Mathematical Spectrum

Volume 12 1979/80 *Number 3*

A Magazine of *Contemporary Mathematics*

Published by the *Applied Probability Trust*

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

Volume 12 of *Mathematical Spectrum* consists of three issues, of which this is the third. The first issue was published in September 1979, and the second in January 1980.

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*?

The front cover of *Mathematical Spectrum* has remained the same since the magazine was first published in 1968: we wonder whether readers are happy with the present layout or would prefer a new look? Any comments will be welcome; in addition, a prize of £15 is offered for a suitable new cover design. To be considered, entries should be sent to the Editor, Mathematical Spectrum, Hicks Building, The University, Sheffield S3 7RH, to arrive not later than 1 September 1980.

The cover must obviously include the title *Mathematical Spectrum* and the volume and issue number, with date, and we should prefer to retain the spectrum symbol. The design should be suitable for printing in not more than two colours or shades on white paper.

Calculating Without a Scientific Calculator

HUGH NEILL
University of Durham

From 1966 to 1972 Hugh Neill was Head of the Mathematics Department at St. Paul's School, London. He is now a Lecturer in Mathematics at the University of Durham, where he devotes a lot of his time to liaison with local teachers and local authority advisers.

Most readers of this magazine will either own or have easy access to a scientific calculator. There was a time, only one or two years ago, when scientific calculators were still very expensive, and when I, in my innocence, thought I would never be able to afford one. However, I could afford at that time a calculator which had the usual four arithmetic operations, a square and a square root key, a reciprocal key and a memory, so I investigated ways in which I could use this relatively simple calculator to evaluate functions such as e^x, $\log x$, $\sin x$, $\cos x$ and so on. I set myself the requirement that the procedures I used had to be quick and reasonably accurate, without defining either of those terms very precisely.

These procedures are now only of academic interest, but the reasons why they work are well within the mathematics of A-level syllabuses. Thus it is these reasons which are of interest, not the procedures themselves. In the article which follows, I have given the procedures first, and then at the end, I have given some of the reasons. The reader is invited not to look at the reasons too soon.

e^x.

enter x → divide by 2 n times → add 1 → square n times → e^x displayed

The value of n, that is the number of divisions by 2 and squarings required will depend on the calculator being used and the value of x, but with x lying between -5 and 5 and a calculator which displays eight or ten figures, a value for n between ten and fifteen should suffice.

$\log x$. This flow diagram is simply the reverse of that for e^x and shows clearly that log and exp are inverse functions. Once again a value of n between ten and fifteen should suffice.

enter x → square root n times → subtract 1 → multiply by 2 n times → $\log x$ displayed

$\sin x$.

enter x in radians → divide by 3 n times. Result y → replace y by $3y - 4y^3$ n times → $\sin x$ displayed

This is probably the most awkward calculation suggested in this article. It may be quicker to calculate $\cos x$ and use $\sin x = \pm\sqrt{1 - \cos^2 x}$. A value of n of about ten or eleven should work.

$\sinh x$. This is exactly the same as $\sin x$, except that 'y is replaced by $3y + 4y^3$' takes the place of the existing third instruction. Alternatively it may be quicker to calculate $\sinh x$ by using the method for e^x together with

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}).$$

$\cos x$.

enter x in radians → divide by 2 n times. Result y → calculate $z = 1 - \frac{1}{2}y^2$ → replace z by $2z^2 - 1$ n times → $\cos x$ displayed

A value $n = 10$ should suffice. This is very much quicker than the method for $\sin x$.

An alternative, like the $\sin x$ method, is the following.

enter x in radians	→	divide by 3 n times. Result y	→	calculate $z = 1 - \frac{1}{2}y^2$	→	replace z by $4z^3 - 3z$ n times	→	$\cos x$ displayed

This is much slower than the first method. The reader may wish to consider why there is no quick method for calculating $\sin x$ like the first method for $\cos x$.

$\cosh x$. This is the same as either method for $\cos x$, except that the third instruction is replaced by $z = 1 + \frac{1}{2}y^2$. Alternatively it is possible to calculate e^x and use

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}).$$

$\tan x$. A direct method which is not very quick is the following.

enter x in radians	→	divide by 2 n times. Result y	→	replace y by $\dfrac{2y}{1-y^2}$ n times	→	$\tan x$ displayed

It is probably better to calculate $\cos x$ and use

$$\tan x = \pm\sqrt{\left(\frac{1}{\cos^2 x} - 1\right)}.$$

$\cos^{-1} x$.

enter x	→	calculate $\sqrt{\frac{1}{2}(1+x)}$ n times. Result y	→	calculate $z = \sqrt{2 - 2y}$	→	multiply by 2 n times	→	$\cos^{-1} x$ displayed

This method is surprisingly quick and should probably be used for all the inverse trigonometric functions instead of the more direct methods.

$\sin^{-1} x$. Given x, calculate $\sqrt{1 - x^2} = \pm y$ and then proceed to find $\cos^{-1} y$ or $\cos^{-1}(-y)$, whichever is appropriate.

$\tan^{-1} x$. Given x, calculate $y = 1/\sqrt{1 + x^2}$, and then calculate $\cos^{-1} y$ or $\cos^{-1}(-y)$, whichever is appropriate.

$\cosh^{-1} x$. This is the same as for $\cos^{-1} x$ except that the third instruction is replaced by $z = \sqrt{2 + 2y}$.

y^x. It is possible to devise special procedures for y^x but they are usually long and complicated. For most values of x it is probably quickest to calculate $\log y$, multiply the result by x and then use either the log function procedure reversed or (and this is the same thing) the exponential function procedure.

A special case occurs with cube roots. A good method of taking the cube root of a positive number N is first to make an intelligent guess, say x_0, to the cube root of N and then to use the procedure

$$x_{n+1} = \sqrt{\sqrt{Nx_n}}$$

until there is no significant change in the digits from x_n to x_{n+1}. For example, if the cube root of 100 is required to 4 significant figures and if 5 is taken as the first approximation x_0, we find that $x_5 = 4{\cdot}642$ is already sufficiently accurate for our needs.

Some discussion

The procedure for e^x may profitably be thought of in two ways.

First it may be seen from the flow diagram that the approximation being made when n is 15 is

$$e^x \approx \left(1 + \frac{x}{2^{15}}\right)^{2^{15}} = \left(1 + \frac{x}{32768}\right)^{32768}.$$

This is justified by the formula, which is sometimes taken as a definition of e^x,

$$e^x = \lim_{m \to \infty} \left(1 + \frac{x}{m}\right)^m.$$

The value 32768 for m is a compromise between m being sufficiently large for the approximation to be reasonable, and m being sufficiently small for the calculation not to build up too many errors.

The second way of looking at this is perhaps more instructive. After x has been entered, and divided by 2 fifteen times, the resulting number, $y = x/2^{15}$, is likely to be very small. The box which says 'add 1' is, in effect, making the approximation

$$e^y \approx 1 + y$$

when y is small, using the first two terms of the series expansion for e^y. What we have then is an approximation for e^y, or $\exp(x/2^{15})$, and we find e^x by using the identity

$$e^x = \left(\exp\left(\frac{x}{2^{15}}\right)\right)^{2^{15}}.$$

What has happened in this case may be described in general terms for a function $f(x)$ and then applied to some of the other examples. To calculate $f(x)$, we first convert x by some means to a small number y. Then the function is evaluated at y using the first few terms of a series. Finally, $f(x)$ has been calculated from $f(y)$ by using internal properties of the function $f(x)$. For the case when $f(x) = e^x$, the internal property used is

$$f(2z) = (f(z))^2.$$

The procedures for $\cos x$ may be described in precisely the same terms. In the first example given, the internal property used is

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

which enables $\cos 2\theta$ to be calculated when $\cos\theta$ is known. Thus, if we wish to calculate $\cos x$, where x is measured in radians, we convert x by successive halving to y, where

$$y = \frac{x}{2^{10}}.$$

Then we use $\cos \approx 1 - \frac{1}{2}y^2$. This is a very accurate formula indeed, there being no cubic terms in the expansion of $\cos y$. We then calculate $\cos 2y$, $\cos 4y$ and so on until we have $\cos(2^{10}y) = \cos x$. This general method is used in many of the procedures given.

It has already been remarked that many of the procedures may be reversed, thus illuminating the ideas of inverse functions and inverse operations, but it is instructive to look at the procedure for $\log x$ from a different point of view.

Suppose that $\log x$ is thought of as

$$\log x = \int_1^x \frac{dt}{t}.$$

Then it seems reasonable that, if the power of t is close to 1, but is not 1, the resulting function is not too different from $\log x$. What we are saying is

$$\log x \approx \int_1^x \frac{dt}{t^\alpha}$$

where α is close to 1. In particular, we take $\alpha = 1 - (1/m)$, where m is large. (The statement

$$\log x = \lim_{m\to\infty} \int_1^x \frac{dt}{t^{1-(1/m)}}$$

is actually true, but it needs more than 'school mathematics' to prove it. However, the fact that it isn't easy to justify something should not stop us from experimenting with it and using it.) Thus

$$\log x \approx \int_1^x \frac{dt}{t^{1-(1/m)}} = \int_1^x (t^{(1/m)-1})\, dt$$
$$= [mt^{1/m}]_1^x = m(x^{1/m} - 1).$$

Now take m to be 2^{15}, and we have the procedure given in the article; and

$$\log x \approx 2^{15}(x^{1/2^{15}} - 1).$$

Four Theorems for a Desert Island

DAVID W. SHARPE
University of Sheffield

David Sharpe is a Lecturer in Pure Mathematics at the University of Sheffield. He is at present the editor and problems editor of *Mathematical Spectrum*. His main mathematical interest is algebra.

The British Broadcasting Corporation has a long-running radio programme entitled 'Desert Island Discs', in which each week a well-known castaway is asked to choose eight gramophone records to accompany him on his desert island. If you were marooned on a desert island, which mathematical theorems would you choose to accompany you? I present here my personal choice. I have restricted myself to four rather than the full eight, in the hope that the reader will make up the number by his own four choices.

1. There are infinitely many prime numbers

My first theorem goes back to the Greek mathematician Euclid, well over 2000 years ago. Euclid is well known for his work on geometry, but it is to his interest in numbers that we turn. We are familiar with the notion of a 'prime number', a natural number greater than 1 which cannot be expressed as the product of two smaller natural numbers. Thus the first few prime numbers are 2, 3, 5, 7, 11, 13, 17, 19, 23. The question arises: how many prime numbers are there? Euclid showed that there are infinitely many. His argument is, for me, one of the most beautiful in the whole of mathematics.

Suppose the result is false, and denote the largest prime by p. If we multiply all the (finitely many) prime numbers together, and add one for luck, we obtain the number

$$n = (2 \times 3 \times 5 \times \cdots \times p) + 1.$$

It may perhaps be thought that this number is prime, in which case a prime larger than p has been found. We cannot, however, be sure that n is prime; after all, $(2 \times 3 \times 5 \times 7 \times 11 \times 13 \times 17) + 1 = 19 \times 26869$. But, in any case, n will have prime factors. (This is another famous result of Euclid's, the 'Fundamental Theorem of Arithmetic', over which we must reluctantly pass.) Now 2 cannot be a factor of n because of the 1 that was added for luck; nor can 3, nor 5, nor any of the prime numbers up to and including p. But from where can the prime factors of n come? It must mean that there is at least one prime number larger than p. Yet p was to be the largest prime. There is something wrong somewhere! The only point where an error has been committed is the assumption that the result is false. The result must therefore be true, and we have proved Euclid's theorem.

This, incidentally, will remind me on my desert island of the important method of proof known as 'proof by contradiction' or, for students of Latin, '*reductio ad* *absurdum*', whereby we assume the opposite of what we would like to prove and derive a contradiction or absurdity. In our case, the absurdity is that, having supposedly written down all the prime numbers, we then proved there was another one!

If I may be allowed to make a party point, this theorem also demonstrates the superiority of pure mathematics over applied mathematics; for no one has ever found infinitely many prime numbers! In 1966, it was discovered somewhat by accident in the University of Illinois in the U.S.A. that $2^{11213} - 1$ is prime. This was done by using the computer. This number was by far the largest known number at that time. In 1971, this was raised to $2^{11937} - 1$. Readers of the Daily Telegraph of 20 November 1978 were no doubt fascinated to learn that two 18-year-old students at California State University had discovered that $2^{21701} - 1$ is prime. As far as I know, that is the story to date.† Should I be in danger of boredom on my desert island, I can always look for a larger prime, although this might be difficult if there is no computer there. But I shall at least know from Euclid's theorem that there must be a larger one.

Readers may be wondering why these three large prime numbers are all of the form $2^n - 1$. Is every prime number of this form? No, because 2 is not; nor is 5, nor 11. It is simply that these are likely candidates to be prime numbers, at least when n is prime. A prime of the form $2^n - 1$ is called a *Mersenne Prime*, after the amateur mathematician Father Marin Mersenne (1588–1648), a friend of Descartes. Readers may like to try to prove that, if $2^n - 1$ is prime, then n must be prime. It would be nice if the converse were true, because that would then give us a formula for producing infinitely many primes. However, $2^{11} - 1 = 23 \times 89$. Life is just not that simple.

2. The infinite series $\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}$

My second result is perhaps better described as a formula rather than a theorem; it is the surprising result that the sum of the infinite series shown is $\pi^2/6$. The Greek mathematicians were not able to grasp the fact that the sum of infinitely many positive numbers could be finite. We may recall the paradox proposed by Zeno in the fifth century BC. The swift-footed Achilles is to race the tortoise, and magnanimously give the tortoise a start. It will take x_1 seconds (say) for Achilles to reach the starting point of the tortoise, during which time the tortoise will have made a little progress. It will take, say, x_2 seconds for Achilles to move from this point to the point reached by the tortoise after x_1 seconds, during which time the tortoise will again have plodded its weary way onwards. And so on. Continuing this indefinitely, we see that it will take $x_1 + x_2 + x_3 + \cdots$ seconds for Achilles to catch up with the tortoise. But are we not here adding infinitely many positive numbers together, so

† I should have known better than to write this! In the same year, it was discovered that $2^{23209} - 1$ is prime, and in 1979 that $2^{44497} - 1$ is prime. 1980 may well be the year of yet bigger primes.

does this not mean that Achilles will never catch up the tortoise! This is where the Greeks were stuck.

Of course, with our superior knowledge, we know perfectly well that infinitely many positive numbers can have a finite sum. After all, if I play around with some infinite series on my desert island, I may realize that

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 2.$$

On the other hand,

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} > 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 2,$$

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = 2\tfrac{1}{2}.$$

Were I to continue like this, I would quickly convince myself that

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

is infinite. Again,

$$1 + \frac{1}{2^2} + \frac{1}{3^2} < 1 + \frac{1}{2^2} + \frac{1}{2^2} = 1 + \frac{1}{2},$$

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \frac{1}{7^2} < 1 + \frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{4^2} + \frac{1}{4^2} + \frac{1}{4^2}$$

$$= 1 + \frac{1}{2} + \frac{1}{4}.$$

Continuing in this way, I could perhaps come to the conclusion that

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots < 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2.$$

But how on earth could I obtain the result $\pi^2/6$?

I am indebted to my colleague Dr R. J. Webster for bringing to my attention a fascinating 'proof' of this formula by the eighteenth century Swiss mathematician Euler. I use the inverted commas advisedly, because, as you will see, Euler's rigour leaves much to be desired. This is how Euler argued.

Suppose I have a quadratic equation $1 + ax + bx^2 = 0$, with roots α, β. Then

$$\frac{1}{\alpha} + \frac{1}{\beta} = \frac{\alpha + \beta}{\alpha\beta} = \frac{-(a/b)}{(1/b)} = -a.$$

Similarly, for a cubic equation $1 + ax + bx^2 + cx^3 = 0$, with roots α, β, γ, it is easy to see again that

$$\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = -a.$$

Fine, said Euler. Suppose we have the equation

$$1 + ax + bx^2 + \cdots = 0$$

with roots $\alpha_1, \alpha_2, \alpha_3, \ldots$. Why not

$$\frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \frac{1}{\alpha_3} + \cdots = -a?$$

Now I seem to remember that

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots,$$

so that

$$\sin\sqrt{x} = \sqrt{x} - \frac{1}{3!}x\sqrt{x} + \frac{1}{5!}x^2\sqrt{x} - \cdots,$$

and

$$\frac{\sin\sqrt{x}}{\sqrt{x}} = 1 - \frac{1}{3!}x + \frac{1}{5!}x^2 - \cdots.$$

This certainly vanishes when $\sqrt{x} = \pi, 2\pi, 3\pi, \ldots$, i.e. when $x = \pi^2, 2^2\pi^2, 3^2\pi^2, \ldots$. So, by what we have above,

$$\frac{1}{\pi^2} + \frac{1}{2^2\pi^2} + \frac{1}{3^2\pi^2} + \cdots = \frac{1}{3!},$$

i.e.

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}.$$

We may note that $1 < \pi^2/6 < 2$. I shall puzzle for hours on my desert island as to why this argument of Euler works.

3. A theorem for a party

I imagine that the most acute problem that I shall face on this desert island is loneliness, and this theorem will remind me of the folks back home. A number of people go to a party and, on arrival, some shake hands with others. The theorem asserts that there are two people who shake hands the same number of times.

Suppose there are n people at the party, labelled $1, 2, \ldots, n$, and denote by x_i the number of times the ith person shakes hands on arrival. We had better suppose that no one is fulsome or forgetful enough to shake hands with someone more than once, so that each x_i is an integer between 0 and $n - 1$. (The reader will notice that no one is so singular as to shake hands with himself.) On the face of it, it is perfectly possible to have n distinct integers between 0 and $n - 1$, until we realize that, should anyone shake hands $n - 1$ times, then he must shake hands with everyone else at the party, and in this case no one can shake hands no times. Thus the integers 0 and $n - 1$ cannot both be included among the x_i, so the n integers $x_1, \ldots, x_n$ may only take up to $n - 1$ distinct values.

At this point we invoke the aptly named 'pigeonhole principle'. If we have n letters to place in fewer than n pigeonholes, then some pigeonhole must contain more than one letter (as is well-known to all employees of the Post Office). This means that two of the x_i's must take the same value, and this delightfully simple theorem is proved.

4. The sailors, the coconuts and the monkey

My last result is again hardly a theorem, but rather a puzzle. And it has obvious application to the situation in hand. The reader will find this well-known problem fully discussed in Chapter 9 of Martin Gardner's delightful book *More Mathematical Puzzles and Diversions* (Penguin Books, Harmondsworth). Five sailors are marooned on a desert island, and spend all one day collecting coconuts, after which they go to sleep. During the night, one sailor wakes up and, not trusting the others, he divides the coconuts into five equal piles, sees that there is one left over, which he gives to the monkey, hides his share and goes back to sleep. A little later, a second sailor wakes up, divides the remaining coconuts into five equal piles, sees that there is one left over which he gives to the monkey, hides his share and goes to sleep. This happens for each sailor in turn. In the morning, the sailors wake up and share the remaining coconuts into five equal piles and see that there is one left over which they give to the monkey. Apart from the information we may glean about human nature, the question is: what is the smallest number of coconuts that the sailors could have collected?

The reader will find Martin Gardner's discussion of this problem very rewarding, especially his concept of 'negative coconuts'. What fascinates me about this problem is that it illustrates the use of congruences. Suppose the sailors collect n coconuts. Then we have to solve the equations

$$\begin{aligned} n &= 5n_1 + 1, \\ 4n_1 &= 5n_2 + 1, \\ 4n_2 &= 5n_3 + 1, \\ 4n_3 &= 5n_4 + 1, \\ 4n_4 &= 5n_5 + 1, \\ 4n_5 &= 5n_6 + 1 \end{aligned}$$

in integers. The last equation tells us that $4n_5 - 1$ must be divisible by 5. In the language of congruences, this is written

$$4n_5 \equiv 1 \pmod 5 \tag{1}$$

(read '$4n_5$ is congruent to 1 modulo 5'). In general, we say that $a \equiv b \pmod c$ (where a, b, c are integers with $c > 0$) if $a - b$ is divisible by c. Now $4n_5 \equiv -n_5 \pmod 5$, because $4n_5 + n_5$ is divisible by 5. Hence (1) can be replaced by

$$n_5 \equiv -1 \pmod 5.$$

We can multiply by 5 to give

$$5n_5 \equiv -5 \pmod{25}.$$

Since $4n_4 = 5n_5 + 1$, we now have

$$4n_4 \equiv -4 \pmod{25}.$$

I feel in my bones that I can cancel the 4 here (perhaps because 4 and 25 have no common factors), to give

$$n_4 \equiv -1 \pmod{25}.$$

It should be clear now how the argument proceeds, and we simply write down the steps without comment.

$$\begin{aligned}
5n_4 &\equiv -5 \pmod{125},\\
4n_3 &\equiv -4 \pmod{125},\\
n_3 &\equiv -1 \pmod{125},\\
5n_3 &\equiv -5 \pmod{625},\\
4n_2 &\equiv -4 \pmod{625},\\
n_2 &\equiv -1 \pmod{625},\\
5n_2 &\equiv -5 \pmod{3125},\\
4n_1 &\equiv -4 \pmod{3125},\\
n_1 &\equiv -1 \pmod{3125},\\
5n_1 &\equiv -5 \pmod{15625},\\
n &\equiv -4 \pmod{15625}.
\end{aligned}$$

As Martin Gardner pointed out, -4 coconuts provide, in theory at least, a solution of the problem. Should you require a positive number of coconuts, then the smallest number of coconuts that the sailors could have collected is 15621. And, should I be unsure whether or not I have manipulated the congruences correctly, I can become an experimentalist and actually collect coconuts myself, always presupposing that my particular desert island has any coconuts—and a monkey.†

If we follow the pattern set by the BBC, I am allowed to take one book apart from the Bible and the collected works of William Shakespeare, which are already on the island. My choice is clear: *Alice's Adventures in Wonderland*, by Lewis Carroll (alias the Rev. Charles Dodgson, sometime mathematics don at Christ Church College, Oxford). And one luxury? Certainly not a computer. Not even a pocket calculator. Simply an infinite (or at least unbounded) supply of pencils and paper. It's so much more convenient than drawing with one's finger in the sand.

† A colleague has raised serious doubts as to whether there would be enough time for all this nocturnal activity.

The 'Simple' Pendulum

D. W. JORDAN and **P. SMITH**
University of Keele

Both these authors teach mathematics at Keele University. Their main interests are in non-linear differential equations, applied analysis and their applications.

A motion is called *periodic* if it repeats itself over and over again in equal times, without change of form. A 'system' which exhibits periodic motion is called an *oscillator*. Many oscillations are visible to us: the daily rotation of the Earth, the annual revolution of the Earth about the Sun, waves on the sea, the pendulum in a clock. Other oscillations we can sense but are not readily visible: sounds produced by musical instruments, the quartz crystal oscillator in a digital clock, the heartbeat. Yet other oscillators we suspect to exist but can only see their effects, as in the physiological timing mechanisms of animals, whose daily and seasonal rhythms are controlled by such 'biological clocks', and not entirely by changes taking place outside themselves. Serious disruption of these rhythms can lead to such symptoms as 'jet lag' after crossing time zones.

Some oscillations are desirable, such as those in the clock, in which good timekeeping is needed. Other oscillations or vibrations may be undesirable and ultimately destructive. The spectacular collapse of the Tacoma Narrows bridge in Washington in 1940 (shown many times on film) displayed vividly what can happen when periodic motion gets out of hand.

To describe periodic motion we need suitable mathematical functions which make the periodicity obvious. The simplest is the cosine (or sine) function. Figure 1(a) shows a graph of $x = \cos t$; $x = \cos(t + \frac{1}{4}\pi)$ is shown in Figure 1(b)—the same shape but moved to the left a distance $\frac{1}{4}\pi$; $x = \cos(t - \frac{1}{2}\pi)$ is shown in Figure 1(c) (equal to $\sin t$ of course). The *period* of all these is 2π, because each time interval 2π contains one complete cycle or oscillation.

If $x = \cos \omega t$, where ω is a positive constant, the period is $2\pi/\omega$ since $\cos[\omega(t + 2\pi/\omega)] = \cos \omega t$. If $0 < \omega < 1$ the graph in 1(a) is stretched to give a longer period and if $\omega > 1$ it is compressed. A general form of similar type is

$$x = a\cos(\omega t + \alpha) \tag{1}$$

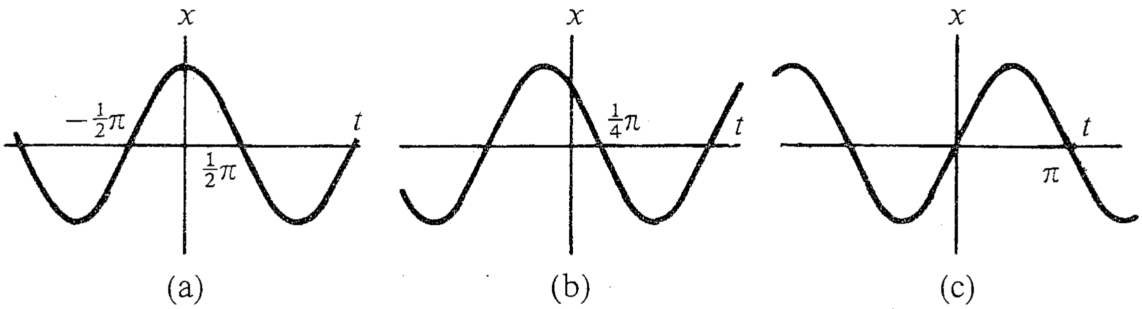


Figure 1. (a) $x = \cos t$ (b) $x = \cos(t + \frac{1}{4}\pi)$ (c) $x = \cos(t - \frac{1}{2}\pi)$.

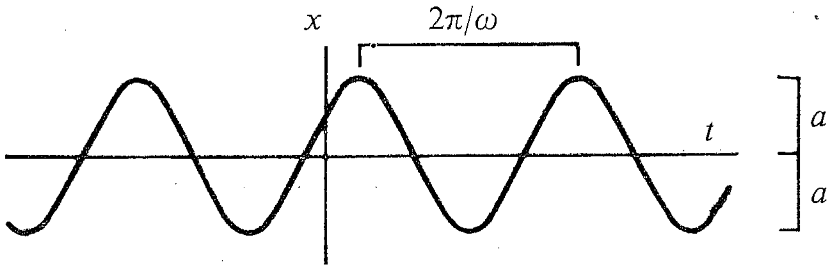


Figure 2. $x = a\cos(\omega t + \alpha)$.

(Figure 2), where a, ω and α are constants ($a > 0, \omega > 0$). Here a is called the *amplitude* (the half-width of the swing) and α is the *phase* (α/ω is the time-displacement of the oscillation). The number of swings per second is called the *frequency*, and is equal to $\omega/2\pi$. An oscillation described by (1) is called *simple harmonic motion*.

If we differentiate (1) twice with respect to t, then

$$\frac{dx}{dt} = -a\omega\sin(\omega t + \alpha) \qquad \text{and} \qquad \frac{d^2x}{dt^2} = -a\omega^2\cos(\omega t + \alpha). \tag{2}$$

Evidently

$$\frac{d^2x}{dt^2} + \omega^2 x = 0 \tag{3}$$

which is the differential equation for simple harmonic motion. If we had set out to solve it, we would have found that all its solutions are of the form (1). This is called the 'general solution' where a and α are arbitrary constants.

This equation can also arise as an approximation to the pendulum equation. Consider the pendulum (Figure 3) idealised as a bob of mass m suspended from 0 by

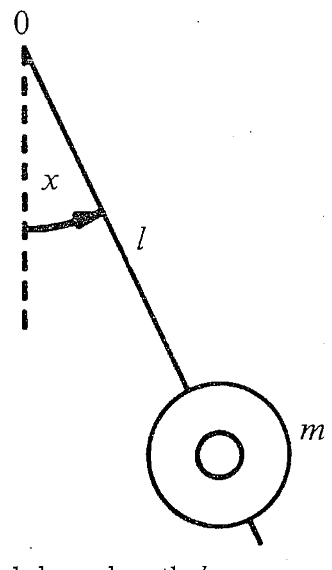


Figure 3. Pendulum: length l, mass m, inclination x.

a light rod of length l. Assume that there is no friction or air resistance. Then if x is the angle made with the downward vertical, taking moments gives

$$-mgl\sin x = ml^2\frac{d^2x}{dt^2}. \tag{4}$$

The equation can be rearranged into

$$\frac{d^2x}{dt^2} + \omega^2\sin x = 0, \tag{5}$$

where $\omega^2 = g/l$. This is the 'equation of motion' of the pendulum and it is difficult to solve, although we know intuitively that the pendulum is capable of only a few physically quite simple types of motion. It can swing to and fro for ever in motion which looks simple harmonic, or it can whirl round and round endlessly. It can also be suspended in equilibrium with $x = 0$ or balanced in unstable equilibrium with $x = \pi$.

Simple solutions can be obtained by supposing that the swings are 'small', in which case $\sin x \approx x$. Equation (5) becomes approximately

$$\frac{d^2x}{dt^2} + \omega^2 x = 0,$$

which has solutions given by (1). It predicts that, provided the swings have small enough amplitude, the period is effectively independent of the amplitude. In fact the approximation works quite well even for swings as large as 30° ($x < 0{\cdot}53$ radians).

Next suppose we want to investigate larger swings, or need a more accurate account of small swings. Even large swings *look* simple harmonic, so suppose that they very nearly are; that is, suppose

$$x \approx a\cos \nu t, \tag{6}$$

where we expect to find that the frequency ν has some dependence on the amplitude a. Then this function should nearly satisfy equation (5). The $\sin x$ term is awkward, so write, for smallish oscillations,

$$\sin x \approx x - \tfrac{1}{6}x^3 \tag{7}$$

in the equation, which is accurate to 4 % at 80° ($x = 0{\cdot}985$ radians). Now put first (7) and then (6) into the pendulum equation (5). This becomes, approximately,

$$\frac{d^2x}{dt^2} + \omega^2\sin x \approx -a\nu^2\cos \nu t + \omega^2(a\cos \nu t - \tfrac{1}{6}a^3\cos^3 \nu t)$$

$$\approx (-a\nu^2 + a\omega^2 - \tfrac{1}{8}a^3\omega^2)\cos \nu t - \tfrac{1}{24}a^3\cos 3\nu t \tag{8}$$

(where we have used the known identity $\cos^3 \nu t = \frac{3}{4}\cos \nu t + \frac{1}{4}\cos 3\nu t$ in the last equation). The expression (8) should be zero for all values of t, but because (6) is only an approximate solution this cannot be made exactly true, no matter how we choose a and ν. But if, in the spirit of the approximation, we attempt to satisfy it only so far

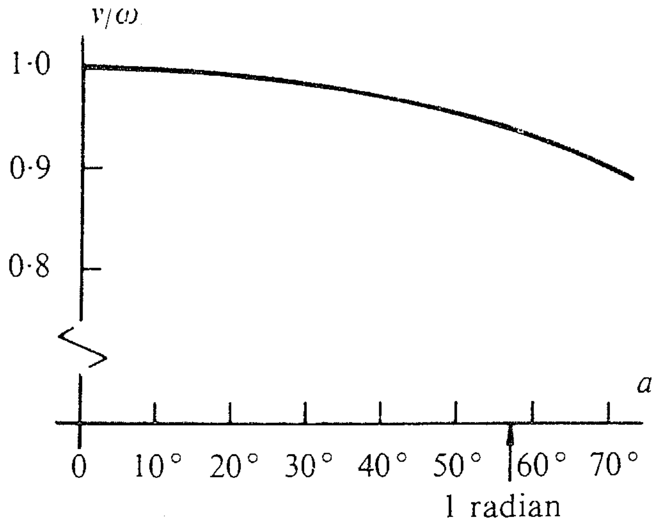


Figure 4. Frequency–amplitude relation for a pendulum: $\nu = \omega\sqrt{1 - \frac{1}{8}a^2}$.

as the cos νt term which we have assumed dominates the behaviour, then we find that

$$-a\nu^2 + a\omega^2 - \tfrac{1}{8}a^3\omega^2 = 0$$

or

$$\nu = \omega\sqrt{1 - \tfrac{1}{8}a^2}, \qquad (\approx \omega(1 - \tfrac{1}{16}a^2) \text{ for } a \text{ small})$$

which shows (approximately) how ν depends on the amplitude (see Figure 4). As a increases, ν decreases, and the period $2\pi/\nu$ increases.

We have not yet been able to reveal any sign of the whirling motions which we know can take place, and for this a completely new approach is needed. In any motion of the pendulum x and dx/dt (the angular velocity) each have some value at any time t. Call y the value of dx/dt at time t. Then, by definition,

$$\frac{dx}{dt} = y \tag{9}$$

and from equation (5)

$$\frac{dy}{dt} = \frac{d^2x}{dt^2} = -\omega^2 \sin x. \tag{10}$$

We eliminate t between these equations as follows:

$$\frac{dy/dt}{dx/dt} = \frac{dy}{dx} = -\frac{\omega^2 \sin x}{y}$$

or

$$y\,dy = -\omega^2 \sin x\,dx.$$

This equation can now be integrated to give

$$\tfrac{1}{2}y^2 = \omega^2 \cos x + c$$

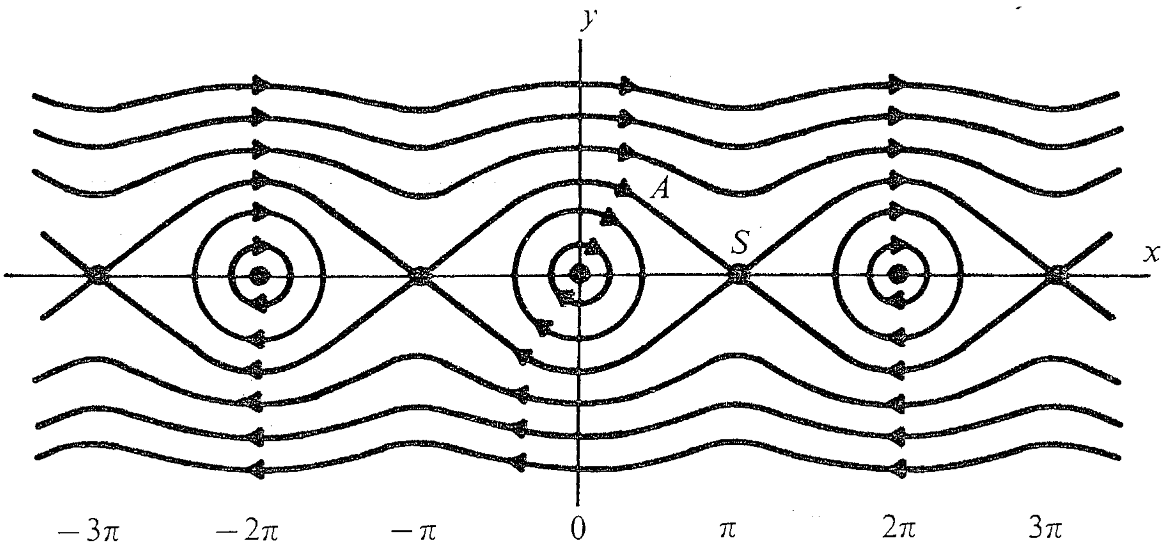


Figure 5. Phase diagram for a pendulum.

where c is an arbitrary constant (essentially this is the energy equation for the pendulum). Suppose now that these curves are plotted in the (x, y) plane, each for a different value of c, as shown in Figure 5. Each one represents a possible motion of the pendulum. The arrows on the curve show the directions that we must follow as time increases: these directions are obtained from equation (9), for this shows that, if $y > 0$, x increases with time. Figure 5 is the *phase diagram* for the pendulum and the individual curves are known as *phase paths*. The phase diagram repeats itself in every interval of 2π in x (remember x is an angle).

All possible physical behaviour of the pendulum can be read off from the phase diagram. The points $(x, y) = (0, 0)$, $(\pm 2\pi, 0)$, $(\pm 4\pi, 0), \ldots$ represent the normal hanging equilibrium positions (these points are really degenerate curves). The points $(0, \pm\pi)$, $(0, \pm 3\pi), \ldots$ represent the pendulum when it is vertical, in equilibrium with the bob above the support. Both of these are called *equilibrium points*. The *closed paths* around $(0, 0)$ indicate the periodic oscillations (known locally as a *centre* on the diagram). The paths round $(\pm\pi, 0)$ indicate unstable equilibrium (why?) and are called locally *saddle points* (can you interpret the curve AS?). The wavy curves above and below the closed curves correspond to the whirling motions in either direction.

The above method applies to many second-order differential equations. For example for the simple harmonic oscillator of equation (3), the phase paths are given by

$$\frac{x^2}{a^2} + \frac{y^2}{\omega^2 a^2} = 1$$

where a is the constant of integration. What does its phase diagram look like?

A real pendulum will experience friction (or *damping* as it is often known for a pendulum). Damping usually depends on velocity and

$$\frac{d^2x}{dt^2} + 2k\frac{dx}{dt} + \omega^2 x = 0 \qquad (k > 0)$$

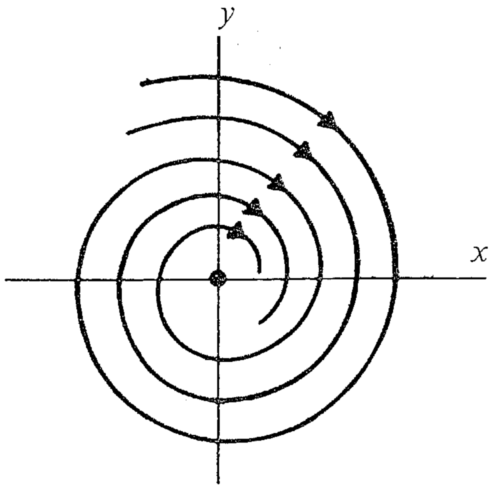


Figure 6. Phase diagram for a damped simple harmonic oscillator.

is a simple example of a *damped simple harmonic oscillator*. The second term $2k(dx/dt)$ represents the friction. The general solution for small k (weak damping) can be written as

$$x = ae^{-kt}\cos(\sqrt{(\omega^2 - k^2)}t + \alpha).$$

The phase diagram in the x and y $(= dx/dt)$ plane now becomes a set of interlaced spirals contracting to the origin, as shown in Figure 6. The injection of a similar friction term into the pendulum equation (5) leads to Figure 7. Can you now interpret the phase paths for a damped pendulum?

The regularity of the pendulum beat suggests that a pendulum could serve as the basis of a timekeeping device, the most obvious disadvantage being that the swings soon die away unless somehow they can be kept going. Galileo was probably the first to suggest a driven pendulum as a timekeeper. The requirements for an accurate clock are extraordinary. An error of 1 minute in a day means accuracy in timing to

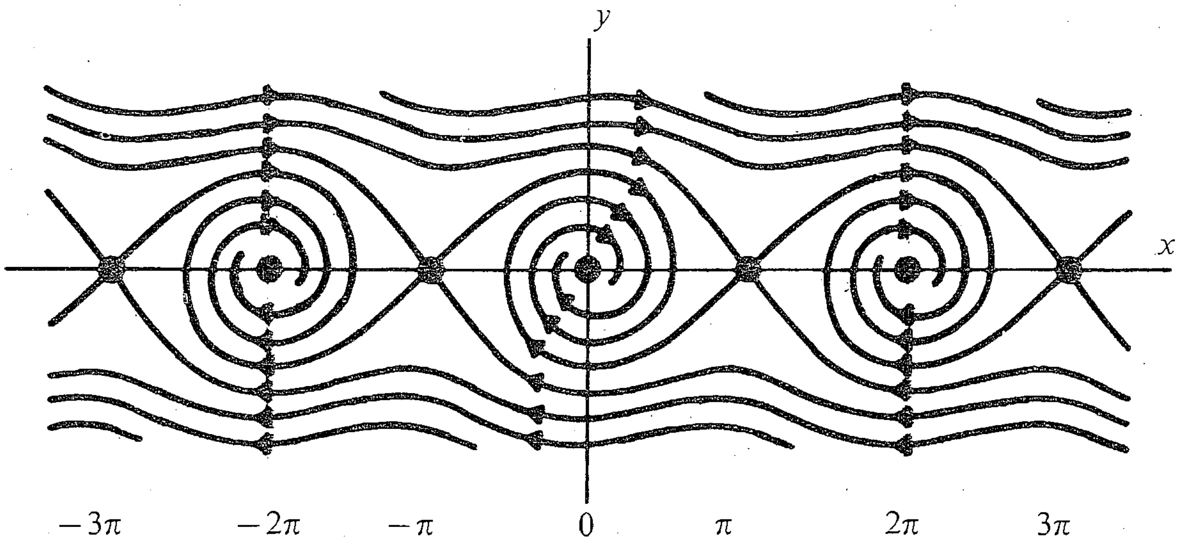


Figure 7. Phase diagram for a damped pendulum.

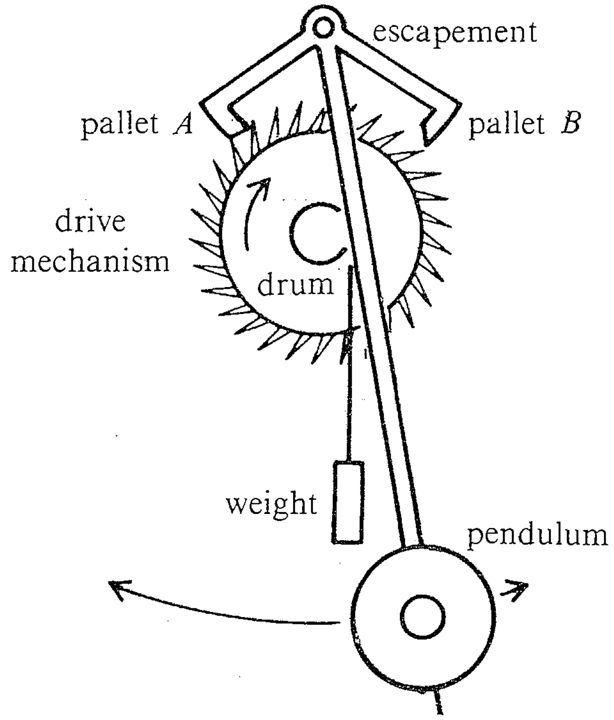


Figure 8. Basic drive and timing mechanism for a clock.

within 0·07%. A pendulum with period 1 second would swing more than 30 000 000 times in a year, so any mechanism must be very stable with respect to wear, disturbance and climatic conditions. This could be achieved without an impossibly high standard of technology because, as we shall see, a clock is largely self-regulating.

Ideally a clock pendulum should swing at a constant small amplitude: small so that changes in amplitude due to wear affect the period as little as possible (see Figure 4). Figure 8 shows the essential controlling mechanism for a clock: the 'escapement'. The clock's source of energy is a slowly descending weight which turns the drum: this simultaneously drives the hands and supplies impulses to the pendulum. As the pendulum swings and the escapement rocks along with it, two impulses are given in each complete swing as it moves to left or right through its vertical position; these are initiated at the moment when one of the 'pallets', A or B, engages with a tooth and the other one disengages. Thus the pendulum is regularly supplied with impulses to keep it swinging, and in turn controls the unwinding of the drum which drives the hands. (To make the teeth engage with the escapement so as to give impulses in the right directions these elements have to be very precisely designed. A description will be found in reference 2.)

The question is: why should it swing with a constant amplitude? If there were no friction it would move erratically up to large amplitudes, but in fact friction arises from several sources. We shall show by means of the phase plane how friction controls the oscillations so that the amplitude will settle to a steady value no matter how we start the clock off or disturb it. Paradoxically friction, which makes a pendulum alone useless, is turned to advantage, and is indeed essential for the clock to work at all.

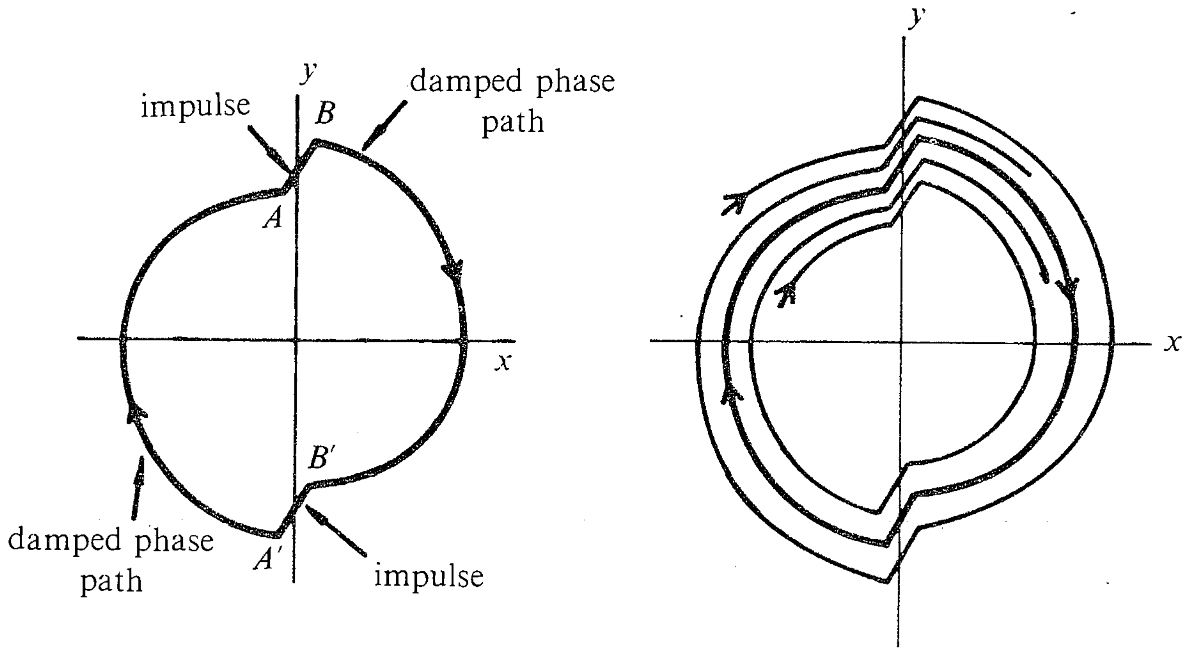


Figure 9. (a) Regular oscillation of the pendulum in a clock. (b) Phase diagram for the motions of a clock pendulum.

To simplify matters, suppose that a constant momentary restoring impulse at each passage through the vertical is the *only* action of the escapement on the toothed wheel and hence on the pendulum; and that between the impulses the pendulum is damped but otherwise swings freely. Then the 'phase path' corresponding to its swings in the normal state looks like Figure 9(a). In the regions $x > 0$ and $x < 0$ respectively the curves consist of segments of appropriate spirals from Figure 6, chosen so that, together with the jumps AB and $B'A'$ caused by the impulses, the pieces fit together to make a continuous closed curve. The closed curve represents a periodic motion, as we said earlier. If the impulses (or the jumps in angular velocity AB, $B'A'$) are of fixed magnitude (they do not have to be equal) there is in fact only one such closed curve that can be constructed.

Now suppose we disturb the pendulum, or start it going in some manner which does not exactly correspond to a point on the curve of Figure 9(a) (for instance if we give it a little push to start it). Then we know in practice that after a swing or two it will settle down into its normal regular state. What happens is illustrated in the phase diagram, Figure 9(b). However we start it, the phase path it follows gradually spirals in or out towards the special closed curve we found before. We can sum up the action in this way. If the amplitude is too large, then the impulses are not big enough to offset the friction, so the amplitude tends to drop. If the amplitude is too small, then the impulses have more effect than the friction and the amplitude increases. Between these two conditions lies the perfect periodic state which is approached (by the spirals), no matter where we start from, within reason. The clock is therefore self-regulating; by means of the escapement it exactly times its own energy inputs, and it settles its own characteristic amplitude by automatically balancing energy output (from the impulses) and energy output (through friction).

The isolated closed curve surrounded by spirals on Figure 9(b) is called a *limit cycle*. Whenever limit cycles occur they are of the greatest importance, because they describe the 'normal condition' of oscillation. The oscillator spontaneously moves into this state no matter what we do to it.

What we have illustrated here in connection with the pendulum are, in effect, some quite general methods of finding what the solutions of equations are like when we are not able to solve them exactly. It is really very unlikely that a random differential equation can be solved exactly; but we can often find the general character of its solutions by a mixture of graphical, physical and approximative methods such as these.

References

1. D. W. Jordan and P. Smith, *Nonlinear Ordinary Differential Equations* (Oxford University Press, 1977).
2. E. Smith, *Clocks* (David and Charles, Newton Abbot, 1977).

Perfect Numbers and Mersenne Primes

E. J. SCOURFIELD
Westfield College, London

The author graduated from the University of Exeter, and obtained her Ph.D. from the University of Glasgow whilst lecturing there. She is now a Lecturer in Mathematics at Westfield College (University of London), and her research interests lie in the theory of numbers. In the hope of encouraging others to share their enthusiasm for mathematics, Dr Scourfield and her colleagues have given talks on various mathematical topics to sixth-formers at schools within reach of Hampstead, and the present article developed from some of her talks.

1. Introduction

The problem of determining perfect numbers is very old, and was certainly investigated by the Greeks over 2000 years ago. At that time certain numbers, such as perfect numbers, were regarded as having a special symbolic and mystical significance, distinct from their ordinary numerical usage. Our definition below of a perfect number was known to Euclid, and the final proposition of the ninth book of his *Elements* (300 BC) is our first theorem.

A positive integer d is said to be a *divisor* (or factor) of the positive integer n if the quotient n/d is an integer. Each integer $n > 1$ has at least two divisors, namely 1 and n; if it has no others, it is called *prime*, and otherwise it is *composite* (the integer 1 being put in a category of its own). The first ten primes are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, and there are infinitely many primes (i.e., the list of primes is unending). A composite integer n has at least two prime divisors p_1, p_2 (which are equal when

$n = p^\alpha$ where p is a prime and $\alpha \geq 2$, so p^2 divides n); since then $1 < p_1 \leq p_2 < n$ and $p_1 p_2$ is also a divisor of n, it follows that $p_1^2 \leq p_1 p_2 \leq n$, whence $p_1 \leq \sqrt{n}$. Thus a composite integer n always has a prime divisor p with $1 < p \leq \sqrt{n}$, a fact that we shall use later.

A *perfect number* is a positive integer that equals the sum of all its positive divisors other than itself. Thus, if $s(n)$ denotes the sum of the divisors of n other than n, then n is perfect if and only if $s(n) = n$. The first two perfect numbers are 6 and 28, for the divisors of 6 are 1, 2, 3, 6 and those of 28 are 1, 2, 4, 7, 14, 28; and

$$s(6) = 1 + 2 + 3 = 6, \qquad s(28) = 1 + 2 + 4 + 7 + 14 = 28.$$

It is easily shown that there are no other perfect numbers less than 40 (say).

If $s(n) < n$, n is said to be *deficient*, and if $s(n) > n$, n is *abundant*. Thus perfect numbers strike the happy medium between abundancy and deficiency, which probably contributed to their mystic appeal in ancient times. Any prime p, or power of p, is deficient, for $s(p) = 1$ always and

$$s(p^\alpha) = 1 + p + p^2 + \cdots + p^{\alpha-1} = (p^\alpha - 1)/(p - 1) < p^\alpha$$

for each integer $\alpha \geq 1$. The first six abundant numbers are 12, 18, 20, 24, 30, 36, with all other integers between 1 and 39, apart from the perfect numbers 6 and 28, being deficient; the first odd abundant number if 945.

It is still not known today whether there are any odd perfect numbers. However if there is one, it is very large (greater than 10^{200}) and must satisfy several conditions; for example, if n is an odd perfect number, then $n = p^{4\alpha+1} N^2$ where α is a non-negative integer and p is a prime of the form $4m + 1$ not dividing the odd integer N (m being a positive integer), and also n has at least 8 different prime factors the largest of which exceeds 100110.

Rather more has been established about even perfect numbers but there are still many unanswered questions. In particular it is not yet known whether there are infinitely many even perfect numbers. It is conjectured that there are, although they occur very rarely; only 27 are known at present, the greatest of which are extremely large and include two discovered by two American students (see Section 3)—an exciting development! From time to time, when faster computers or new techniques become available, a new perfect number is discovered, and there is no reason to believe that this process will terminate.

The first four even perfect numbers 6, 28, 496, 8128 were known to the Greeks and recorded by Nicomachus about 100 AD, when he noted that they end alternately in 6 and 8, and there is just one in the intervals between 1, 10, 100, 1000, 10000. In consequence it was thought in the Middle Ages that the mth perfect number has exactly m digits and ends in 6 if m is odd and 8 if m is even. By the middle of the fifteenth century it was known that the fifth perfect number is 33550336, which disposed of the first of these assertions. The second was disproved when, a century or so later, the sixth perfect number (stated in Section 3) was found to end in 6 also. However, as we shall see in Theorem 4 below, a modification of the second assertion is true.

Before we continue (in Section 3) outlining the historical landmarks in the search to find even perfect numbers, we obtain in Section 2 a precise mathematical description of them which itself raises new questions.

2. Characterization of even perfect numbers

We shall assume the fundamental theorem of arithmetic, which says that every integer exceeding 1 can be written uniquely (apart from the order) as a product of primes; the reader will find a proof of this result in reference 2, page 3, or reference 4, page 6, for example, and all theorems below are proved in reference 4 too.

We begin by proving Euclid's result referred to in Section 1.

Theorem 1. *If* $2^n - 1$ *is prime, then*

$$N = 2^{n-1}(2^n - 1)$$

is a perfect number.

Proof. Let $q = 2^n - 1$ be a prime, so that its only positive divisors are $1, q$; then the positive divisors of $N = 2^{n-1}q$ are exactly

$$1, 2, 2^2, \ldots, 2^{n-1}, q, 2q, \ldots, 2^{n-1}q\,(= N).$$

Hence, adding all these divisors except the last,

$$\begin{aligned} s(N) &= 1 + 2 + 2^2 + \cdots + 2^{n-1} + q(1 + 2 + \cdots + 2^{n-2}) \\ &= 2^n - 1 + q(2^{n-1} - 1) = q + q(2^{n-1} - 1) = N, \end{aligned}$$

and so N is perfect.

Note that when $n = 2, 3$, $2^n - 1$ is prime, and so the above result then yields the perfect numbers 6, 28, but when $n = 4$, $2^n - 1 = 3{\cdot}5$ which is composite. This prompts us to ask when $2^n - 1$ is prime.

Theorem 2. *If* $2^n - 1$ *is prime, then* n *is prime.*

Proof. Suppose that n is composite, so that $n = kl$ where k, l are integers satisfying $1 < k \leq l < n$. Then from the factorization

$$a^l - 1 = (a - 1)(a^{l-1} + a^{l-2} + \cdots + a + 1)$$

with $a = 2^k$, we obtain

$$2^n - 1 = (2^k)^l - 1 = (2^k - 1)(2^{k(l-1)} + \cdots + 2^k + 1). \tag{1}$$

Since $l \geq k \geq 2$, $2^k - 1 \geq 3$ and $2^{k(l-1)} + \cdots + 2^k + 1 \geq 2^k + 1 \geq 5$, and hence it follows from (1) that $2^n - 1$ is composite whenever n is composite. Thus if $2^n - 1$ is prime, n itself must be prime.

The converse of Theorem 2 does not hold, for $2^n - 1$ can be composite when n itself is prime; for example, 11 is prime but $2^{11} - 1 = 23\,.\,89$ which is composite.

A number of the form $M_p = 2^p - 1$, where p is prime, is called a *Mersenne number* or, when prime, a *Mersenne prime*. Father Marin Mersenne (1588–1648) was a Minim Friar in Paris with a great interest in mathematics, and he corresponded assiduously with the leading French mathematicians of the day about, amongst other things, perfect numbers and primes of the form M_p; but the reason for naming these primes after him is not too clear, for he did not himself obtain any significant results about them although he made assertions (see Section 3).

If it could be shown that there are infinitely many Mersenne primes, then it would follow from Theorem 1 that there are infinitely many even perfect numbers. However, this yields another unanswered and seemingly very difficult question. It is not even known whether there are infinitely many composite Mersenne numbers, but of course there must be infinitely many Mersenne numbers in the set of primes or in the set of composite integers (or infinitely many in each). We can, however, show that if all the Mersenne primes were known, then we should know all the even perfect numbers, for Theorem 1 has a converse, established by Leonhard Euler (1707–1783); we give below a shorter proof than Euler's due to L. E. Dickson.

Theorem 3. *Every even perfect number is of the form*

$$2^{p-1}(2^p - 1) \qquad \textit{where } 2^p - 1 \textit{ is a prime.}$$

Proof. Let N be an even perfect number, so that

$$N = 2^{n-1}r \qquad \text{where } n \geq 2,\ r \text{ is odd and } s(N) = N.$$

Every divisor of N can be written uniquely as a product of a divisor of 2^{n-1} and a divisor of r. Let T denote the sum of all the positive divisors of r (including r); then

$$(1 + 2 + 2^2 + \cdots + 2^{n-1})T$$

is the sum of all the positive divisors of N (including N) and so equals $s(N) + N$. Since $s(N) = N$, we deduce that

$$2^n r = 2N = (1 + 2 + 2^2 + \cdots + 2^{n-1})T = (2^n - 1)T$$

whence

$$T = r + (r/(2^n - 1)). \tag{2}$$

Since T, r are integers, $r/(2^n - 1)$ must be an integer which is, moreover, a divisor of r other than r (for, as $n \geq 2$, $2^n - 1 \geq 3$). If $r \neq 2^n - 1$, then 1, $r/(2^n - 1)$, r are three distinct divisors of r, and so by (2) and the definition of T

$$r + \frac{r}{2^n - 1} = T \geq r + \frac{r}{2^n - 1} + 1 > r + \frac{r}{2^n - 1}$$

which gives a contradiction; hence $r = 2^n - 1$. If r were composite, it would have a divisor d satisfying $1 < d < r$, and so by (2)

$$r + 1 = T \geq r + d + 1 > r + 1,$$

giving a contradiction again; thus r must be prime, whence n itself must be a prime p by Theorem 2. This completes the proof of the theorem.

Let us now state and prove the result on the last digit of an even perfect number, referred to at the end of Section 1.

Theorem 4. *The last digit of an even perfect number is* 6 *or* 8.

Proof. By Theorem 3, an even perfect number N is of the form

$$N = N_p = 2^{p-1}(2^p - 1)$$

where $2^p - 1$ is prime, and by Theorem 2, p must also be prime. The first two perfect numbers 6, 28, corresponding to $p = 2, 3$ respectively, end in 6 or 8 and so we need only consider N_p when $p > 3$. Each prime $p > 3$ is odd and may be written in one of the forms $4m + 1$, $4m + 3$ for some integer $m \geq 1$. Hence

$$N_p = 2^{4m}(2^{4m+1} - 1) = 16^m(2 . 16^m - 1)$$

or

$$N_p = 2^{4m+2}(2^{4m+3} - 1) = 4 . 16^m(8 . 16^m - 1).$$

For each $m \geq 1$, the last digit of 16^m is 6, and hence the numbers

$$2 . 16^m - 1, \qquad 16^m(2 . 16^m - 1), \qquad 8 . 16^m - 1, \qquad 4 . 16^m(8 . 16^m - 1)$$

have last digit 1, 6, 7, 8 respectively. Thus N_p ends in 6 if $p = 4m + 1 \geq 5$ or $p = 2$, and in 8 if $p = 4m + 3 \geq 3$.

3. The search for even perfect numbers

At the end of Section 1, we wrote down the first five even perfect numbers, and we can now justify that statement by applying Theorem 3. The first six primes are $p = 2$, 3, 5, 7, 11, 13 and the corresponding number $2^p - 1$ is prime except for $p = 11$. Hence the first five perfect numbers 6, 28, 496, 8128, 33550336 are found by putting $p = 2, 3, 5, 7, 13$ in turn in the expression $2^{p-1}(2^p - 1)$. In mediaeval times, it would have been quite a feat of arithmetic to establish from scratch the primality of $M_{13} = 2^{13} - 1 = 8191$; if M_{13} were composite, it would have a prime divisor less than $\sqrt{8191} = 90{\cdot}5..$, and so the person doing the calculation would have checked (at least) that none of the 24 primes less than 91 divides 8191.

In 1588, P. A. Cataldi (1548–1626), an Italian mathematician, showed that $M_{17} = 131071$ and $M_{19} = 524287$ are both primes; since $724 < \sqrt{M_{19}} < 725$, to investigate the primality of M_{19}, he first constructed a table of primes less than 750, and then verified that none of the 128 primes less than 725 divides M_{19}, a laborious process! He thus obtained the sixth perfect number,

$$2^{16}(2^{17} - 1) = 8589869056,$$

and the seventh, $2^{18}(2^{19} - 1)$, which is larger still. It was nearly 200 years before the next Mersenne prime M_{31} was verified to be non-composite, for the numbers M_p for $p > 19$ are too large for Cataldi's method to be practical, and new labour-saving devices (described in Section 4) had to be developed. Short-cuts to show that certain M_p are composite were found before this.

After Cataldi, the next stage in the search for even perfect numbers occurred in France during the working life of the amateur and gifted mathematician Pierre de Fermat (1601–1665), a lawyer by profession and a councillor in the local parliament of his native town of Toulouse. Although he did not find any new Mersenne primes, he developed methods to establish the non-primality of several Mersenne numbers (see Section 4), and he played a major role in founding modern number theory. Other prominent French mathematicians alive at that time (when Cardinal Richelieu was in power and France was the political and cultural leader of Europe) were Descartes, Desargues and Pascal. Fermat was in correspondence with them, with Mersenne often acting as intermediary and encouraging them to work on problems he proposed.

In 1644, Mersenne made the assertion (based on earlier incorrect tables) that the first eleven even perfect numbers are $2^{p-1}(2^p - 1)$ for $p = 2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257$, the first seven being known already. The larger primes here give rise to extremely large numbers M_p, much too big for mathematicians then to have been able to determine whether they were prime or composite: for example, M_{127} has 39 digits. Hence it is not surprising that the list contains five errors—two of inclusion and three of omission. It was not until 1947 that the complete list of all primes M_p for $p \leq 257$ was established beyond doubt.

The two correct but unsubstantiated perfect numbers in Mersenne's list correspond to $p = 31$ and 127. In 1772, Euler used methods he had developed (see Section 4) to prove that $M_{31} = 2147483647$ is prime, so that $2^{30}M_{31}$ is perfect. For over 100 years, M_{31} remained the largest integer to have been proved non-composite. Then in 1876, E. Lucas (1842–1891) established by his more complicated methods that M_{127} is prime, and this in turn was the largest known prime for three-quarters of a century.

The first error in Mersenne's list was also found by Lucas in 1876, for his methods showed that M_{67} is composite without, however, giving its factors, which were found by F. N. Cole in 1903; M_{67} is the product of two primes, one with 9 digits and the other with 12. In the 1880's, J. Pervušin, P. Seelhof, and J. Hudelot each showed that M_{61} is prime, an omission from Mersenne's list. Similarly, R. E. Powers and E. Fauquembergue proved in 1911–1914 that M_{89} and M_{107} are prime. The final error, that M_{257} is composite, was confirmed beyond doubt by 1932 by D. H. Lehmer, but its prime factors, the smallest of which is known to exceed 2^{39}, have not yet been found. Hence a corrected list of the primes $p \leq 257$ for which M_p is prime and $2^{p-1}M_p$ is perfect is

$$p = 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127,$$

and M_p is composite for all other primes $p \leq 257$ although its prime factors are not known in all cases yet.

Since the advent of high-speed computers, much more progress has been possible; some new Mersenne primes have been found and some large composite Mersenne numbers factorized. In the twenty years 1951–1971, after much work in mathematics and computing by various people in several countries, but especially in the U.S.A., twelve new Mersenne primes M_p were found, given by

$$p = 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937,$$

and the corresponding numbers $2^{p-1}M_p$ are perfect; there are no other Mersenne primes M_p with $257 < p < 20000$. When $p = 19937$, M_p has some 6002 digits and the corresponding perfect number has 12003 digits.

Some much larger Mersenne numbers M_p are known to be composite, for example M_p with $p = 16188302111$. In Section 4, we shall mention some results that can be used to show that certain M_p, some of them very large, are composite. Such methods were used by D. J. Wheeler in 1953 to disprove a conjecture made by E. Catalan shortly after Lucas had shown in 1876 that M_{127} is prime. Catalan conjectured (on the slender numerical evidence that it holds for $p = 2, 3, 5, 7$) that if $q = M_p$ is prime, then M_q is also prime; however Wheeler showed that M_{8191} is composite, so Catalan's assertion fails for $p = 13$ (when $q = 8191$).

We come now to some exciting new discoveries made during 1978/9 by two young Californian students, Laura Nickel and Curt Noll, when they were less than 20 years old. As announced in some of the British national daily newspapers on 20 November 1978, after 450 hours' work on a computer they showed that, when $p = 21701$, M_p is prime, has 6533 digits and begins and ends as $488\ldots751$. Subsequently† they have found a still larger new Mersenne prime, namely M_p for $p = 23209$. Also Nelson and Slowinsky of the Livermore Laboratories in California are investigating all the primes p satisfying $23209 < p < 50000$, and so far this has led to the discovery that M_p is prime, and so $2^{p-1}M_p$ is perfect, for $p = 44497$. The tests used to establish the Mersenne primes found by computer in the last 30 years are based on a modification by D. H. Lehmer of Lucas' test, (IV) of Section 4 below, and the current investigations utilize a new CRAY 1 computer at Livermore. By these methods the number of known perfect numbers has been increased from 24 in 1971 to 27 now.

Hence the process of finding new Mersenne primes and perfect numbers is a continuing one, contrary to the prediction of Peter Barlow (London, 1811). Referring to Euler's prime M_{31}, he wrote that the corresponding perfect number is '. . . the greatest perfect number known at present, and probably the greatest that ever will be discovered; for as they are merely curious, without being useful, it is not likely that any person will attempt to find one beyond it.'

4. Some tests of primality for Mersenne numbers

The results below will be stated without proof, for the proofs are beyond the scope of this article; the interested reader should consult reference 4, for example. The notation $a|b$ means that the integer a is a divisor of the integer b; if a does not divide b, we write $a\nmid b$.

† The author is grateful to Professor D. H. Lehmer for informing her by letter about these latest discoveries.

In this section we are concerned with the question of how one can show that a large Mersenne number is prime or composite. Fermat in 1640 found our first criterion:

(I) *Any prime q that divides M_p $(p > 2)$ is of the form*

$$q = 2kp + 1, \qquad \text{where } k \text{ is a positive integer.}$$

Examples.

(i) $p = 11$: $q = 22k + 1$ is prime for $k = 1$ and $23 | M_{11} = 23 . 89$, $89 = 22 . 4 + 1$ being prime.

(ii) $p = 23$: $q = 46k + 1$ is prime for $k = 1$ and $47 | M_{23}$.

(iii) $p = 29$: $q = 58k + 1$ is prime for $k = 1, 4$ and $59 \nmid M_{29}$ but $233 | M_{29} = 536870911$.

Thus only two divisions are needed to show that M_{29} is composite. Similarly it can be shown that M_p is composite for $p = 37, 41, 43, 47, 53, 59$ and for some much larger primes p, for example $p = 16035002279$ (for then p and $q = 2p + 1$ are prime and $q | M_p$).

Had Cataldi known the result (I), he would have only had to test whether M_{19} was divisible by any of the six primes of the form $38k + 1$ that are less than 725. However, to prove M_{31} prime, 157 divisions by primes of the form $62k + 1$ would be needed. Euler reduced the number of divisions necessary to 84 by combining (I) with the following result:

(II) *Every divisor of M_p is of one of the forms $8k + 1$, $8k - 1$.*

Thus (I) and (II) together imply that any prime q dividing M_{31} is of one of the forms $248k + 1$, $248k + 63$, and there are 84 such primes less than $\sqrt{M_{31}}$.

Euler also established a result that can be used to show that certain M_p are composite:

(III) *If $p = 4m + 3$ $(m \geq 1)$ is prime and if also $q = 2p + 1$ is prime, then $q | M_p$ and M_p is composite.*

Examples. $p = 4m + 3$ and $q = 2p + 1$ are primes when

$$p = 11, 23, 83, 131, 179, 191, 239, 251, \text{ etc.},$$

and for some much larger values of p, such as

$$p = 16035002279, \qquad 16188302111$$

when M_p has nearly $5 . 10^9$ digits.

It is still not known whether there are infinitely many primes of the form $2p + 1$ with $p = 4m + 3$ and prime.

In conjunction with (I) and (II), one can use the known fact that no two Mersenne numbers have a common factor; hence if r divides M_p, r cannot divide M_q for $q \neq p$. But these criteria are impractical when it comes to determining whether M_p is prime for $p \geq 61$, so Lucas devised another test which he used to show that M_{127} is prime but M_{67} is composite; his test forms the basis for establishing many later results, and for much of the computer work described in Section 3.

The proof of Lucas' criterion is rather more difficult than the proofs of (I)–(III), and even the statement is more complicated. In its original form, Lucas' test applied only to those M_p with $p = 4m + 3$, but he extended it later to all primes $p > 2$, and the result that we state is in the following more general form:

(IV) *Define the integers* $r_1, r_2, \ldots, r_{p-1}$ *by the recurrence relation*

$$r_1 = 4, \qquad r_{n+1} = r_n^2 - 2 \quad \textit{for } n = 1, 2, \ldots, p-2.$$

Then for $p > 2$, M_p *is prime if* $M_p | r_{p-1}$ *and composite otherwise.*†

For larger values of p, there is less work involved in determining whether $M_p | r_{p-1}$ than in applying (I) and (II) combined to the same M_p, but even then Lucas' application of (IV) to show that M_{127} is prime demanded some formidable calculations. That is why the discovery of larger Mersenne primes had to await the advent of the computer. Lucas' criterion can be used to establish the non-primality of a given M_p, but it gives no indication of any factors of M_p. Hence some M_p are known to be composite without any prime factor being found—these have to be determined by other methods. Thus the factors of M_{257} are not yet known, and M_{101} was proved to be composite by 1913, some 50 years before its two prime factors, one of 13 digits and one of 18 digits, were discovered.

The search to find new Mersenne primes, and to answer some of the open questions that we have raised, will continue long after this article has appeared in print, and the 'new' discoveries given here will probably soon be out of date—displaced by something newer. When present knowledge, and perhaps additional and, as yet, unknown techniques, can be combined with the generous use of a new generation of more powerful computers, further progress on finding larger Mersenne primes and factorizing other Mersenne numbers will undoubtedly be made. However some questions raised in this article, such as those asking whether there are infinitely many integers with a specified property (that of being a Mersenne prime, for example), are likely to be answered by theoretical methods. So far these questions have defied all attempts by some distinguished mathematicians to solve them, and now await new ideas—perhaps from the next generation of mathematicians, represented by our younger readers.

References

1. L. E. Dickson, *History of the Theory of Numbers*, Volume 1, Chapter 1 (Carnegie Institution, Washington, 1919).
2. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 3rd edn (Clarendon Press, Oxford, 1954).
3. O. Ore, *Number Theory and its History* (McGraw-Hill, New York, 1948).
4. D. Shanks, *Solved and Unsolved Problems in Number Theory*, Volume 1 (Spartan, Washington, 1962).

† In his article in volume 10 of *Mathematical Spectrum*, John Strange referred on p. 22 to this test of Lucas, and to the result that M_p is prime for $p = 11213$.

Dear Editor,

Pythagorean numbers

I recently noticed that, for any positive whole value of n, the following statements can be used to find six Pythagorean numbers:

$$a = n(n+2), \qquad b = \frac{1}{a}(en), \qquad c = b + n,$$
$$d = a + n + 1, \qquad e = \tfrac{1}{2}(d^2 - 1), \qquad f = e + 1.$$

For example, when $n = 1$, the following values are obtained:

$$a = 3, \quad b = 4, \quad c = 5 \quad \text{and} \quad a^2 + b^2 = c^2,$$
$$d = 5, \quad e = 12, \quad f = 13 \quad \text{and} \quad d^2 + e^2 = f^2;$$

when $n = 2$,

$$a = 8, \quad b = 15, \quad c = 17 \quad \text{and} \quad a^2 + b^2 = c^2,$$
$$d = 11, \quad e = 60, \quad f = 61 \quad \text{and} \quad d^2 + e^2 = f^2.$$

I am primarily a biologist, not a mathematician, and therefore I do not know if these connections have been observed before. I should be interested to know whether they have or not. I should also be interested to know whether these equations can be used to find all Pythagorean numbers which cannot be found by using

$$m, \qquad \tfrac{1}{2}(m^2 - 1), \qquad \tfrac{1}{2}(m^2 + 1),$$

or whether there are others.

Yours sincerely,

PAUL HEATH
(9 Hazeldene Meads, Brighton BN1 5LR)

Dear Editor,

Tests for divisibility

I have seen tests for the divisibility of integers by 7, 11, 13, 17 and 19. These tests can be extended to all numbers less than 100 which are not divisible by 2 or 5.

Suppose that the integer $10a + b$, where $0 \leq b \leq 9$, is to be tested for divisibility by the integers $10v + 1$, $10v + 3$, $10v + 7$, $10v + 9$, where $0 \leq v \leq 9$. We first note that

$$\begin{aligned}
(10a + b) - (10v + 1)b &= 10(a - vb),\\
(10a + b) + 3(10v + 3)b &= 10(a + (1 + 3v)b),\\
(10a + b) - 3(10v + 7)b &= 10(a - (2 + 3v)b),\\
(10a + b) + (10v + 9)b &= 10(a + (1 + v)b).
\end{aligned}$$

Since each of the pairs $(10, 10v + 1)$, $(10, 10v + 3)$, $(10, 10v + 7)$, $(10, 10v + 9)$ has highest common factor 1, it follows that

(i) $10v + 1$ divides $10a + b$ if and only if $10v + 1$ divides $a - vb$,

(ii) $10v + 3$ divides $10a + b$ if and only if $10v + 3$ divides $a + (1 + 3v)b$,
(iii) $10v + 7$ divides $10a + b$ if and only if $10v + 7$ divides $a - (2 + 3v)b$,
(iv) $10v + 9$ divides $10a + b$ if and only if $10v + 9$ divides $a + (1 + v)b$.

When $10a + b$ is fairly large, these statements reduce the four divisibility problems to rather simpler ones. Repeated application of the method finally yields a problem which can be solved at sight. An example will help to clarify the procedure.

We test 792814732 for divisibility by 37. Here $a = 79281473$, $b = 2$, $v = 3$. Reference to (iii) shows that the question is equivalent to the divisibility of

$$a - (2 + 3v)b = a - 11b = 79281473 - 22 = 79281451$$

by 37. Next, $a = 7928145$ and $b = 1$. Again, by (iii) we may instead consider the divisibility of

$$7928145 - 11 = 7928134.$$

Proceeding in this way we finally see that the problem reduces to the divisibility of 34 by 37. So 792814732 is not divisible by 37. The detailed working can be set out as follows.

$$\begin{array}{r}
792814732 \\
22- \\
\hline
79281451 \\
11- \\
\hline
7928134 \\
44- \\
\hline
792769 \\
99- \\
\hline
79175 \\
77- \\
\hline
7840 \\
0- \\
\hline
784 \\
44- \\
\hline
34
\end{array}$$

Yours sincerely,
A. J. GRANVILLE
(Clifton College, Bristol)

Dear Editor,

A Bayesian look at the jury system

I have read the article on the jury system in Volume 11 Number 2 with much interest (although I am not clear whether it was written with British or Australian courts in mind).

The estimated value of 0·8 to 0·99 for $P(G^+)$ is most unlikely to be true for the U.K. I believe that 0·9 is about right as the proportion of those *indicted* who are in fact guilty, but not as the proportion of those subject to a *jury verdict* who are in fact guilty. The following figures are based only on recollection, but I think they are fairly reliable: 50 % of those indicted plead 'guilty'; 15 % change their 'not guilty' pleas to 'guilty' during the trial; 25 % are discharged on the judge's direction. This leaves only 10 % to be actually tried by jury, and in the U.K. about half of these (5 % of those indicted) are acquitted by the jury. (If our '90 % guilty' hypothesis is right, the other 5 % might have been discharged on the judge's direction.) Since we are discussing jury verdicts, the appropriate value for $P(G^+)$ is more like 0·5 than 0·9.

According to my calculations, this value for $P(G^+)$ changes $P(G^+|J^-)$ from 0·3125 to 0·048, and $P(G^-|J^+)$ from 0·0012 to 0·0104. It seems quite plausible that 1/20 of those acquitted are in fact guilty, and 1/100 of those convicted are innocent.

I am also concerned at the inferences in the article about the situation where $P(G^+) \rightarrow 1$. If we use the same values for $P(J^-|G^+)$ and $P(J^-|G^-)$, it is assumed that the juries' reactions will not respond to the more convincing prosecution evidence which would appear if only guilty people were indicted; should we not expect the former probability to decrease and the latter to increase?

Yours sincerely,
NEVILLE K. UPTON
(City of Birmingham Polytechnic)

Dear Editor,

A Bayesian look at the jury system

Mr Upton raises some important practical areas of law procedure which I did not really have in mind at the time of writing the article; questions of discharge on the judge's direction, changing pleas, etc., were not considered. I have no practical data on the proportions involved but, if his recollection is correct, a value for $P(G^+)$ below the suggested 0·9 would be reasonable.

The change in value of $P(G^+|J^-)$ to 0·048 using his lower value of 0·5 for $P(G^+)$ would probably be a welcome result, but the increase in $P(G^-|J^+)$ from 1 in 1000 to 1 in 100 might cause great alarm among those people concerned with such possible miscarriages of justice.

Concerning values for $P(J^-|G^+)$ and $P(J^-|G^-)$ and the possibly 'more convincing prosecution evidence' we have no knowledge, but it would be of interest to consider different values and their implications anyway.

Mr Upton correctly dispels the apparent paradox of $P(G^+|J^-) \rightarrow 1$ as $P(G^+) \rightarrow 1$. This was one of my reasons for writing the article, since I thought some readers might like to analyse the situation in this way (and indeed several have done so).

The very nature of the questions to which we seek answers (namely, $P(G^+|J^-)$ and $P(G^-|J^+)$) are such that more definitive conclusions are highly unlikely. Perhaps the use of lie-detectors and/or other devices with defendants after the jury outcomes are known might provide some direct measures of the quantities of interest, but the moral and practical implications of such an approach may be insurmountable.

Yours sincerely,
J. W. HILLE
(State College of Victoria, Frankston)

Problems and Solutions

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

Problems

12.7. (Submitted by Paul Brennan, Cambridge) Let a, b be positive real numbers and let n be an integer with $n > 2$. Show that

$$(a^n + b^n)^2 < (a^2 + b^2)^n.$$

12.8. (Submitted by K. H. Yim, University of Liverpool) Let k be an integer with $k \geq 2$ and let $n_1, \ldots, n_r$ $(r \geq 2)$ be integers greater than or equal to k whose sum is denoted by n. Show that

$$\binom{n_1}{k} + \binom{n_2}{k} + \cdots + \binom{n_r}{k} < \binom{n-(r-1)(k-1)}{k},$$

where ( ) denotes the binomial coefficient.

12.9. A particle is slightly displaced from the top of a smooth, fixed sphere. Determine where the particle leaves the sphere.

Solutions to Problems in Volume 12, Number 1

12.1. If $n_1, \ldots, n_r$ are integers greater than one with sum n, show that

$$\binom{n_1}{2} + \binom{n_2}{2} + \cdots + \binom{n_r}{2} \leq \binom{n-r+1}{2},$$

with strict inequality when $r > 1$, where ( ) denotes the binomial coefficient.

Solution

We use induction on r. The result is true when $r = 1$ with an equality. Now let $r > 1$ and assume the result for $r - 1$. Then

$$\binom{n_1}{2} + \binom{n_2}{2} + \cdots + \binom{n_{r-1}}{2} \leq \binom{n - n_r - (r-1) + 1}{2}.$$

Thus

$$\binom{n_1}{2} + \binom{n_2}{2} + \cdots + \binom{n_r}{2} \leq \binom{n - n_r - r + 2}{2} + \binom{n_r}{2}$$

$$= \tfrac{1}{2}[(n - n_r - r + 2)(n - n_r - r + 1) + n_r(n_r - 1)]$$

$$= \binom{n-r+1}{2} - (n_r - 1)(n - r + 1 - n_r).$$

Now $n_r > 1$ and $n - n_r > r - 1$, so

$$(n_r - 1)(n - r + 1 - n_r) > 0.$$

The result follows for r with a strict inequality.

12.2. For which real numbers x is

$$\{x + \sqrt{(x^2 + 1)}\}^{1/3} + \{x - \sqrt{(x^2 + 1)}\}^{1/3}$$

an integer?

Solution

Put

$$y = \{x + \sqrt{(x^2 + 1)}\}^{1/3} + \{x - \sqrt{(x^2 + 1)}\}^{1/3}.$$

Then

$$y^3 + 3y = 2x.$$

Suppose that y is an integer. Then

$$x = \tfrac{1}{2}(y^3 + 3y).$$

Conversely, suppose that $x = (n^3 + 3n)/2$, where n is an integer. Then

$$\begin{aligned} &\tfrac{1}{2}(n^3 + 3n) = \tfrac{1}{2}(y^3 + 3y) \\ &\Rightarrow y^3 - n^3 + 3(y - n) = 0 \\ &\Rightarrow (y - n)(y^2 + yn + n^2 + 3) = 0 \\ &\Rightarrow y = n \end{aligned}$$

because the quadratic equation $y^2 + yn + n^2 + 3 = 0$ in y does not have real roots. Hence the real numbers x for which y is an integer are given by $x = (n^3 + 3n)/2$, where n is an integer.

Also solved by John Collinson (Open University) and Michael Dumbrell (Corpus Christi College, Oxford).

12.3. A circle is tangent internally to the circumcircle of triangle ABC and also to the sides AB, AC at P, Q respectively. Prove that the midpoint of segment PQ is the centre of the incircle of triangle ABC.

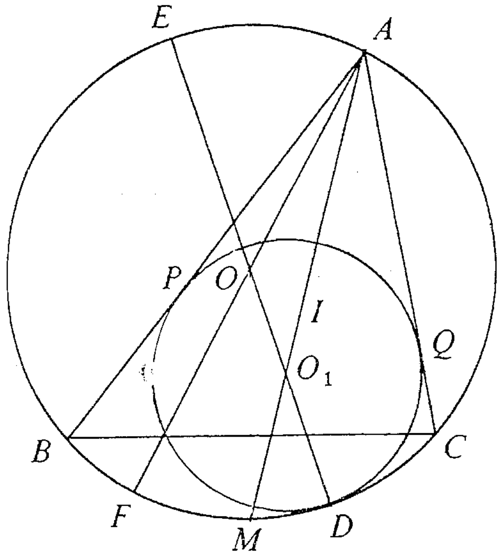
Solution 1 (by R. C. Lyness)

Denote by O, R the centre and radius of the circumcircle of triangle ABC, by O', ρ the centre and radius of the circle tangent to the circumcircle and to AB, AC, and by I, r the centre and radius of the incircle of triangle ABC. Denote by M the point in which the bisector of angle BAC again meets the circumcircle. Both O_1 and I lie on AM. Denote by D the point where the two circles are tangent, and let DOE be the diameter of the circumcircle. Then

$$\begin{aligned} AO_1 \,.\, O_1M &= EO_1 \,.\, O_1D \\ &= (R + OO_1)(R - OO_1) \\ &= R^2 - OO_1^2 \\ &= R^2 - (R - \rho)^2 \\ &= \rho(Rr - \rho). \end{aligned}$$

From the right-angled triangle $AO_1P, AO_1 = \rho \operatorname{cosec}(A/2)$; and from the right-angled triangle AFM, where AOF is the diameter of the circumcircle,

$$\begin{aligned} AM &= 2R\cos \angle FAM \\ &= 2R\cos(A/2 - (\pi/2 - C)) \\ &= 2R\cos\frac{C - B}{2}. \end{aligned}$$



Hence

$$\left(\rho\,\mathrm{cosec}\frac{A}{2}\right)\left(2R\cos\frac{C-B}{2}-\rho\,\mathrm{cosec}\frac{A}{2}\right)=\rho(2R-\rho)$$

$$2R\left(\cos\frac{C-B}{2}-\sin\frac{A}{2}\right)=\rho\left(\mathrm{cosec}\frac{A}{2}-\sin\frac{A}{2}\right)$$

$$2R\left(\cos\frac{C-B}{2}-\cos\frac{C+B}{2}\right)=\rho\,\mathrm{cosec}\,\frac{A}{2}\cos^2\frac{A}{2}$$

$$4R\sin\frac{B}{2}\sin\frac{C}{2}=\rho\,\mathrm{cosec}\,\frac{A}{2}\cos^2\frac{A}{2}.$$

But

$$r=4R\sin\frac{A}{2}\sin\frac{B}{2}\sin\frac{C}{2},$$

so that

$$r=\rho\cos^2\frac{A}{2}.$$

From triangle PAO_1, $PA=\rho\cot(A/2)$. Denote by I' the midpoint of PQ. Then

$$PI'=PA\sin(A/2)=\rho\cos(A/2)=QI',$$

and the distance of I' from AB, AC is $\rho\cos^2(A/2)$, the same as the distance of I from AB, AC. Since I, I' both lie on AM, it follows that $I'=I$, as required.

Solution 2 (by R. C. Lyness)

The line $B'C'$ on the diagram is drawn tangentially to the incircle of triangle ABC in such a way that $\angle AB'C'=C$. Then $\angle C'B'B=A+B$, so that $\angle IB'B=(A+B)/2$, where I is the centre of the incircle. Since $\angle B'AI=A/2$, then $\angle AIB'=B/2$. Hence triangles $AB'I$ and AIB are similar, so that

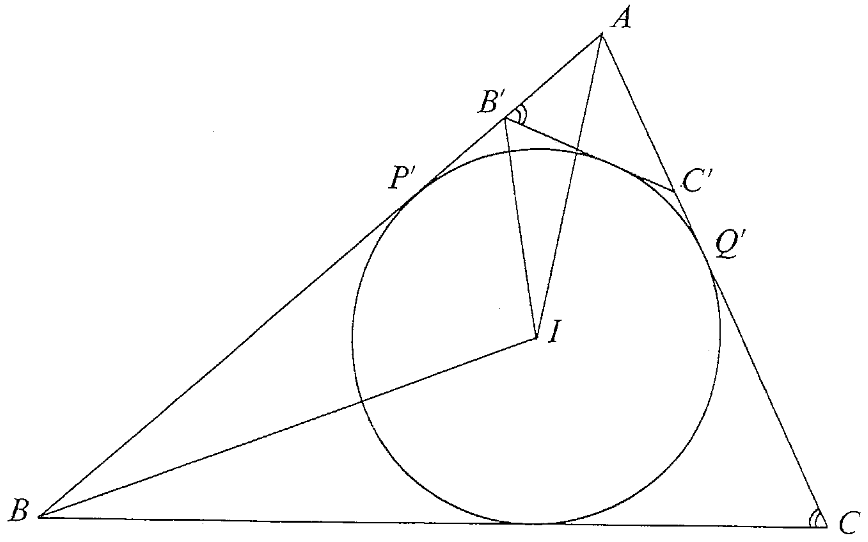
$$\frac{AI}{AB'}=\frac{AB}{AI}$$

$$\Rightarrow AI^2=AB\,.\,AB'.$$

Similarly,

$$AI^2 = AC\,.\,AC'.$$

We now invert with respect to the circle centre A radius AI. The straight line $B'C'$ inverts to the circle through A, B, C, i.e. to the circumcircle of triangle ABC. Hence the incircle inverts to a circle touching the circumcircle and also touching AB, AC, i.e. to the circle given in the question. Denote by P', Q' the points on AB, AC where the incircle touches these lines. Then A, P', I, Q' lie on a circle, which inverts to a straight line through P, I, Q. Since AI bisects $\angle BAC$, I will be the midpoint of PQ, as required.



Book Reviews

Elementary Number Theory. By UNDERWOOD DUDLEY, W. H. Freeman and Company Limited, San Francisco, 1978. Pp. ix + 249. £8·70 (second edition).

'Natural number' is surely the original mathematical concept and the most fundamental. Speculations about the nature and properties of whole numbers doubtless constitute the oldest form of mathematical thought. The ancient civilizations of Sumeria, Babylon and Egypt as well as those of China and India all reveal a considerable knowledge of the properties of the system of natural numbers, but it was the Greeks who first studied arithmetic, properly so called. Indeed Euclid's *Elements* is as much a text-book on the theory of numbers as on geometry and was the first systematic presentation of the subject.

Arithmetic deals primarily with questions about the natural numbers, most of which are easy to state and to comprehend; for example, the statement, known as Goldbach's Conjecture, that every even number greater than 4 is the sum of two odd prime numbers. Yet that conjecture is still unproved, in spite of the efforts of many of the greatest mathematicians using the formidable tools of modern analysis (calculus), as are many other deceptively simple statements. Indeed it is true to say that a great deal of modern algebra and analysis was created primarily in order to solve such problems, which therefore, like geometry and other branches of 'applied' mathematics, is one of the great sources of developments in mathematics.

Because it is a natural source of stimulating problems, arithmetic has for centuries not only fascinated many gifted amateurs, but it has also been the favourite subject of many of the most powerfully equipped professional mathematicians, foremost amongst whom was Gauss, who considered his work in the theory of numbers to be his greatest achievement.

One of the educationally exciting aspects of the subject is that this vital source of research problems is also accessible to pupils at school, for whom it offers a stimulating encounter with serious mathematics. The book under review affords an excellent introduction to the subject, ideally suited to sixth-form reading and with a wealth of problems to challenge and entice the reader.

The early chapters develop the basic concepts of divisibility, prime numbers and congruences and these are used to prove, for example, that a number p is a prime if and only if $(p-1)!+1$ is divisible by p (Wilson's Theorem).

One of the most remarkable theorems in the subject is the law of quadratic reciprocity. A number $a \neq 0$ is called a quadratic residue for a prime number p if there is a number x such that $x^2 - a$ is divisible by p (that is, if a has a square root in the finite 'modular' or 'clock' arithmetic of remainders on division by p). The French mathematician Legendre introduced a symbol denoted by (a/p) which is $+1$ if there is such an x and is -1 if not. Gauss computed the value of the symbol (p/q) for over 10 000 values of the primes p and q (for example $(5/17) = -1$ and $(3/13) = +1$) and, at the age of 19, finally proved (as had been guessed by Euler and Legendre) that there is a relation between (p/q) and (q/p); namely that $(p/q)(q/p) = (-1)^{(p-1)(q-1)/4}$.

Gauss's first proof is still a source of inspiration to research mathematicians and in all he gave seven proofs. Since then, many more have been found (one, by Cauchy, using the theory of heat conduction) and the attempts to understand what lies behind it and to generalize it (for example, to the case of cubes instead of squares) have led to major advances in mathematics. The proof reproduced in this book is actually Gauss's third.

There follows an interlude on number bases and decimal representation, which could be read with profit by all intending teachers.

The book concludes with two major themes: Diophantine equations and the distribution of prime numbers.

The phrase 'Diophantine equation' is derived from Diophantus of Alexandria who was the first to obtain a formula for all the integral solutions of the equation $x^2 + y^2 = z^2$. This book devotes a chapter, on 'Pythagorean Triangles', to that equation and then goes on to consider equations, like $x^3 + y^3 = z^3$, which are generalizations of the Pythagorean equation. The 'amateur' Fermat (1601–1665: his profession was the law) asserted that for, $n > 2$, the equation $x^n + y^n = z^n$ has no integral solutions except the trivial ones like $x = z, y = O$. He wrote his 'theorem' in the margin of his copy of Diophantus' *Arithmetic* and claimed that he had a proof. No one has succeeded in finding a proof since that time, in spite of the efforts of many of the greatest mathematicians; so perhaps he was mistaken, though the romantically inclined may prefer to think of it as a mathematical 'lost chord'. This book uses Fermat's method of infinite descent to deal with the case $n = 4$.

Other problems of a similar kind, for example the solutions of the equation $x^2 - Ny^2 = 1$ and that of representing a number as a sum of squares are also discussed. (On p. 147 the author asserts that the largest integer that needs a sum of eight cubes to represent it is 454; that is not correct, 8042 requires eight cubes, though seven cubes suffice from 8043 onwards.)

There are many natural questions one can ask about the sequence of prime numbers such as: is there a formula for the nth prime number; are there infinitely many primes; are there infinitely many primes p such that $p + 2$ is also prime (for example 101 and 103 or $76 \cdot 3^{139} - 1$ and $76 \cdot 3^{139} + 1$); how many primes are there less than a given number x? The answer to the first is that 'it depends on what you mean by a formula'; the answer to the second is 'yes'; the answer to the third is almost certainly 'yes', though no proof has yet been found; finally $x/\ln x$ ($\ln x = \log_e x$) is a good approximation to the number of primes, but how good an approximation it is remains a tantalizing question. These and related questions are lucidly presented and discussed.

The book is well produced, there are few errors (the index gives Godfrey Harold Hardy's names as Geoffrey Henry) and the author has succeeded splendidly in encouraging the reader to learn the subject by doing problems, to which end he has provided an admirable collection.

I recommend the book strongly and hope that every school library will purchase a copy.

University of Durham J. V. ARMITAGE

Algebra. By SAUNDERS MACLANE AND GARRETT BIRKHOFF. Collier Macmillan, Second Edition, 1979. Pp. xvi + 586. £14·25.

The second edition of this classic book has a number of alterations and improvements. There is a new chapter on Galois theory in place of one on affine geometry, but perhaps more important the original first chapter on category theory is now Chapter 4 and the book thereby becomes more readable than before.

University of Durham H. NEILL

Finite Mathematics with Applications. By A. W. GOODMAN and J. S. RATTI. Collier Macmillan, London, 1979. Pp. xiv + 584. £11·95 (third edition).

This beautifully produced and attractive American book contains many good examples for fifth- and sixth-form teachers in Britain, but fits uneasily into existing British courses. It would be a useful school library book, but it may be considered too expensive to be good value.

University of Durham H. NEILL

The Cambridge Elementary Mathematical Tables. By J. C. P. MILLER AND F. C. POWELL. Cambridge University Press, Second Edition, 1979. Pp. 48. £0·50.

This second edition has the same number of pages as the well-known first edition and contains one or two more statistical tables instead of the cube tables.

University of Durham H. NEILL

INDEX

Volume 10 (1977/78) Volume 11 (1978/79) Volume 12 (1979/80)

Contents

ISSN 0025–5653

PRICES (*postage included*)

Prices for Volume 12 (*Issues Nos.* 1, 2, *and* 3):

Subscribers in Britain and Europe: £1·60
Subscribers overseas: £3·20 (US$6.40; $A5.50)
(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.

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