given by

$$x_1 = x' + a, \qquad y_1 = y'.$$

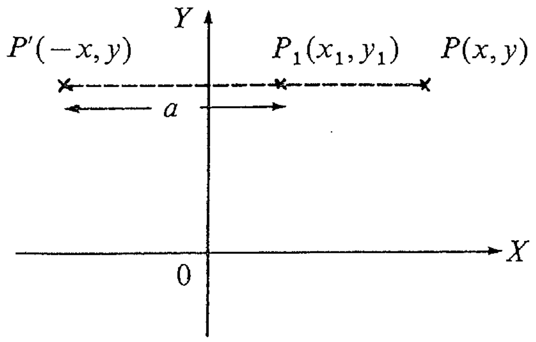


Figure 4

Now the set H of transformation equations

$$x_1 = \pm x + a, \qquad y_1 = y \tag{15}$$

form a group, as you can prove. The translation group G is a subgroup of H. Because the leap from P to P' by the inversion (i) cannot be made by continuous variation of the parameter a from its value 0 (the identity) at P, we say that *the group manifold consists of two disconnected pieces*. We say further that *the group G given by the equations* (13) *is continuously connected to the identity*; but this is not so for the set of transformations (14) which is *not* a group.

We may argue similarly for the orthogonal group. The group manifold consists of two disconnected pieces, namely the set of transformations with determinant equal to 1 (which we have considered in Examples 2.5 and 2.6(e)), and the set with determinant equal to -1 (which we did not consider). The first set is called *the group of proper rotations* and is given by continuous parameter variation from the identity value i.e. *is continuously connected to the identity*. The second set is not.

In the case of the Lorentz group in four dimensions the group manifold consists of four pieces, only one of which is continuously connected to the identity.

4. Summary

In addition to the properties possessed by all groups, the transformation groups which we have considered possess a topological character—they are continuous, have disconnected parts and may or may not be compact. These topological properties assume great importance in applications of the theory of Lie groups, of which the transformation groups are examples. Our next article, to be published in Volume 14 Number 2 of *Mathematical Spectrum*, will take up the discussion of Lie groups.

Letters to the Editor

Dear Editor,

Delivery costs minimised!

A. V. Boyd (*Mathematical Spectrum* Volume 13, Number 1, p. 29) posed the problem of finding the location of a warehouse on a circle to minimise the sum of distances along the circle to n shops, also located on the circle. He presented solutions only for $n = 3,4$.

The general solution is also quite simple. The warehouse should be located at a circular median which, by analogy with the median on a line, is defined as a point, the diameter through which divides the data into equal sized groups. Each such diameter has, of course, two endpoints. Usually, only one of these, that for which the sum, S, of distances to all data points is the less, is a circular median. Mardia (reference 1, p. 31) proves that a point with the property of minimising this sum is necessarily a circular median. However, this is not a sufficient condition.

Determining which circular medians minimise the sum is computationally very easy. A procedure for doing so is illustrated in the following example. In Figure 1, there are eight points on the circumference of a circle; two of these (3,7) are diametrically opposed. The interpoint distances, proceeding anticlockwise from point 1 are 1, 5, 2, 6, 2, 2, 3, 3; the circumference is 24 units.

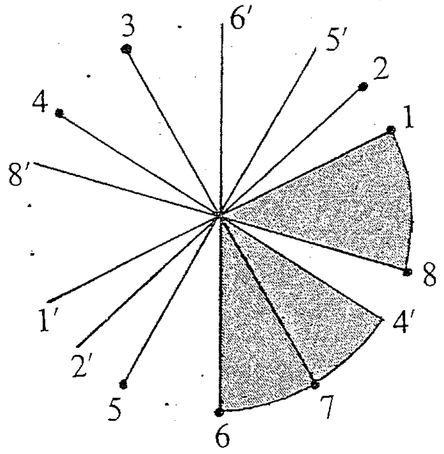


Figure 1

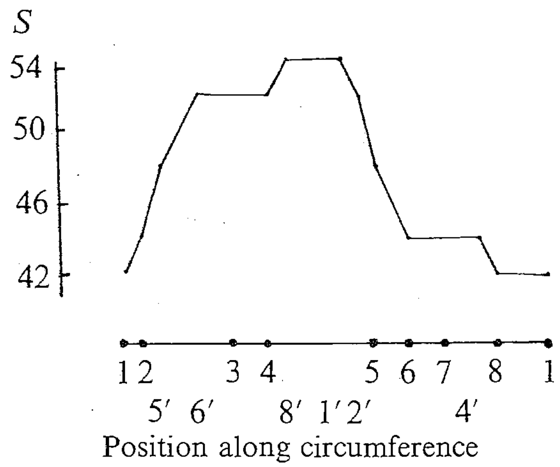


Figure 2

The sum of distances, S, to all 8 points, expressed as a function of position on the circumference, is graphed in Figure 2. As the reader may confirm for himself, this sum need only be calculated directly at one point. At remaining diameter endpoints, S may be calculated by difference or reflection. Linear interpolation determines S for all other points on the circle.

In this example, all points on the circle in the two shaded regions of Figure 1 are medians. However only those on the arc joining points 1 and 8 minimise S. Any point on this arc is then an optimum warehouse location.

Reference

1. Mardia, K. V. (1972) *Statistics of Directional Data*. (Academic Press, London, 1972).

Yours sincerely,

David Griffiths
(CSIRO Division of Mathematics and Statistics
P.O. Box 218, Lindfield, NSW 2070, Australia)

Dear Editor,

Fermat's principle and curvature

Fermat's principle asserts that the optical distance travelled by a disturbance is either a maximum or a minimum. In the case of reflection at a curved surface the choice depends upon the curvature of the surface at the point of incidence. This is best illustrated using two properties of the ellipse, namely, that the focal radii to a point on the ellipse make equal angles with the normal to the ellipse at that point and that the sum of their lengths is constant. With these it can be seen that the optical path $AP+PB$ in Figure 1 is a minimum because

$$AP+PB=AQ+QB<AR+RB. \tag{1}$$

Similarly it can be seen that the same optical path in Figure 2 is a maximum.

It is evident from these diagrams that there is a critical curvature, namely that of the ellipse with foci at A and B, at the point of incidence P, about which the optical path changes from a maximum to a minimum.

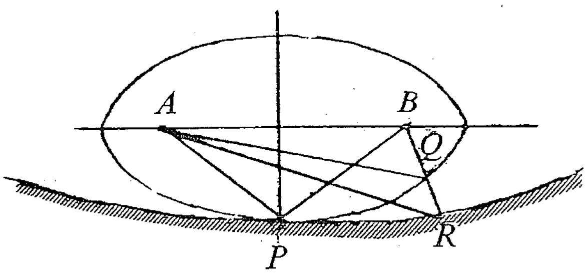


Figure 1

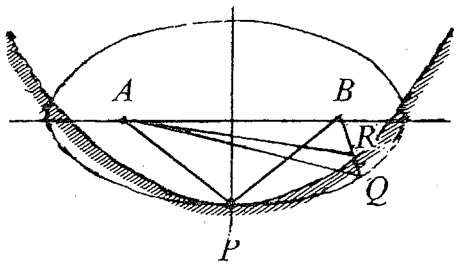


Figure 2

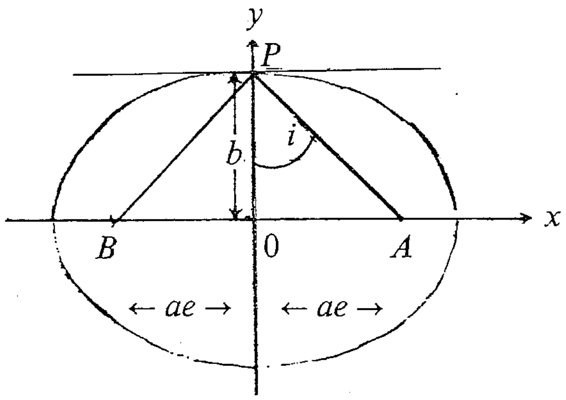


Figure 3

We now determine this critical curvature in terms of the angle of incidence i and the perpendicular distance b of the object A from the tangent line to the curve of the mirror at the point of incidence P. For the two-dimensional case we have, in the notation of Figure 3,

$$\sin i = ae/a = e, \tag{2}$$

where e is the eccentricity of the ellipse with foci at the object A and its image B in the perpendicular to the mirror at P, and a is its semi-major axis.

The curvature ρ of the ellipse at any point (x,y) is given by

$$\rho = \frac{1}{a^2b^2\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}\right)^{3/2}}. \tag{3}$$

At the point of incidence P $(0, b)$ this reduces to

$$\rho = \frac{b}{a^2} = \frac{1-e^2}{b} = \frac{\cos^2 i}{b} \tag{4}$$

on using (2). This is the critical value sought. If the curvature of the mirror at the point of incidence is less than this the optical path will be a minimum, otherwise it will be a maximum.

Yours sincerely,
J. M. H. PETERS
(Department of Mathematics, Liverpool Polytechnic)

Dear Editor,

Mathematics: its stylistics and its ethics

It is more than the picaresqueness of Professor John Pym's English style to which Mr C. W. Puritz directs attention in his letter in Volume 13 No. 3. He no more than points, but nevertheless importantly, to the need for mathematicians, students and professionals, pure and applied, to give and to be seen to give, more and deeper consideration to the many moral implications and decisions inherent in all kinds of mathematical practice. Yet his letter is unconvincing; it commits him to censorship of printing; it implies that students in English departments of schools and colleges must not continue to study the Bible, or such works as those of Browning, Burns, Chaucer, Defoe, Fielding, Orwell, Shakespeare, Smollett and Spenser; it implies that you should not publish this letter. Mr Puritz convicts you and Professor Pym of introducing licence, as, in his day, John Milton was condemned; but, on the one hand as he makes clear in *Areopagitica* (1644), Milton was firmly opposed to licensing (that is, 'that no Book, pamphlet, or paper, shall be henceforth Printed, unless the same be first approved and licensed'), whilst on the other hand, Mr Puritz would, at any rate of books, pamphlets and papers mathematical (of which he, at least stylistically, disapproved), forbid their printing and publication, so silencing their authors. Mr Puritz would have licensing of the literature of mathematics.

Yours sincerely,
G. N. COPLEY
(4 Manor Close, Templecombe, Somerset BA8 0LA)

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 home address and also the postal address of your school, college or university.

Problems

14.1. A straight rod is cut at random into three pieces. What is the probability that one of these pieces is at least half as long as the original rod?

14.2. (Submitted by Victor Bryant, University of Sheffield—see his article in this issue) There are five people in a room. Given any two of them, there is just one other who is a friend of them both. Show that there is just one person who is a friend of all the rest. Having established the result for five people, extend your solution to the case of seven people. [In fact, the result applies to any odd number of people.]

14.3. (Submitted by Ian Macdonald, University of Stirling) Given four distinct primes p_1, p_2, p_3, p_4 and four integers q_1, q_2, q_3, q_4, prove that the determinant

$$\begin{vmatrix} p_1\alpha_1 + q_1 & p_2\alpha_2 + q_2 \\ p_3\alpha_3 + q_3 & p_4\alpha_4 + q_4 \end{vmatrix}$$

takes all integral values as the integers $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ vary.

Solutions to Problems in Volume 13, Number 2

13.4. A positive integer is said to be *normal* if the digits occurring in its decimal representation are all different (e.g. 123 is normal, but 122 is abnormal). Determine the positive integer or integers n such that there are equally many normal and abnormal numbers smaller than n.

Solution

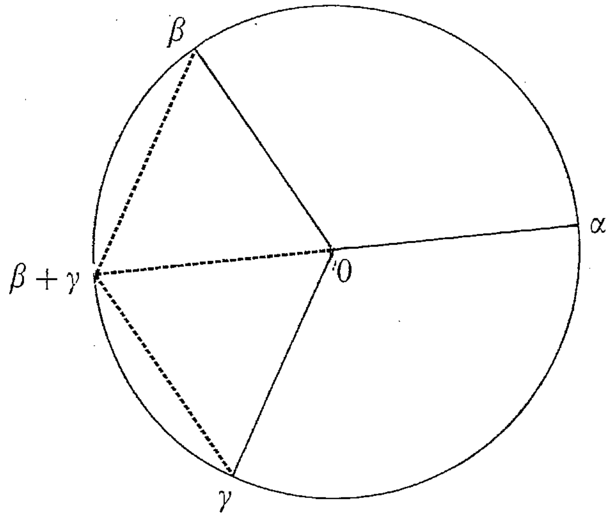
We received no solutions to this problem, which was submitted to us by Roger Webster. A few moments' thought should convince you that a sophisticated argument will not be the order of the day, and that the best way of tackling the problem will be to start counting normal and abnormal numbers. The table gives the numbers of such in various ranges. Thus the numbers of normal and of abnormal numbers smaller than 11221 are both 5610. It should be clear that the abnormal numbers will occur more frequently than the normal numbers from this point on, and that this is the only such positive integer.

Range	Number of normal numbers	Number of abnormal numbers
1–9	9	0
10–99	$9 \times 9 = 81$	9
100–999	$9 \times 9 \times 8 = 648$	252
1000–9999	$9 \times 9 \times 8 \times 7 = 4536$	4464
10000–10999	$8 \times 7 \times 6 = 336$	664
11000–11220	0	221

13.5. Let n be a natural number ≥ 3, and let α, β, γ be complex numbers such that $\alpha^n = \beta^n = \gamma^n = 1$, $\alpha + \beta + \gamma = 0$. Prove that n is a multiple of 3.

Solution

We received just one solution to this problem, from Sudhir Jain, a third-year student at Imperial College, London. We present essentially his solution here. The points α, β, γ will be on the circumference of the unit circle in the Argand diagram. Because $\alpha = -\beta - \gamma$, and



complex add according to the parallelogram law in the Argand diagram, it is clear that the triangle formed by 0, β and $\beta + \gamma$ is an equilateral triangle, as is the triangle formed by 0, γ, $\beta + \gamma$. Thus the angle between 0β and 0γ is 120°. Thus

$$\gamma = e^{(1/3)2\pi i}\beta,$$

so

$$\gamma^n = e^{(1/3)2\pi n i}\beta^n.$$

But $\beta^n = \gamma^n = 1$, so

$$e^{(1/3)2\pi n i} = 1,$$

and n must be a multiple of 3.

A variant of this argument, which avoids appeal to geometry, is to say that

$$\left(\frac{\alpha}{\gamma}\right)^n = \left(\frac{\beta}{\gamma}\right)^n = 1, \qquad \frac{\alpha}{\gamma} + \frac{\beta}{\gamma} + 1 = 0,$$

so we can replace α, β, γ by α/γ, β/γ, 1 and thus assume that $\gamma = 1$. If we put $\alpha = x + iy$, $\beta = x' + iy'$ (x, y, x', y' real), then, because now $\alpha + \beta = -1$, we have

$$y + y' = 0 \qquad \text{and} \qquad x + x' = -1.$$

Since $|\alpha| = |\beta| = 1$, we also have $x, x' = \pm\sqrt{1-y^2}$. It follows that

$$x = x' = -\sqrt{1-y^2} = -\tfrac{1}{2},$$

so that $y = \pm(\sqrt{3}/2)$. Thus α, β are $(-1 \pm i\sqrt{3})/2 = e^{\pm 2\pi i/3}$. Since $\alpha^n = 1$, we easily see that $3|n$.

13.6. Prove that $p(n) \geq 2^{[\sqrt{n}]}$ for all integers $n > 2$, where $[\sqrt{n}]$ denotes the integral part of $\sqrt{n}$ and $p(n)$ denotes the number of ways of partitioning n as a sum of positive integers.

Solution

This problem arose out of the article by C. M. Shiu and P. Shiu on partitioning problems in Volume 13, Number 2. We received no solutions to the problem, perhaps not surprisingly, so here is our solution. The result can be verified directly for $n \leq 9$. For example, when $n = 9$ $2^{[\sqrt{n}]} = 2^3 = 8$, and 9 certainly has at least 8 partitions, for example

$$\begin{aligned}
9 &= 9\\
9 &= 8+1\\
9 &= 7+2\\
9 &= 6+3\\
9 &= 5+4\\
9 &= 7+1+1\\
9 &= 6+2+1\\
9 &= 6+1+1+1.
\end{aligned}$$

In fact, you could easily write down others. Now suppose that $n > 9$. If $a_1, \ldots, a_r$ are distinct numbers selected from $1, 2, \ldots, [\sqrt{n}]$, then

$$\begin{aligned}
n - (a_1 + \cdots + a_r) &\geq n - (1 + 2 + \cdots + [\sqrt{n}])\\
&= n - \tfrac{1}{2}[\sqrt{n}]([\sqrt{n}] + 1)\\
&\geq n - \tfrac{1}{2}\sqrt{n}(\sqrt{n} + 1)\\
&= \tfrac{1}{2}\sqrt{n}(\sqrt{n} - 1)\\
&> \sqrt{n} \text{ when } n > 9\\
&\geq [\sqrt{n}].
\end{aligned}$$

Hence

$$n = a_1 + \cdots + a_r + (n - a_1 - \cdots - a_r)$$

provides a partition of n into positive integers such that every two distinct choices of $a_1, \ldots, a_r$ provide distinct partitions. The number of ways of choosing $a_1, \ldots, a_r$ is given by a binomial coefficient, so

$$\begin{aligned}
p(n) &\geq \binom{[\sqrt{n}]}{0} + \binom{[\sqrt{n}]}{1} + \cdots + \binom{[\sqrt{n}]}{[\sqrt{n}]}\\
&= (1+1)^{[\sqrt{n}]}\\
&= 2^{[\sqrt{n}]},
\end{aligned}$$

as required.

Book Reviews

Mathematics: The Loss of Certainty. By MORRIS KLINE. Oxford University Press, 1980. Pp. 366. £11·50.

Morris Kline has deservedly established a reputation as a clear, lucid interpreter of the history of mathematics. In this book he maintains his high standard and quality of writing as he outlines, from the time of the Greeks to the present day, the ways in which mathematicians have attempted to explain and secure the foundations of their subject. He traces the history through moments of near triumph when the foundations seemed almost secured and periods of utter bewilderment when the shattering of accepted ideas threw the whole of mathematics into confusion. The inextricable but often elusive links between mathematics and nature are carefully described and explored in a narrative which, whilst always clear and easy to read, does demand a certain minimum pre-knowledge of the history of mathematics.

For the intelligent sixth former who wished to seek answers to questions such as, 'What is mathematics?' or 'Where does mathematics begin?' this book would provide interesting and absorbing reading. The history of mathematics is usually sadly neglected both at school and university levels—some systematic study of a standard history of mathematics text together with this book would give many students a most valuable background and such a study might profitably be undertaken in the break between school and university mathematics. Many potential mathematics undergraduates (and indeed many mathematics teachers) would find Kline's latest book valuable, informative and stimulating.

University of Durham — M. L. CORNELIUS

Mathematics: Problem Solving through Recreational Mathematics. By BONNIE AVERBACH and ORIN CHEIN. W. H. Freeman & Co. Ltd., Oxford, 1980. Pp. 400. £8·30 Hard covers.

This is an interesting American book. It is essentially the text of a one- or two-semester college course developed to involve the student actively in mathematics. The authors have made a definite decision to break away from the familiar pattern of students' watching the teacher, learning some techniques and then practising them on similar problems over and over again. The aim is to get the student thinking for himself and developing his own problem-solving strategies. They have used games, puzzles and recreational problems because these are fun and readily provide motivation. The book, as a result, is a useful source of such puzzles and problems, though many similar ones can be found elsewhere. This book is different in that it also looks at the strategies for attacking the problems and attempts to explain the problem-solving procedures.

Each chapter deals with a different topic area (e.g. solve it with logic; solve it with networks; solitaire games and puzzles) and is more or less independent of others. The chapters are carefully structured. Each starts off with sample problems for the student to try, and these are then solved and discussed in the text. There are practice problems to 'firm up the mathematical ideas presented' and also more challenging exercises.

Personally, I have enjoyed having this book to dip into, and have used several problems with my sixth formers. They have enjoyed them, and we have had good class discussions about the strategies involved in the solutions. At first it seemed expensive, but it has a hard cover (with useful grids for games printed on the inside covers) and should be a useful source book for a departmental library.

Accrington and Rossendale College — K. CROSS

CONTENTS

1 A new cover for *Mathematical Spectrum*

1 Bonds of friendship: VICTOR BRYANT

6 Mathematics in India in the Middle Ages: CHANDRA B. SHARMA

9 How infectious diseases spread in households: NIELS BECKER

14 Blind student tests new computer equipment

16 Continuous transformation groups: W. D. HALFORD

26 Letters to the editor

29 Problems and solutions

32 Book reviews

ISSN 0025–5653

PRICES (*postage included*)

Prices for Volume 14 (*Issues Nos.* 1, 2, *and* 3):

Subscribers in Britain and Europe: £2·30
Subscribers overseas: £4·60 (US$11.00; $A9.50)
(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:

All back issues are available; information concerning prices may be obtained from the Editor.

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

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

Published by the Applied Probability Trust
Printed in England by Galliard (Printers) Ltd, Great Yarmouth