MATHEMATICAL SPECTRUM

A MAGAZINE FOR STUDENTS AT SCHOOLS
COLLEGES AND UNIVERSITIES

Volume 15 1982/83 Number 2

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 15 of *Mathematical Spectrum* will consist of three issues, of which this is the second. The first issue was published in September 1982 and the third will appear in May 1983.

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.

Mathematical Spectrum Award for Volume 14

We remind our readers that each year two prizes are available to contributors who are still at school or are students in colleges or universities. A prize of £20 is for an article published in the magazine and another of £10 is for a letter or the solution of a problem.

There were no articles in Volume 14 by authors eligible for the £20 prize. The £10 prize has been awarded to Nigel McCann for his letter *Roots of polynomials* (Volume 14, page 55). We look forward to further contributions from readers.

A Survey of Mathematical Puzzles III

KEITH AUSTIN, *University of Sheffield*

Keith Austin is a Lecturer in Pure Mathematics at the University of Sheffield. He has presented a monthly collection of mathematical puzzles on Sheffield's independent radio station and writes Brain-teasers for the *Sunday Times*. He is interested in the question of whether puzzles can be used to improve mental fitness in the same way that exercises are used to improve physical fitness.

In the second article in this series† we considered finite mathematical puzzles, and so in this final article we shall look at three other types of mathematical puzzles. The answers to the puzzles will be given at the end of the article.

1. Puzzles involving time, distance, quantity and money

Whereas finite puzzles have dominated the scene in recent years, the puzzles in this section were previously the more popular. Perhaps it is a sign of the times that we have moved from more leisurely days to the precisely and explicitly defined way of life of the computer age.

Generally, a finite puzzle requires the sorting and rearranging of the information given. This usually involves no mathematical theory but does take quite a time. On the other hand, the puzzles in this section often require an elementary knowledge of arithmetic, in particular the idea of proportion, but the time needed for solution is

† The first and second articles appeared in Volume 14 Number 2 and Volume 15 Number 1 of *Mathematical Spectrum*.

short once the method is clear. Two consequences of this are: (a) it can be difficult to draw the line between a puzzle and a problem in the theory of proportion, (b) these puzzles are more suitable than the finite puzzles when there is only a short time for the solution. For example, I had a monthly puzzle programme on our local radio station, Radio Hallam, where I gave the puzzle and then, while a record was being played, the listeners rang in with their answers. In these programmes I used a good number of puzzles involving time, distance, etc.

The puzzles in this section are usually solvable by algebra, but it is an interesting exercise to try to avoid algebra and find an answer using only arithmetic and logic. For the purposes of a radio programme, algebraic solutions are not suitable for giving over the air.

Examples.

1. A man returns from work on the train and arrives at his local station at 5.00. His wife then drives him home and they arrive at 5.12. One day, his wife does not leave home until 5.00 and so he sets off walking to meet her. When they meet he gets into the car and they arrive home at 5.20. On another day he walks all the way home from the station. At what time does he arrive home?
2. Mrs. Arnold and Mrs. Brown bought some apples in the morning; Mrs. Arnold paid $6\frac{1}{2}$p for each apple and Mrs. Brown 7p; Mrs. Arnold bought 2 apples. In the afternoon Mrs. Arnold bought an apple for 1p and Mrs. Brown bought one for 2p. At the end of the day Mrs. Brown found that the average price she had paid for apples overall was less than Mrs. Arnold's average. How many apples did Mrs. Brown buy in the morning?

2. Whole number puzzles

These puzzles concern objects which are not divided, and so we have the additional clue that the answers are whole numbers.

Examples.

1. A delivery van contains some £5 shirts, some £1 records and some 5p postcards. There are 100 items altogether, and the total cost is £100. How many are there of each item?
2. At a party the guests divide into groups of 10 and find there are 9 over. They divide into groups of 9 and find 8 over, into groups of 8 and find 7 over . . ., into groups of 3 and find 2 over, into groups of 2 and find 1 over. There are fewer than 3000 at the party. How many are there at the party?

3. Puzzles about knowing

These puzzles usually involve a group of people who have hats put on their heads, numbers pinned on their backs, or stamps stuck on their foreheads. They cannot see the objects on their own persons, but they can see those on other people—either all the other people or, if they are standing in a queue, some other people. This is a modern type of puzzle and has developed considerably in recent years.

Examples.

1. There are a number of people in a room. Each is wearing a red or blue hat, but he cannot see the colour of his own hat. At 12.00 they are told that there is at least one red hat in the room. At 1.00 they are asked if each person who knows the colour on his own hat will raise his hand. At 2.00 the same request is made. Similarly at 3.00, 4.00 and so on. What happens?

2. Alan has 7 pinned on his back and Bryan has 5. Each can see the other's number but not his own. They are told the numbers are positive integers and the total is 10 or 12. Alan is asked if he knows his own number, and Bryan hears his answer. Bryan is asked if he knows his number and Alan hears his answer. So they go on, each being asked in turn. What happens?

Further reading

The monthly magazine *Scientific American* until recently had a regular article entitled 'Mathematical Games' by Martin Gardner which considered puzzles. Over the years, these articles have been collected into books, and some of these are available in paperback from Penguin Books. You will probably find *Scientific American* in your local library.

ANSWERS TO THE PUZZLES

1. Puzzles involving time, distance, quantity and money

1. Let the distance from the station to home be 1 unit of distance. Then the car's speed is 5 units per hour (u.p.h.). It takes 10 minutes for the car and the man to cover the 1 unit and so the sum of their speeds is 6 u.p.h. So the man's speed is 1 u.p.h. and he arrives home at 6.00.

2. 1 apple. So Mrs. Arnold bought 3 for 14p giving an average of $4\frac{2}{3}$p and Mrs. Brown bought 2 for 9p giving an average of $4\frac{1}{2}$p. If Mrs. Brown bought 2 at 7p then she would have 3 for 16p giving an average of $5\frac{1}{3}$p, and the purchase of more at 7p would only increase that average.

2. Whole number puzzles

1. 19 shirts, 1 record, 80 postcards. We require a number of shirts and postcards such that the total number of items equals the total cost in pounds. Also we must have the postcards in 20's. 20 postcards and 4 shirts means 24 items for £21. Now a shirt produces 4 more pounds than items so take 4 lots of the 24 items for £21 giving 96 items for £84. Now add 3 shirts making the total 99 items for £99. Finally we have 1 record.

Note. This is a traditional puzzle set by Arab, Chinese and Indian puzzlers and dating back to the 6th century in some form or other.

2. 2519 guests. If there had been one more guest then the total would have been divisible by 10, 9, 8, . . ., 3 and 2. Hence it would have been divisible by 2520 and so that would have been the number of guests.

3. Puzzles about knowing

1. What happens depends on the number of red hats in the room. If there is 1 red hat then its wearer raises his hand at 1.00. If there are 2 red hats then their wearers know at 1.05 that there is not 1 red hat, as no hand went up at 1.00, and so they raise their hands at 2.00. Similarly, if there are 3 red hats, then their wearers raise their hands at 3.00, and so on. If there are any blue hats then their wearers raise their hands one hour after the red hat wearers raise theirs.

2. We draw a diagram showing all the possibilities. We only use the fact that is known to both Alan and Bryan, namely that the sum is 10 or 12.

Alan's number (rows), Bryan's number (columns); √ denotes a possibility

11	√										
10		√									
9	√		√								
8		√		√							
7			√		√						
6				√		√					
5					√		√				
4						√		√			
3							√		√		
2								√		√	
1									√		√
	1	2	3	4	5	6	7	8	9	10	11

Bryan's number

Alan knows the answer is in column 5 but as there are 2 ticks he says he does not know his number. From Alan's answer, everybody can cross off columns 10 and 11 as they only contain one tick. (Cover them with a piece of paper.) Bryan knows the answer is in row 7, but as there are 2 ticks he says he does not know his number. From Bryan's answer, everybody can cross off rows 1, 2, 10 and 11 as they only contain one tick. Alan knows the answer is in column 5 so he says he does not know his number. Everybody crosses off columns 1, 2, 8 and 9. Bryan says he does not know his number. Everybody crosses off rows 3, 4, 8 and 9. Alan says he does not know his number. Everybody crosses off columns 3, 4, 6 and 7. Bryan says he knows his number. Alan says he does not know his number.

Continued Fractions

K. E. HIRST, *University of Southampton*

Dr Hirst has been a Lecturer in Mathematics at the University of Southampton since 1965. His research interests began in analysis and number theory and have since developed into the field of mathematical education. He has lectured several times at the annual conference of the Association of Teachers of Mathematics and has published articles in their journal *Mathematics Teaching*.

His chief recreation is music, particularly choral singing.

The representation of a number with which you will be most familiar is the decimal expansion, where numbers are expressed by the use of powers of 10. For example

$$\tfrac{13}{8} = 1{\cdot}625; \qquad \tfrac{17}{7} = 2{\cdot}\overline{428571},$$

where the digits following the decimal point are repeated periodically;

$$\sqrt{2} = 1{\cdot}4142\ldots,$$

a decimal which has no regular pattern.

Continued fractions provide another way of representing numbers, and I shall introduce them using the above examples.

$$\tfrac{13}{8} = 1 + \tfrac{5}{8} = 1 + \frac{1}{\tfrac{8}{5}} = 1 + \frac{1}{1+\tfrac{3}{5}}$$

$$= 1 + \frac{1}{1+\frac{1}{\tfrac{5}{3}}} = 1 + \frac{1}{1+\frac{1}{1+\tfrac{2}{3}}}$$

$$= 1 + \frac{1}{1+\frac{1}{1+\frac{1}{\tfrac{3}{2}}}} = 1 + \frac{1}{1+\frac{1}{1+\frac{1}{1+\tfrac{1}{2}}}}$$

$$= 1 + \frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\tfrac{1}{1}}}}}$$

This is obviously a cumbersome notation, although it makes clear the choice of the name 'continued fraction'. Other notations in common use are

$$1 + \frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1} \qquad \text{or} \qquad [1;1,1,1,1,1].$$

With the example above we could choose to stop at the penultimate stage, to obtain

$$1 + \frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{2} \qquad \text{or} \qquad [1;1,1,1,2].$$

There is always this choice with a terminating continued fraction. For the second example you should be able to check that $\frac{17}{7} = [2;2,3] = [2;2,2,1]$. For the third example we need to know that $\sqrt{2}$ lies between 1 and 2. We then find that

$$\sqrt{2} = 1 + (\sqrt{2}-1) = 1 + \frac{1}{\dfrac{1}{\sqrt{2}-1}} = 1 + \frac{1}{\sqrt{2}+1}$$

$$= 1 + \frac{1}{2+(\sqrt{2}-1)} = \cdots = 1 + \frac{1}{2+}\,\frac{1}{2+}\cdots\frac{1}{2+(\sqrt{2}-1)};$$

the process will continue without stopping. We write $[1;\bar{2}]$ for the formal periodic continued fraction that we have obtained in this way.

This may at first sight seem a rather simple-minded kind of thing to do, but it turns out to have useful properties, and connections with other parts of mathematics. The first of these is the Euclidean algorithm for finding the highest common factor (HCF) of two numbers.

A very readable book on this topic is *Continued Fractions* by C. D. Olds, published by Random House in their New Mathematical Library series. Most of the details omitted below will be found there, with some other interesting material.

The Euclidean algorithm

I shall illustrate the working of this with a numerical example. The general method and the theory can be found in almost any book on elementary number theory, and is in Olds's book. It is also a topic which used to be in school mathematics syllabuses many, many years ago.

To find the HCF of 707 and 301 the algorithm proceeds as follows:

$$\begin{aligned} 707 &= 2 \times 301 + 105 \\ 301 &= 2 \times 105 + 91 \\ 105 &= 1 \times 91 + 14 \\ 91 &= 6 \times 14 + 7 \\ 14 &= 2 \times 7. \end{aligned}$$

The HCF of 707 and 301 is 7. We can use the successive divisions to find the continued fraction for 707/301:

$$\frac{707}{301} = 2 + \frac{105}{301}, \quad \frac{301}{105} = 2 + \frac{91}{105}, \quad \frac{105}{91} = 1 + \frac{14}{91}, \quad \frac{91}{14} = 6 + \frac{7}{14}, \quad \frac{14}{7} = 2.$$

So the continued fraction is

$$\frac{707}{301} = 2 + \frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{6+}\,\frac{1}{2}.$$

To evaluate the continued fraction on the right we start from the right-hand end and proceed as follows:

$$2 + \frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{6+}\,\frac{1}{2} = 2 + \frac{1}{2+}\,\frac{1}{1+}\,\frac{2}{13} = 2 + \frac{1}{2+}\,\frac{13}{15} = 2 + \frac{15}{43} = \frac{101}{43},$$

which is 707/301 in its lowest terms. What we shall now do is develop a bit of the theory to see why we always get a fraction in its lowest terms, and to find out a better way of evaluating continued fractions without starting at the right-hand end each time.

General continued fractions

We can adapt the procedure used for rational numbers such as 707/301 to work for a non-rational number x. We can write $x = a_1 + r_1$, where a_1 is an integer and

$0 < r_1 < 1$. Since $0 < r_1 < 1$; $1/r_1 > 1$. Writing $x_1 = 1/r_1$, $x = a_1 + (1/x_1)$, and x_1 is non-rational. We can now do the same thing with x_1 to obtain

$$x = a_1 + \frac{1}{a_2 +}\,\frac{1}{x_2},$$

where $x_2 > 1$ is a non-rational number, and a_2 is a positive integer. This process can be continued to give

$$x = a_1 + \frac{1}{a_2 +}\,\frac{1}{a_3 +}\cdots\frac{1}{a_n +}\,\frac{1}{x_n}.$$

Now suppose we consider the infinite continued fraction

$$a_1 + \frac{1}{a_2 +}\,\frac{1}{a_3 +}\cdots\frac{1}{a_n +}\cdots.$$

We certainly cannot work this out from the right-hand end, because there isn't one. This raises two questions:

1. Can we attach a meaning to an infinite continued fraction?
2. Will it be equal in any sense to the number x with which we started?

These questions are answered by the use of convergents.

Convergents

The finite continued fraction

$$c_n = a_1 + \frac{1}{a_2 +}\cdots\frac{1}{a_n}$$

is called the *nth convergent* of the corresponding infinite (or longer finite) continued fraction. Because $a_1, a_2, \ldots, a_n$ are integers, this will always work out as a rational fraction p_n/q_n, and if the first few are worked out we can see how they are related. Now

$$c_1 = a_1 = \frac{a_1}{1} = \frac{p_1}{q_1},$$

$$c_2 = a_1 + \frac{1}{a_2} = \frac{a_1a_2 + 1}{a_2} = \frac{p_2}{q_2}.$$

You should do the algebra yourself to verify that

$$c_3 = \frac{a_1a_2a_3 + a_1 + a_3}{a_2a_3 + 1} = \frac{p_3}{q_3}.$$

However, we can rearrange this in the form

$$\frac{p_3}{q_3} = \frac{a_3(a_1a_2 + 1) + a_1}{a_3(a_2) + 1} = \frac{a_3p_2 + p_1}{a_3q_2 + q_1}.$$

You should also verify the algebra to show that

$$c_4 = \frac{p_4}{q_4} = \frac{a_4 p_3 + p_2}{a_4 p_3 + p_1}.$$

In fact it can be proved in general that

$$p_{n+1} = a_{n+1} p_n + p_{n-1}, \qquad q_{n+1} = a_{n+1} q_n + q_{n-1}, \tag{1}$$

and this is done in Olds's book.

We are now in a position to explain why a finite continued fraction always works out in its lowest terms. We do this by showing that, for $n = 1, 2, 3, \ldots$,

$$p_{n+1} q_n - q_{n+1} p_n = (-1)^{n+1}. \tag{2}$$

This uses the method of proof by mathematical induction. For $n = 1$, the left-hand side is

$$p_2 q_1 - q_2 p_1 = (a_1 a_2 + 1) . 1 - a_2 . a_1 = 1 = (-1)^{1+1}.$$

We then have

$$\begin{aligned} p_{n+2} q_{n+1} - q_{n+2} p_{n+1} &= (a_{n+2} p_{n+1} + p_n) q_{n+1} - (a_{n+2} q_{n+1} + q_n) p_{n+1} \\ &= -(p_{n+1} q_n - q_{n+1} p_n). \end{aligned} \tag{3}$$

This relationship shows that, *if*

$$p_{n+1} q_n - q_{n+1} p_n = (-1)^{n+1},$$

then

$$p_{n+2} q_{n+1} - q_{n+2} p_{n+1} = -(-1)^{n+1} = (-1)^{n+2}.$$

The result is therefore established by induction. It is a nice example of proof by induction because we are clearly not 'assuming what we want to prove', since the essential step is at (3) which relates successive left-hand sides of (2) without involving the right-hand side at all.

Now a number which is a divisor of both p_n and q_n is also a divisor of

$$p_{n+1} q_n - q_{n+1} p_n,$$

and therefore divides $(-1)^{n+1}$. The only possible common divisors of p_n and q_n are therefore ± 1, so that p_n/q_n will always be in its lowest terms.

Using the formulae (1) and (2), you can now verify that

$$c_{n+1} - c_n = \frac{(-1)^{n+1}}{q_n q_{n+1}} \tag{4}$$

and

$$c_{n+1} - c_{n-1} = \frac{a_{n+1}(-1)^n}{q_{n+1} q_{n-1}}. \tag{5}$$

Putting $n=1$ in (4), $n=2$ in (4) and $n=1$ in (5), we have

$$c_2 - c_1 = 1/q_1q_2 > 0$$
$$c_3 - c_2 = -1/q_2q_3 < 0$$
$$c_3 - c_1 = a_3/q_3q_1 > 0,$$

since $a_3 > 0$ and all the q's are greater than 0.

This tells us that $c_1 < c_3 < c_2$. Applying successive values of n in turn, we have

$$c_3 < c_4 < c_2,$$
$$c_3 < c_5 < c_4,$$
$$c_5 < c_6 < c_4,$$
$$c_5 < c_7 < c_6, \text{ etc.}$$

So we find the convergents related as follows:

$$c_1 < c_3 < c_5 < \cdots < c_{2n+1} < \cdots < c_{2n} < \cdots < c_6 < c_4 < c_2.$$

You should work out some numerical examples to illustrate this.

Now, since $a_{n+1} \geq 1$,

$$q_{n+1} = a_{n+1}q_n + q_{n-1} \geq q_n + q_{n-1}, \tag{6}$$

so

$$q_3 \geq q_2 + q_1 \geq 2,$$
$$q_4 \geq q_3 + q_2 \geq 3,$$
$$q_5 \geq q_4 + q_3 \geq 5,$$
$$q_6 \geq q_5 + q_4 \geq 8, \text{ etc.}$$

(You may recognise the sequence $2, 3, 5, 8, \ldots$ as the Fibonacci sequence†; it is connected with the continued fraction

$$1 + \frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\cdots$$

all of whose entries are equal to 1.) It follows from (6) that the q's form an increasing sequence of integers, and hence

$$c_{n+1} - c_n = \frac{(-1)^n}{q_{n+1}q_n} \to 0 \qquad \text{as } n \to \infty. \tag{7}$$

† The Fibonacci sequence is characterised by the property that each term is the sum of the two preceding terms. It has many interesting properties and applications, a few of which are given in S.M.P. Book D, Chapter 12. See also R. J. Webster's article in *Mathematical Spectrum*, Volume 3, Number 2 (1970) entitled 'The Legend of Leonardo of Pisa'.

Finally we can show that the sequence of convergents does in fact converge and has for its limit the number x which we started from. We have

$$c_n = a_1 + \frac{1}{a_2 +}\,\frac{1}{a_3 +}\cdots\frac{1}{a_{n-1} +}\,\frac{1}{a_n}$$

$$x = a_1 + \frac{1}{a_2 +}\,\frac{1}{a_3 +}\cdots\frac{1}{a_{n-1} +}\,\frac{1}{x_{n-1}}$$

$$c_{n+1} = a_1 + \frac{1}{a_2 +}\,\frac{1}{a_3 +}\cdots\frac{1}{a_{n-1} +}\,\frac{1}{a_n +}\,\frac{1}{a_{n+1}},$$

where $x_{n-1} > 1$ and is a non-rational number. Recall that to continue the process from the expression for x above, we write $x_{n-1} = a_n + r_n$, where a_n is a positive integer, and where the remainder r_n satisfies $0 < r_n < 1$. For the next step we have

$$\frac{1}{r_n} = x_n = a_{n+1} + r_{n+1}.$$

So we conclude that

$$a_n < x_{n-1} < a_n + \frac{1}{a_{n-1}},$$

i.e.

$$\frac{1}{a_n} > \frac{1}{x_{n-1}} > \frac{1}{a_n +}\,\frac{1}{a_{n+1}}.$$

This tells us that x lies between c_n and c_{n+1}, with $c_n < x < c_{n+1}$ if n is odd and $c_{n+1} < x < c_n$ if n is even. Since we have shown that $c_{n+1} - c_n \to 0$ it follows that $c_n \to x$ as $n \to \infty$.

Approximations

One of the main applications of continued fractions is to the provision of good approximations. I shall illustrate this in connection with the number π. The first few terms were obtained by the German–Swiss mathematician J. H. Lambert in 1770. (It was Lambert who proved that π is an irrational number.) We have

$$\pi = 3 + \frac{1}{7 +}\,\frac{1}{15 +}\,\frac{1}{1 +}\,\frac{1}{292 +}\,\frac{1}{1 +}\cdots$$

(Olds quotes the first 23 places.)

The formulae (1) enable us to work out the first few convergents. They can be set out in a table as follows:

n	1	2	3	4	5
a_n	3	7	15	1	292
p_n	3	22	333	355	103993
q_n	1	7	106	113	33102

As an illustration, with $n=4$, we have

$$p_5 = a_5 p_4 + p_3,$$

i.e. $103993 = 292 \times 355 + 333$.

Now $c_1 = 3$, and $c_2 = 22/7$ (which is a very familiar approximation to π), $c_5 = 355/133 = 3{\cdot}14159292\ldots$, and the first six decimal places are correct. The next convergent $c_6 = 103993/33102 = 3{\cdot}14159265301\ldots$, for which the first *nine* decimal places are correct.

In general, if p/q is a decimal approximation to a number x (so that $q = 10^n$), we have

$$\left| x - \frac{p}{q} \right| < \frac{1}{q}.$$

However, if p/q is a convergent to the continued fraction for x, then it follows from (7) that

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2},$$

and the presence of the q^2 means that we get very good approximations.

An interesting application of this result is the following. Consider the sequence of numbers

$$\sin\theta, \sin 2\theta, \sin 3\theta, \ldots, \sin n\theta, \ldots \qquad (8)$$

If θ is a rational multiple of π, then (8) contains infinitely many zeros. If θ is not a rational multiple of π, then (8) contains no zero terms, but it does contain terms which are arbitrarily close to 0. For, if p_n/q_n is the sequence of convergents of the continued fraction for the number θ/π, then

$$\left| \frac{\theta}{\pi} - \frac{p_n}{q_n} \right| < \frac{1}{q_n^2},$$

so that

$$|q_n\theta - p_n\pi| < \frac{\pi}{q_n},$$

and therefore $q_n\theta - p_n\pi \to 0$ as $n \to \infty$. The numerator p_n is an integer so $\sin p_n\pi = 0$. It follows that $\sin q_n\theta \to 0$ as $n \to \infty$. (Notice that this does not mean that $\sin n\theta \to 0$ as $n \to \infty$. The numbers q_n do not comprise the whole sequence of integers. In fact it is a nice exercise to use the identity for $\sin(n+1)\theta$ to show that $\sin n\theta$ does not have a limit at all.)

Periodic continued fractions

When we worked out the continued fraction for $\sqrt{2}$ at the beginning of this article, we saw that it was periodic. One of the remarkable facts about a continued fraction is that it is eventually periodic if and only if it represents a quadratic irrational number $(P+\sqrt{D})/Q$, where P, Q and $D>0$ are integers. The various parts of this result were proved in the late eighteenth and early nineteenth centuries by the French mathematicians Lagrange, Legendre, and Galois. Particularly interesting are the continued fractions for square roots themselves. As an example to illustrate their behaviour,

$$\sqrt{43}=[6;\overline{1,1,3,1,5,1,3,1,1,12}].$$

Now $6^2<43<7^2$, and the first digit is 6. The last digit of the repeating part is $12=2\times 6$, and this doubling occurs for all square roots. The remainder of the repeating part is symmetric about the centre digit, and this symmetry is always

TABLE 1

D	Continued fraction for $\sqrt{D}$	D	Continued fraction for $\sqrt{D}$
2	$[1;\overline{2}]$	17	$[4;\overline{8}]$
3	$[1;\overline{1,2}]$	18	$[4;\overline{4,8}]$
		19	$[4;\overline{2,1,3,1,2,8}]$
5	$[2;\overline{4}]$	20	$[4;\overline{2,8}]$
6	$[2;\overline{2,4}]$	21	$[4;\overline{1,1,2,1,1,8}]$
7	$[2;\overline{1,1,1,4}]$	22	$[4;\overline{1,2,4,2,1,8}]$
8	$[2;\overline{1,4}]$	23	$[4;\overline{1,3,1,8}]$
		24	$[4;\overline{1,8}]$
10	$[3;\overline{6}]$		
11	$[3;\overline{3,6}]$		
12	$[3;\overline{2,6}]$		
13	$[3;\overline{1,1,1,1,6}]$		
14	$[3;\overline{1,2,1,6}]$		
15	$[3;\overline{1,6}]$		

present, sometimes with two central digits. Studying a table of continued fractions for square roots shows many other patterns, some of which you can prove for yourself. (See Table 1.) Quite extensive tables have been compiled, in particular by W. Patz in 1941, who took square roots from 1 to 10000. Needless to say, computers have been involved, and in 1974 it was reported that the square root of the prime number

$$p=26437680473689$$

has a continued fraction with a period length equal to 18331889.

Minimum Roadway Problems

CYRIL ISENBERG, *University of Kent at Canterbury*

Dr Isenberg obtained his B.Sc. and Ph.D. degrees at King's College London. From 1965 to 1967 he was a Resident Research Associate at the Argonne National Laboratory in Chicago. In 1967 he returned to Britain as a lecturer in theoretical physics at the University of Kent at Canterbury. His research interests are in the fields of solid-state and liquid-state physics.

Since 1973 he has popularized the science and mathematics of soap films by lecturing on this subject.

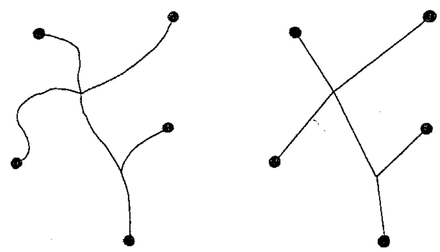
1. Introduction

One of the mathematical results that one encounters early in life is that concerning the shortest path joining two points. It is, of course, the straight line. Any other path will have a greater length. If we attempt to generalize this result, in order to determine the minimum path connecting three, or four, or more points, the problem becomes rapidly more complicated. These problems, however, are important in such applications as the construction of roadways connecting a number of towns, and pipelines or cables joining a number of centres. In all these cases the cost of construction is proportional to the length, so minimizing the length also minimizes the cost.

Let us restrict the discussion to the construction of roadways. Can we derive any properties of minimum roadway configurations, linking a number of towns, before attempting the more difficult task of finding complete solutions? If we consider a number of towns connected by some roadway system (see Figure 1(a)) it is clear that the minimum roadway configuration cannot have any curved roads, as any curved section of road can be replaced by a shorter straight-line length of road. When all the curved roads are replaced by straight-line roads, we obtain the configuration shown in Figure 1(b).

2. Three-town problems

In order to gain further insight into the general problem, let us examine the three-town problems. The simplest problem would seem to occur when all the towns, A, B and C, say, are arranged at the vertices of an equilateral triangle (see Figure 2) with sides of length d. The minimum roadway system might be along two



(a) (b)

Figure 1. Roadway configurations; the straight-line road system has the shorter length.

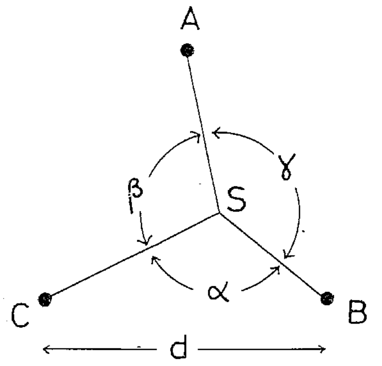


Figure 2. Three towns arranged at the vertices of an equilateral triangle.

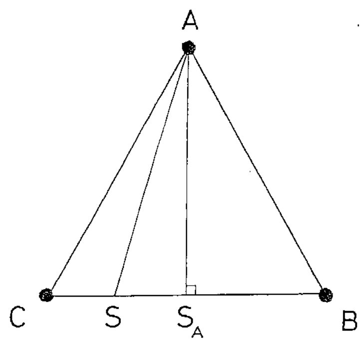


Figure 3. A roadway joining A, B and C with S along BC.

sides—with $L = 2d$—or along three straight-line roads meeting at a point S with angles between the roads of α, β, γ, as indicated in Figure 2—or one side and one road joining the opposite vertex to that side, as in Figure 3. The minimum length in this case is clearly when S is at S_A, the midpoint of CB. This gives $L = (1 + \frac{1}{2}\sqrt{3})d = (1{\cdot}866\ldots)d$, which is clearly shorter than the two-side solution with $L = 2d$.

We might now look at a three-road system with S on AS_A, as indicated in Figure 4. If $SS_A = x$ then the total length L of the roadway is given by

$$L = AS + 2CS, \tag{1}$$

$$= (\tfrac{1}{2}d\sqrt{3} - x) + 2\sqrt{x^2 + \tfrac{1}{4}d^2}. \tag{2}$$

This will be minimized when $dL/dx = 0$, that is, when

$$0 = -1 + 2x(x^2 + \tfrac{1}{4}d^2)^{-1/2}, \tag{3}$$

or

$$x = \tfrac{1}{6}\sqrt{3}d. \tag{4}$$

Consequently $\alpha = 120°$, and, by symmetry, $\gamma = \beta = 120°$. The value of this minimum is, from (2), $\sqrt{3}d$.

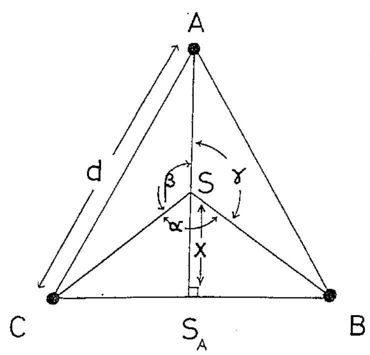


Figure 4. A roadway configuration with S along AS_A.

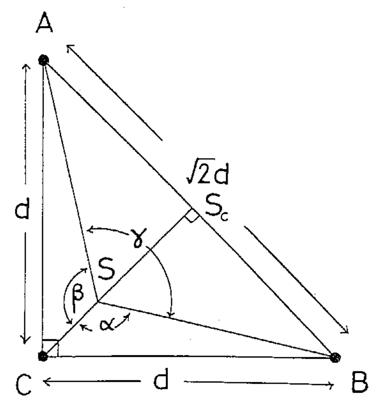


Figure 5. Three towns arranged at the vertices of an isosceles right-angled triangle.

This argument does not prove that this is *the* minimum roadway, but indicates that this may well be the case. It is certainly shorter than any system with S on the boundary or at any other point on the bisector of an angle of the triangle. Further analysis will be needed to prove that it is *the* true minimum.

Before attempting a general proof for the three-town problems, let us examine a further example. Consider the towns at the corners of an isosceles right-angled triangle with sides of length d and $\sqrt{2}d$, as in Figure 5. Similar arguments to those presented for the equilateral triangle would indicate that the position of S might be along the bisector CS_C of $\hat{C}$ where S_C is the midpoint of AB. We can obtain a general expression for the total length $L(x)$ of the roadway for S along CS_C as a function of $x = SS_C$, and minimize L with respect to x. Again we obtain $\gamma = \alpha = \beta = 120°$, the total length of roadway being $\frac{1}{2}(\sqrt{6} + \sqrt{2})$.

At this stage we might conjecture that, for any arrangement of three towns, the minimum path has this 120° property. We shall now show that this is indeed the case, using a geometrical proof that was brought to my attention by Mr R. D. Nelson of Ampleforth College, York (see reference 2).

Consider any triangle ABC with a 'roadway' system of length L formed by three lines meeting at S, as in Figure 6. Now rotate the shaded triangle anticlockwise about C through 60°, so that A is now at A' and S is now at S'. Then $S'C = SC$ and $S'CS = 60°$, so triangle CSS' is an equilateral triangle. Also,

$$L = AS + BS + CS = A'S' + SS' + SB \qquad (5)$$

as, by definition, $A'S' = AS$ and $S'S = SC$, $CS'S$ being an equilateral triangle. Now A' and B are fixed points. As S is varied,

$$L = A'S' + S'S + SB$$

will be minimized when $A'B$ is a straight line (see Figure 7). In Figure 7, $\alpha = \widehat{CSB}$ and $S'SB$ is a straight line, so

$$\alpha = 180° - 60° = 120°. \qquad (6)$$

Also, using the notation in Figure 7,

$$\begin{aligned} \beta &= \beta', && \text{by definition,} \\ &= 180° - 60°, && \text{as } A'B \text{ is a straight line,} \\ &= 120°. && \end{aligned} \qquad (7)$$

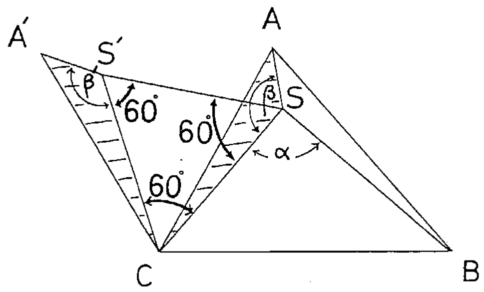


Figure 6. CAS rotated through 60° about C.

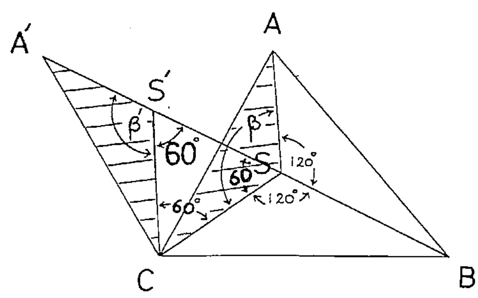


Figure 7. The minimum-path position for S.

Hence, from (6) and (7),

$$\gamma = 360° - \alpha - \beta = 120°. \tag{8}$$

Thus the $120°$ angles are a general property of the minimum path connecting A, B and C.

These results can be demonstrated experimentally by drilling three small holes in a horizontal wooden table. They represent the three points A, B and C. Three equal weights mg are now hung below the holes from strings that are tied together on the table, the height of the weights above the ground being h_1, h_2 and h_3. The three weights will come to equilibrium when their total potential energy is minimized. This will occur when $mg(h_1 + h_2 + h_3)$ is minimized, that is, when $(h_1 + h_2 + h_3)$ is minimized. This will clearly occur when the sum of the three horizontal lengths of string on the table is minimized. So the final equilibrium configuration of the horizontal strings is one in which their total length is minimized. Also, as the tension in each string is mg, the equilibrium configuration occurs when pairs of adjacent strings intersect at $120°$.

These proofs, however, are only valid provided that the angles of the triangle are less than or equal to $120°$, when S lies inside or on the triangle. For triangles with an angle equal to $120°$, the minimum path length is the sum of the two shortest sides of the triangle, i.e. those adjacent to the largest angle. This latter result can also be proved to hold for triangles with an angle greater than $120°$. Summarizing, a triangle with no angle greater than $120°$ has a minimum path formed by three lines meeting at a point S inside the triangle at $120°$. All other triangles have a minimum path formed by the two sides adjacent to the largest angle of the triangle.

3. Many-town problems

We can show, using the result for the minimum roadway joining three towns, that the general minimum roadway connecting a number of towns can only have roadway intersections with three roads, each road making an angle of $120°$ with the adjacent roads.

If the general minimum roadway consisted of a roadway intersection with more than three roads, as in Figure 8(a), one could, using the results for three points, find a roadway configuration with a smaller length. For example, in Figure 8(b), using points P and Q in triangle POQ, with $\widehat{POQ} < 120°$, we can form a roadway configuration with a smaller length by replacing the full lines PO and QO by the

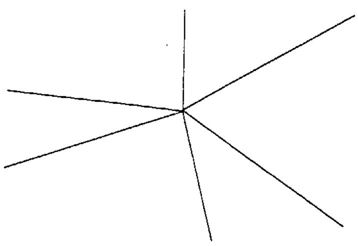


Figure 8(a). Intersection point with more than three roads.

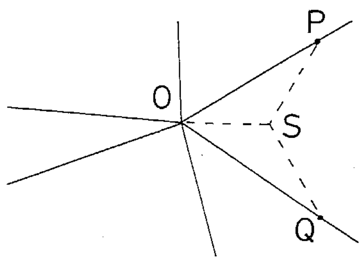


Figure 8(b). A roadway with a shorter length than that in Figure 8(a).

Figure 9. Jacob Steiner (1796–1863).

broken lines, PS, SO and SQ, meeting at 120°. Repeating this procedure with all the roads meeting at O, we finally obtain only three roads meeting with angles of 120°.

These minimum-path, or roadway, problems were investigated during the nineteenth century by the Swiss mathematician Jacob Steiner (Figure 9). The points at the three-way intersection are often known as *Steiner points*.

Let us examine the four-town problem with the four towns arranged in a square array with sides of length d (Figure 10(a)). On the grounds of symmetry, we might be tempted to guess that the minimum roadway has an X-configuration, with length $2\sqrt{2}d$. However, we know that the roadway intersection must consist of three roads meeting at 120°. Consequently, the only possible solutions are the two shown in Figure 10(b) by the full and broken lines with angles 120°. These give roadways of length $(1+\sqrt{3})d = (2{\cdot}73\,..)d$, which is shorter than the X-configuration of length $2\sqrt{2}d = (2{\cdot}82\ldots.)d$. It is of interest to investigate how the minimum roadway configurations alter when the towns are arranged in a rectangular array with $AD = d$ and $AB = (1+\omega)d$. If ω is increased from 0 the two configurations in Figure 11 will have different lengths, one being a local minimum and the other an absolute

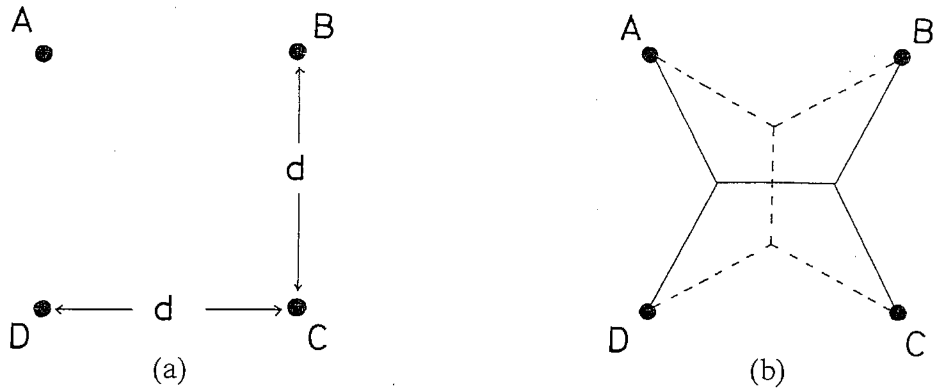


Figure 10. (a) The square array of four towns. (b) The minimum roadway configurations.

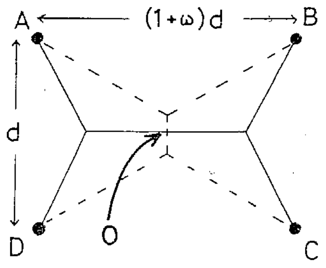


Figure 11. The two minimum roadways for a rectangular array of towns.

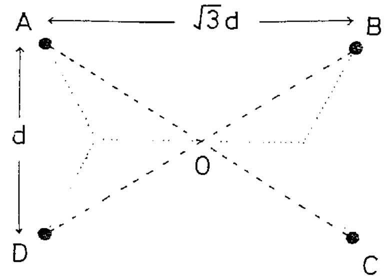


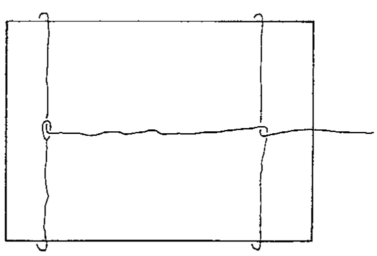
Figure 12. The configuration with the critical value of $\omega = (\sqrt{3} - 1)d$.

minimum. The length of the full road is $(1 + \sqrt{3} + \omega)d$, and that of the broken road $(1 + \sqrt{3}(1 + \omega))d$. Consequently the broken road is the local minimum. As ω is increased, the two three-way broken intersections will eventually meet, leading to a four-way intersection at O in Figure 11. When this occurs $\omega = \sqrt{3} - 1$, and the configuration will no longer be a minimum configuration as we can form a roadway with a shorter length by replacing AO and DO by three roads meeting at 120° as indicated by the dotted lines in Figure 12. This configuration is the same as the absolute minimum. So, for $\omega \geq (\sqrt{3} - 1)$ there is only one minimum roadway. A similar situation arises when $\omega < 0$, the rectangular array of towns being such that $AB < AD$. As ω is reduced, the full configuration in Figure 10(b) is longer than the broken configuration. When the point is reached where the full configuration consists of two coincident three-way intersections, the configuration will no longer be a minimum configuration and the only minimum configuration will be the broken configuration. This will occur when $\omega = -(1 - (\sqrt{3}/3))$.

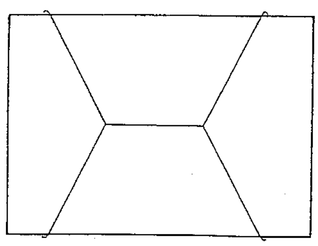
4. Analogue methods and soap films

It is perhaps unexpected to find that the square array of towns has a minimum roadway configuration that does not have the full symmetry of the square. However, it should not be thought so unusual, as we regularly encounter the solution to this problem in another context.

We solve this problem every time we tie up a parcel with string. Usually we start by wrapping the string around the parcel with the configuration shown in Figure 13(a). When we pull on the string, in order to minimize the length of string being



(a)



(b)

Figure 13. Tying up a parcel: (a) initial configuration, string loose (b) string taut.

used to tie up the parcel securely, we produce a constant tension in the string. This leads to the configuration in Figure 13(b). The three-way string-intersection has three equal forces, produced by the constant tension in the string, in equilibrium, so three strings must meet at 120°. Thus minimizing the length of string leads to all the features associated with minimum roadway problems.

Another analogous system, which has the advantage that it can easily be applied to any number of points, is that based on the properties of soap films (see references 1 and 3). A soap film has the property that its energy is proportional to its area. When a soap film comes to equilibrium, it will take up a configuration that minimizes its energy and hence its area. For example, a soap film contained by a circular ring will not bulge out at equilibrium but will form the minimum-area surface, the disc contained by the circular ring.

In order to make use of this minimum-area property to solve minimum-path problems, we must convert the minimum-area property into a minimum-path property. To do this, let us first focus attention on the simplest problem: the minimum path connecting two points. Consider a soap film contained between two parallel clear perspex plates, with two pins perpendicular to the plates, separating the plates, and at some distance from each other. Then, by symmetry, the film will be perpendicular to the plates and bounded by the two plates, beginning on one pin and ending on the other (Figure 14). Consequently the film will be in the form of a tape, with constant width equal to the distance between the plates. The area of the soap film is proportional to its length. When it comes to equilibrium it will have minimum area and also minimum length. Thus the tape will end up, in equilibrium, with the straight-line configuration in Figure 14. The analogous solution can, by the same reasoning, be extended to any arrangements of pins or points. The soap film solution to the four-town problem is shown in Figure 15.

We have seen that the four-town problem can have two minimum configurations. In order to determine the minimum roadway with the smallest length we need to calculate the length of each path and determine the one with the smallest length using the 120° property. In problems with many towns there may be many minimum configurations. This analogue method is based on producing soap films between the plates by dipping the plates into a bath of soap solution at different

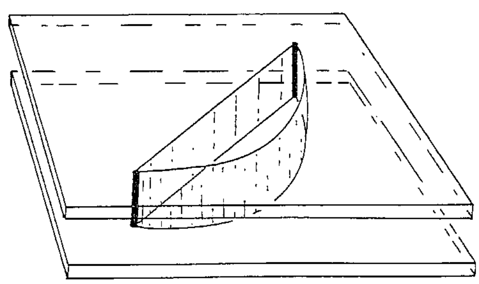


Figure 14. Soap films bounded by parallel plates and two pins. The curved surface is a non-equilibrium soap film and the straight surface is the minimum-area, minimum-path, surface.

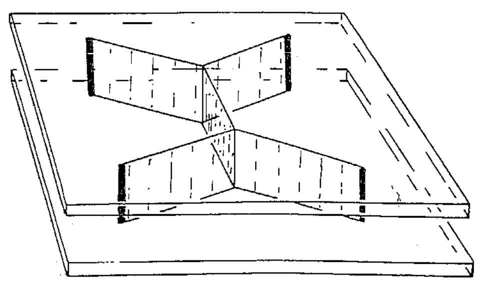


Figure 15. The minimum-path soap film joining four pins arranged in a square array.

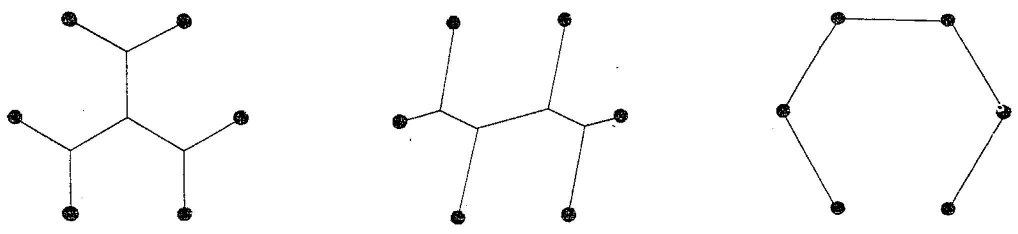


Figure 16. Minimum configurations for six towns arranged in a regular hexagon.

angles or by perturbing an equilibrium soap film by blowing it into another equilibrium configuration. There is no simple method of determining analytically all the minimum configurations (see reference 4).

An interesting example of a problem with three minimum configurations occurs when solving the minimum path joining six towns arranged at the vertices of a regular hexagon. The three configurations, which can easily be obtained using soap films, are shown in Figure 16.

If the sides of the hexagon have unit length, the lengths of the minima can be calculated, using the $120°$ property, to be $3\sqrt{3}$, $2\sqrt{7}$ and 5 respectively. It is interesting to note that the configurations have, 3, 2 and 1-fold symmetry about the axis of symmetry perpendicular to the plane of the hexagon. The configuration of smallest length in this case is 16(c). However, in other problems it may well be one of the internal roadways.

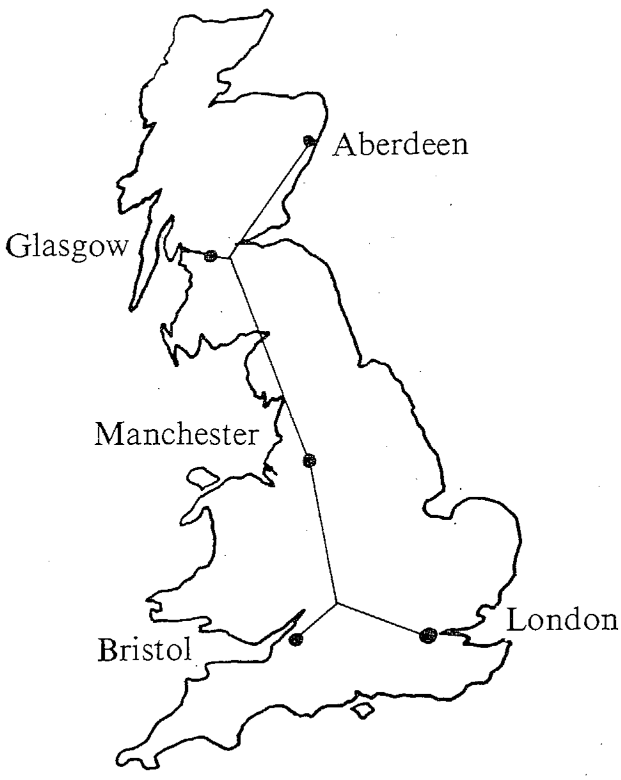


Figure 17. The minimum roadway system linking London, Bristol, Manchester, Glasgow and Aberdeen.

Let us now apply this analogue method to the practical problem of linking London, Bristol, Manchester, Glasgow and Aberdeen by the shortest length of roadway. It is necessary to draw a map of Britain on one of the parallel perspex plates and insert pins perpendicular to the plates at these towns. After dipping the plates into soap solution we obtain the roadway configuration shown in Figure 17, with one Steiner point to the east of Glasgow and one south of Birmingham. In this application, there is only one minimum roadway configuration.

References

1. R. Courant and H. Robbins, *What is Mathematics?* (Oxford University Press, New York, 1941).
2. H. S. M. Coxeter, *Introduction to Geometry* (Wiley, New York, 1963).
3. E. N. Gilbert and H. O. Pollak, Steiner minimal trees, *SIAM J. Appl. Math.* **16** (1968), 1.
4. C. Isenberg, *The Science of Soap Films and Soap Bubbles* (Tieto, Clevedon, 1978).

Is Complex Analysis Useful?

I. D. MACDONALD, *University of Stirling*

The author graduated from the University of Aberdeen. After taking a Ph.D. at Manchester he was an Assistant Lecturer in Sheffield. Then he had a spell in Australia, where he was Foundation Professor of Mathematics in the University of Newcastle. At the moment he is Senior Lecturer in the University of Stirling. His research interest is the theory of groups.

Because great stress is placed on applications of knowledge in these parsimonious times, we must consider the uses and applications of complex analysis in the real world. (By complex analysis we mean firstly properties of the complex number system, and second and more particularly functions of a complex variable.) Uses are not quite the same thing as usefulness in the real world. The uses of complex analysis are subtle and sophisticated. One must not expect it to have the immediacy of application that arithmetic, say, has in the market place. Indeed, complex analysis is used in other areas of mathematics and in physics, and these other areas sometimes have applications to the real world. In any event, usefulness is a powerful source of motivation.

We shall freely quote theorems and results of complex analysis without proof. Indeed this note may be regarded as a summary of a whole course in complex analysis.

1. Simplification of elementary algebra

(i) The quadratic equation

$$ax^2 + bx + c = 0,$$

where a, b, c are real numbers and $a \neq 0$, always has two complex roots. Exactly the same statement can be made if a, b, c are allowed to be complex numbers. This is a drastic simplification of statements like: 'a quadratic equation sometimes has two roots and sometimes it has none, and you can distinguish the two cases as follows . . .'.

It may not be surprising that the equation with real coefficients has two complex roots. But note carefully that we have no need of any system of 'super-complex' numbers, for instance the quaternions, to say that the equation with complex coefficients has two roots.

Similarly, complex cubic equations always have three roots, no matter whether the coefficients are real or complex. (We are of course assuming use of the convention that multiple roots are to be counted the correct number of times, here and throughout our discussion of equations.)

(ii) These are special cases of a general result, the so-called 'fundamental theorem of algebra', which states that the equation

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = 0$$

has precisely n complex roots (with convention about multiplicity). Here $a_0 \neq 0$, and $a_0, a_1, \ldots, a_n$ are complex numbers.

For $n \geq 5$ neither the fundamental theorem nor the special case that the above equation with a_0, $a_1, \ldots, a_n$ real has n complex roots can be proved by purely algebraic methods. Techniques from complex analysis, contour integration or something essentially equivalent to it, have to be employed in every known proof. Again complex analysis is seen to simplify elementary algebra.

2. Simplification of elementary (real) function theory

(i) The real function $(1 + x)^{-1}$ has a power series; this series converges when $|x| < 1$ and diverges when $|x| \geq 1$; the function is of course not defined at $x = -1$, that is at one number whose modulus is 1.

Compare and contrast this with the function $(1 + x^2)^{-1}$, which exists for all real numbers x. The function has a power series expansion which converges if and only if $|x| < 1$. There may come a time when we feel that we need to 'explain' divergence in terms of the roots of the equation $x^2 = -1$. (The theory of differential equations offers similar scope for 'explanations' of divergence.)

(ii) Consider the power series for the real functions sin, cos, exp, and so on. The series for sin, found by Euler about 1730, looks like this:

$$\sin x = x - x^3/3! + x^5/5! - \cdots.$$

Just let us suppose, without any justification at all, that we could replace x in that series by ix, where x is real and $i^2 = -1$:

$$\sin ix = ix + ix^3/3! + ix^5/5! + \cdots.$$

The series for e^x with x real is

$$e^x = 1 + x + x^2/2! + x^3/3! + \cdots,$$

and, if we make the definition

$$\sinh x = (e^x - e^{-x})/2,$$

then we can develop series for $\sinh x$ (and $\cosh x$) and we seem to find that, if x is real, then $\sin ix = i\sinh x$. A parallel argument gives $\cos ix = \cosh x$. And on similar assumptions $\sin x$ and $\cos x$ can be expressed in terms of $\sinh ix$ and $\cosh ix$.

The conclusion is that, if the assumptions that we have made about manipulations involving i could be validated, then there would be a single theory of trigonometric and hyperbolic functions, certainly for real x and maybe for complex x too if that could be made to mean anything; such a theory would be a branch of the theory of the exponential function. This would be a considerable unification of elementary function theory.

Let us develop this a little further. An argument which is the same in spirit as those above soon gives the formula $e^{ix} = \cos x + i\sin x$, (of which we note the special case $e^{i\pi} = -1$), from which we also have $e^{-ix} = \cos x - i\sin x$. Thus

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}, \qquad \sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

Let us further suppose, with no more justification than above, that if φ, ψ are real numbers then $e^{i\varphi}e^{i\psi} = e^{i\varphi + i\psi}$. Then

$$\begin{aligned}\sin(\varphi + \psi) &= (e^{i(\varphi+\psi)} - e^{-i(\varphi+\psi)})/2i \\ &= (2e^{i\varphi}e^{i\psi} - 2e^{-i\varphi}e^{-i\psi})/4i \\ &= ((e^{i\varphi} - e^{-i\varphi})(e^{i\psi} + e^{-i\psi}) + (e^{i\varphi} + e^{-i\varphi})(e^{i\psi} - e^{-i\psi}))/4i \qquad (*) \\ &= \sin\varphi\cos\psi + \cos\varphi\sin\psi,\end{aligned}$$

where at (*) we use the purely algebraic result

$$2(ac - bd) = (a-b)(c+d) + (a+b)(c-d)$$

if a, b, c, d are complex numbers, which is readily deducible from the laws of algebra. Therefore the usual formula for $\sin(\varphi + \psi)$ is deduced from the assumed property of e^{ix}; similarly for $\cos(\varphi + \psi)$, $\sinh(\varphi + \psi)$, etc.

(iii) Next we look at integration (of real functions), still under the assumption that i can be manipulated like a real number. Put

$$I_1 = \int e^x \cos x\, dx, \qquad I_2 = \int e^x \sin x\, dx.$$

Then

$$
\begin{aligned}
I_1 + iI_2 &= \int e^{(1+i)x}dx \\
&= \frac{1}{1+i}e^{(1+i)x} \\
&= \frac{1}{1-i^2}(1-i)e^x(\cos x + i\sin x) \\
&= \tfrac{1}{2}e^x((\cos x + \sin x) + i(-\cos x + \sin x)),
\end{aligned}
$$

where some of the steps are totally lacking in justification, and 'therefore' (taking real and imaginary parts)

$$
I_1 = \tfrac{1}{2}e^x(\cos x + \sin x), \qquad I_2 = \tfrac{1}{2}e^x(-\cos x + \sin x).
$$

At this stage we might ask ourselves whether these results are correct—would we get the same answers if we integrated I_1 and I_2 by parts? More significantly, *should* they be correct?

(iv) Here is a definite integration performed in the same spirit. We start with the result

$$
\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi},
$$

which we assume to be known. Consider

$$
I = \int_{-\infty}^{\infty} e^{-x^2}\cos 2ax\,dx,
$$

where a is real. Since

$$
\begin{aligned}
e^{-x^2}\cos 2ax &= \tfrac{1}{2}(e^{-x^2+2iax} + e^{-x^2-2iax}) \\
&= \tfrac{1}{2}e^{-a^2}(e^{-(x-ia)^2} + e^{-(x+ia)^2}),
\end{aligned}
$$

we write down

$$
I = \tfrac{1}{2}e^{-a^2}\left(\int_{-\infty}^{\infty} e^{-(x-ia)^2}dx + \int_{-\infty}^{\infty} e^{-(x+ia)^2}dx\right).
$$

At this point we feel it would be nice to substitute $t = x - ia$ in the first integral and $t = x + ia$ in the second; after all, the limits of integration are infinite and surely the factors $\pm ia$ will make no difference at infinity. If this is agreed to, then we have

$$
\begin{aligned}
I &= \tfrac{1}{2}e^{-a^2}\left(\int_{-\infty}^{\infty} e^{-t^2}dt + \int_{-\infty}^{\infty} e^{-t^2}dt\right). \\
&= e^{-a^2}\sqrt{\pi}.
\end{aligned}
$$

Let us summarise our treatment of function theory. Considerable justification of manipulation of i and of complex numbers in general is necessary in this context. The situation is quite different from the algebraic use of i in the same way as a real number, which we justify by appeal to axioms. Here we are using i in operations with the limiting processes of analysis such as summing infinite series, differentiating, integrating, and, far from being justified, such use of i does not even have a meaning at this stage. However, it can be given both a meaning and a justification—that is what the theory of complex functions is all about—and in fact all the 'results' we have 'derived' above will then be true. At the moment the rigorous context is totally lacking.

3. Evaluation of (real) definite integrals

Complex function theory is decidely useful in evaluating definite integrals in real function theory. No doubt such integrals could be evaluated by real methods, but in many cases the complex method is appreciably easier. It is also to some extent systematic. We quote some examples from the book *A First Course on Complex Functions* by G. J. O. Jameson (Chapman and Hall, London, 1970).

$$\text{(p. 119)} \qquad \int_{-\infty}^{\infty} \frac{x \sin x}{x^2 + a^2}\, dx = \frac{\pi}{2} e^{-a}, \qquad (a > 0).$$

$$\text{(p. 122)} \qquad \int_{-\infty}^{\infty} (e^{ax}/(1 + e^{x}))dx = \pi/\sin a\pi: \qquad (0 < a < 1).$$

$$\text{(p. 123)} \qquad \int_{0}^{2\pi} \frac{dt}{a + \cos t} = 2\pi/\sqrt{(a^2 - 1)}, \qquad (a > 1).$$

4. Applications outside pure mathematics

The chief applications outside pure mathematics are to physics and, since physics is the most mathematically sophisticated of the natural sciences, this may not be unexpected. Physicists find complex functions useful in the theory of electricity. What is basically needed is the material about the exponential and trigonometric functions outlined above, and it is used in a rather formal way. There is a deeper application in the theory of fluid flow. This arises from transformations of the plane and the use of complex functions in that geometrical sense. By applying a suitable transformation, the fluid flow round some obstruction may be replaced by the flow round some other object in a way that is easier to treat; for instance, the air flow across an aeroplane wing may be replaced by the air flow round a cylinder. These are two-dimensional problems, and there do not seem to be any satisfactory analogous methods for three-dimensional problems.

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

15.4. (Submitted by Paul Brennan, London SE2) Given positive real numbers a, b and positive integers m, n with $m > n$, prove that

$$(a^m + b^m)^n < (a^n + b^n)^m.$$

15.5. (From the 1981 Hungarian Olympiad) Determine all real numbers α which satisfy the double inequality

$$\tfrac{1}{3}\tan\alpha \leq \tan 3\alpha \leq 3\tan\alpha.$$

15.6. (From *le petit Archimède*) What is the integer part of the number

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{10{,}000}}?$$

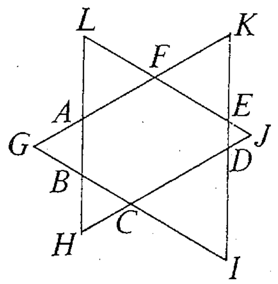
Solutions to Problems in Volume 14, Number 3

14.7. The angles of the convex hexagon $ABCDEF$ are equal. Prove that

$$AB - DE = EF - BC = CD - FA.$$

Solution

It is perhaps a sad commentary on the state of elementary geometry today that we received no solutions to this easy problem. The internal angles of the hexagon $ABCDEF$ will



all be 120°. We produce the sides of the hexagon as shown. Then the triangles on the sides of the hexagon are equilateral, and opposite sides of it are parallel. Hence

$$AB + BC = AB + BH = KE + ED = FE + ED,$$

so that

$$AB - DE = EF - BC.$$

Also,

$$BC + CD = BC + CI = LF + FE = AF + FE,$$

so that

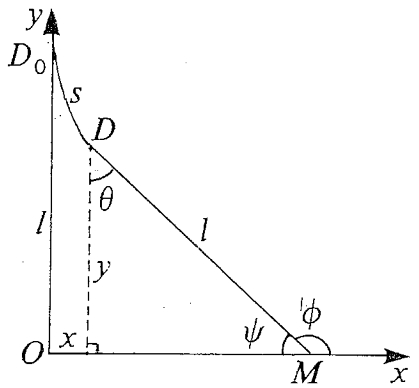
$$EF - BC = CD - FA,$$

and the problem is solved.

14.8. A man takes his obstinate dog for a walk. The man walks in a straight line, and the lead is always taut. Initially, the lead is at right angles to the direction in which the man walks. How far has the dog walked when the man has walked a distance d?

Solution

The problem was an extension of Problem 14.4. Rather than using the result of that problem, we shall solve the present problem from scratch. The man starts at the origin O and walks in the direction $\overrightarrow{Ox}$. The dog starts at the point $D_0(0, l)$, so that the lead is of length l.



Suppose that, at some stage in the walk, the man has reached point M and the dog point $D(x, y)$, as shown in the diagram. Then

$$y = l\cos\theta.$$

Now

$$\frac{dy}{dx} = \tan\phi = \tan(\pi - \psi) = \tan\left(\frac{\pi}{2} + \theta\right) = -\cot\theta,$$

so

$$\frac{dx}{d\theta} = \frac{dx}{dy}\frac{dy}{d\theta} = -\tan\theta(-l\sin\theta) = \frac{l\sin^2\theta}{\cos\theta}.$$

If s is the distance walked by the dog to the point D, then

$$\begin{aligned} s &= \int_0^\theta \sqrt{\{(dx)^2 + (dy)^2\}} \\ &= l\int_0^\theta \sqrt{\left\{\frac{\sin^4\theta}{\cos^2\theta} + \sin^2\theta\right\}}\, d\theta \\ &= l\int_0^\theta \tan\theta\, d\theta \\ &= l[-\log(\cos\theta)]_0^\theta \\ &= -l\log\cos\theta. \end{aligned}$$

Also,

$$x=\int_0^\theta \frac{l\sin^2\theta}{\cos\theta}\,d\theta$$

$$=l\int_0^\theta (\sec\theta-\cos\theta)\,d\theta$$

$$=l\left[\log\left(\tan\left(\frac{\pi}{4}+\frac{\theta}{2}\right)\right)-\sin\theta\right]_0^\theta$$

$$=l\left(\log\left(\tan\left(\frac{\pi}{4}+\frac{\theta}{2}\right)\right)-\sin\theta\right).$$

Now

$$d=x+l\sin\theta,$$

so that

$$d=l\log\left(\tan\left(\frac{\pi}{4}+\frac{\theta}{2}\right)\right)$$

and

$$\tan\left(\frac{\pi}{4}+\frac{\theta}{2}\right)=e^{d/l}.$$

Also,

$$\cot\left(\frac{\pi}{4}+\frac{\theta}{2}\right)=e^{-d/l},$$

so that

$$\frac{1}{\sin\left(\frac{\pi}{4}+\frac{\theta}{2}\right)\cos\left(\frac{\pi}{4}+\frac{\theta}{2}\right)}=e^{d/l}+e^{-d/l}.$$

Thus

$$\cosh\frac{d}{l}=\frac{1}{\sin\left(\frac{\pi}{2}+\theta\right)}=\frac{1}{\cos\theta}.$$

Hence

$$s=-l\log\cos\theta=l\log\left(\cosh\frac{d}{l}\right),$$

which is the distance walked by the dog when the man has walked a distance d. One wonders whether the dog is aware of this fact!

The only solution we received to this problem was from Lloyd Taylor of the University of Nottingham, who submitted the original problem 14.4.

14.9. After the second ballot for the election of the Deputy Leader of the Labour Party, it was announced that Mr Benn had received from members of the Parliamentary Party 10·241 % of the electoral college vote. A television commentator said that it would be interesting to know how many MPs had actually voted for Mr Benn, but that this would not be known until the next day. Is this correct? The total electoral vote of the Parliamentary Labour Party was 30 % divided in proportion to those who voted for each candidate, and there were, at that time, 254 Labour MPs, not all of whom voted. How many MPs voted for Mr Benn if we assume that the figure of 10·241 was correct to $\pm 0{\cdot}0005$?

When the actual votes were disclosed later, it was found that the figure of 10·241 had been miscalculated; it should have been 10·240. Can you find the actual figures from this information?

Solution

Since no readers ventured into this political minefield so far as to submit a solution, we must supply the solution given by Dr D. J. Roaf, who submitted the problem.

Let B be the total number of votes cast for Mr Benn and let N be the total vote. Then

$$\frac{10{\cdot}2405}{30} \leq \frac{B}{N} \leq \frac{10{\cdot}2415}{30}$$

so that

$$0{\cdot}02405 \leq \frac{3B-N}{N} \leq 0{\cdot}02415,$$

$$41{\cdot}5801 > \frac{N}{3B-N} > 41{\cdot}4078.$$

We now have the following possibilities:

$$\begin{array}{ll}
3B-N=1, & 41{\cdot}5801 > N > 41{\cdot}4078,\ N \text{ not an integer} \\
3B-N=2, & 83{\cdot}1602 > N > 82{\cdot}8156,\ N=83 \\
3B-N=3, & 124{\cdot}7403 > N > 124{\cdot}2234,\ N \text{ not an integer} \\
3B-N=4, & 166{\cdot}3204 > N > 165{\cdot}6312,\ N=166 \\
3B-N=5, & 207{\cdot}9005 > N > 207{\cdot}0390,\ N \text{ not an integer} \\
3B-N=6, & 249{\cdot}4806 > N > 248{\cdot}4468,\ N=249 \\
3B-N=7, & \qquad\qquad\qquad\qquad N > 280.
\end{array}$$

But, if $N=83$, then $3B=N+2=85$, and B is not an integer. If $N=166$, then $3B=N+4=170$, and again B is not an integer. If $N=249$, then $3B=255$ and $B=85$. Therefore $B=85$ and $N=249$.

With 10·240 instead of 10·241, we have

$$\frac{10{\cdot}2395}{30} \leq \frac{B}{N} \leq \frac{10{\cdot}2405}{30},$$

so

$$41{\cdot}7537 > \frac{N}{3B-N} > 41{\cdot}5800,$$

which gives $B=57$ and $N=167$ or $B=71$ and $N=208$. The latter was the actual vote.

Book Reviews

The Creative Use of Calculators. By J. P. KILLINGBECK. Penguin Books Ltd, London, 1981. Pp. 220. £1·95.

Dr Killingbeck divides his exploration of some of the uses of calculators into four main areas. Firstly, he examines everyday calculations that can be carried out on a simple calculator such as simple and compound interest, mortgage calculations, investments and so on. Secondly, he concentrates his attention on the application of calculators to slightly more formal work in mathematics and science, examining problems such as the use of the binomial theorem for approximations, iterative procedures for extending the number of decimal places in an answer, complex numbers, differentiation, and several topics in physics. Dr Killingbeck devotes quite some effort to demonstrating how simple calculators can be used to calculate functions which do not appear on their keyboards, and takes pains to develop the mathematical explanations behind the techniques employed. Thirdly, he examines the uses of scientific calculators, looking at trigonometric functions, power series and applications of exponential and logarithmic functions in mathematics and science. In this section, he covers quite a lot of interesting mathematical material without clearly explaining the relevance of the calculator to its study. The final section of the book is a chapter on programmable calculators which gives some hints on what the programmable machine should ideally be able to perform, and discusses some of the suitable applications such as numerical integration and iterative methods for solution of equations. It is a pity in many ways that Dr Killingbeck does not look into some of these ideas in connection with his section on scientific calculators, because a great deal of profitable work on these topics can be performed without a programmable calculator.

Apart from these four categories dealing with the actual applications of calculators, Dr Killingbeck includes an introductory chapter on the general design features of calculators, where he discusses the different kinds of logic (simple algebraic logic, algebraic operating system, reverse Polish logic) that need to be used for different makes of calculator. This section also examines the facilities available on simple calculators and compares the relative advantages and disadvantages of different kinds of keyboard, power supply and display.

There is also a brief but interesting chapter on how the calculator works, describing the construction of the silicon chip, binary arithmetic and how it is employed in the calculator and other general operating features. This short section, though necessarily somewhat superficial, provides the reader with some idea of what is 'going on' in a calculator when he uses it. The link demonstrated between the mathematics and the physical operations of the machine may spur the enthusiastic reader to delve deeper into the subject elsewhere.

Exercises are provided at the end of each chapter on the material covered, and solutions appear at the end of each question for readers to consult if they do not wish to attempt the problems. As Dr Killingbeck points out, it is never easy to read mathematics without attempting to 'do' it, and he quite rightly recommends that the reader should work his way through the book, and through the exercises in particular, with his calculator to hand. Indeed, there are so many topics touched upon in the book that the significance of the material and the methods used might be lost to the reader who does not try the problems on his own machine. The book also contains appendixes of formulae which may be helpful to someone without a scientific calculator, and a short list of books recommended for further reading, together with the author's comments on each.

In general, the book is fairly readable, but because of the nature of the subject and of some of the material covered, it requires some concentration and willingness to work through the problems presented. The earlier chapters are easy to follow and Dr Killingbeck has taken care to describe in detail calculator procedures (even including key sequences) and the logic behind them. However, as the book proceeds and the ideas and techniques become more demanding, a reader without some mathematical training might find the work less comprehensible.

Whilst there is material in the book that would be of interest to any calculator owner, the majority of the work is aimed at readers who have a reasonable mathematical background and who perhaps have already studied some of the material in a different context or without reference to calculators. As such, the book would make good background reading for students engaged in the study of mathematics or science to A-level or equivalent, particularly if calculators are not being used 'creatively' in their courses. Similarly, it would be of some benefit to teachers who are looking for ideas to improve or incorporate the use of calculators in their reading.

Carmel R.C. School, Darlington J. QUINN

Great Moments in Mathematics (Before 1650). By HOWARD EVES. The Mathematical Association of America (distributed by John Wiley and Sons Ltd), 1980. Pp. xiv + 270. £14·00.

In Britain mathematics is a serious business; while we frequently find courses on the appreciation of painting or music being run in our schools and colleges, I have never heard of one on the appreciation of mathematics. One reason is because the study of mathematics, even at a superficial level, demands a technical knowledge of the subject. Usually those who become interested in the history of mathematics already have a considerable knowledge of this. That this is not necessary is demonstrated by the contents of this book, which demands little knowledge over O level, plus a keen intellect and an enquiring mind. This book grew out of a course of lectures given to American college students, who then had little acquaintance with advanced mathematics. By the author's own account, it is a very pale copy of what must have been a series of exciting talks.

Twenty great moments are recounted; the ideas which they embrace have occurred during anything from a flash of inspiration to a few centuries of evolution and gestation. This is the first of two volumes, and it deals with mathematics discovered before the advent of the calculus of Newton and Leibniz. It begins with the discovery of counting and continues until the construction of analytical geometry by Descartes and Fermat. In between these events, Pythagoras' theorem, irrational numbers, Euclid's 'Elements', the work of Archimedes, Napier's logarithms, Kepler's laws and many other important discoveries are outlined and evaluated.

This book is essentially a discussion of topics each of which is worthy of the description 'great moment in mathematics'. This element of showmanship can be forgiven as the chapters are authoritative, informative and interesting to the extent that they will absorb readers with widely different levels of understanding. The topical approach will not suit all teachers and students of mathematical history; indeed this is not a book from which to acquire a systematic knowledge of this subject, but it does facilitate the development and comprehension of certain areas through problems which are often challenging, but which require only relatively elementary methods for their solution. Thirty pages are given to discussing the solutions of the exercises. There is an extensive, and therefore useful, index. At the end of each chapter there is a list of further reading; most of the items mentioned are well-known works which can be found in college and university libraries.

Although this book is typically American in conception, and it has a very definite niche in that system of mathematical education, it will be appreciated by enquiring students of mathematics whose technical expertise is as yet restricted, such as sixth formers and even more advanced students. Their teachers may also find this a useful source of material, as might also intending teachers, and all interested in the history of mathematics, from the interested layman with sufficient mathematical background to the professional mathematician. This book goes some way towards making the appreciation of mathematics a reality. I look forward to reading the second volume.

F. GARETH ASHURST

Diversions in Modern Mathematics. By BARRY LEWIS. Heinemann Educational Books Ltd, London, 1981. Pp. 136. £4·95.

This interesting book has developed from a series of lectures designed to supplement and enrich a sixth-form mathematics course. There are seven chapters, each divided into two linked parts (some links being much closer than others). The material of each chapter might originally have filled a double-period lecture comfortably, though in the text there are now about a dozen questions to be answered in the course of each chapter, and five of the chapters end with collections of further investigations. A helpful appendix gives the answers to these questions, with enough guidance to make the book accessible to sixth formers working on their own.

The topics covered are (1) magic squares and figurative numbers, (2) tessellations and polyhedra, (3) construction of the rational numbers and field extensions, (4) construction and applications of the winding function, (5) paradoxes of the infinite and prime numbers, (6) congruences and Euclidean constructions, and (7) the development of (non-Euclidean) geometry, and geometry and reality. As the author readily admits, most of this can be found in the classics of recreational mathematics and elsewhere, but this does not detract from the carefully considered and efficient but friendly presentations given in this book.

There are a few mathematical blemishes, most seriously in Chapter 3, where the integers are constructed as equivalence classes of pairs of natural numbers, and then the rationals are obtained similarly from the integers. This involves defining new operations of addition and multiplication at each stage, but there is no discussion of whether these operations are well defined, i.e. independent of the particular number of pairs which are chosen to represent the new numbers. Also the definition of a field on p. 39 is incomplete, because it does not mention the link between addition and multiplication via the distributive law. Some of the views expressed seem rather odd: the uniqueness of prime factorisation is said on p. 72 to need 'a very subtle argument' to establish it, yet question 11 on p. 79 can easily be extended to give a straightforward proof. On the other hand, on the basis of the well-known $n^2 - n + 41$ (prime for $n < 41$) and $n^2 - 79n + 1601$ (prime for $n < 80$) it is said that 'one can easily construct polynomial functions which will give as many primes as required'. True, but they will not turn out to be nearly as simple as these exceptional examples.

Despite these quibbles, occasional obscure diagrams, and a few misprints, this is a stimulating and instructive little book, which may well prove to be a welcome eye-opener to many students. I was particularly pleased by two results, both new to me.

(1) If $a\ (\neq 10^n)$ is any natural number, then there is a power of a which begins with any sequence of digits we choose.

(2) For the table

$$\begin{array}{rrrrrl} 4 & 7 & 10 & 13 & 16 & \ldots \\ 7 & 12 & 17 & 22 & 27 & \ldots \\ 10 & 17 & 24 & 31 & 38 & \ldots \\ 13 & 22 & 31 & 38 & 45 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{array}$$

$2N + 1$ is prime if and only if N does not occur in the table. Intrigued? Read the book!

City of London School T. J. HEARD

Basic Facts: Mathematics. By CHRISTOPHER JONES and PETER CLAMP. Collins, London and Glasgow, 1982. Pp. 251. £1·50.

Possibly a useful pocket book for senior school students. With one or two exceptions, the entries (arranged alphabetically) appear to be lucid and helpful.

TEACHING STATISTICS IN SCHOOLS THROUGHOUT THE WORLD

A new publication of the International Statistical Institute, 1982, pp. xvi + 250, editor V. Barnett.

This volume reviews the exact nature of statistical education in approximately 20 different countries. Individuals with first-hand experience of the prevailing circumstances have given personal descriptions of the present situation, the way it has developed, and possible future prospects, in both developed and developing countries. The general structure of school education in each is outlined in order to clarify understanding of the provisions made for teaching statistics. Information is provided concerning the types of schools, the pupils catered for, principles of administration of the educational system, methods of teacher training, patterns for examinations, prospects for curriculum reform, etc.

The book is intended to be of service not only to those involved in teaching statistics as a separate discipline within schools, but also to those teachers and educators who are involved with teaching statistics as part of other disciplines, for example, biology, chemistry, mathematics, physics, and the social sciences.

The price of the volume is US \$10 or £6. Orders should be directed to:

The International Statistical Institute,
428 Prinses Beatrixlaan,
P.O. Box 950,
2270 AZ Voorburg,
Netherlands.

CONTENTS

ISSN 0025–5653

PRICES (*postage included*)

Prices for Volume 15 (Issues Nos. 1, 2 and 3):

Subscribers in	*Price per volume*
North, Central and South America	US\$11.50 or £6.00
Australia	\$A10.20 or £6.00
Britain, Europe and all other countries	£2.75

(Note: These overseas prices apply even if the order is placed by an agent in Britain.)

A discount of 10% will be allowed on all orders for five or more copies of Volume 15 sent to the same address

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