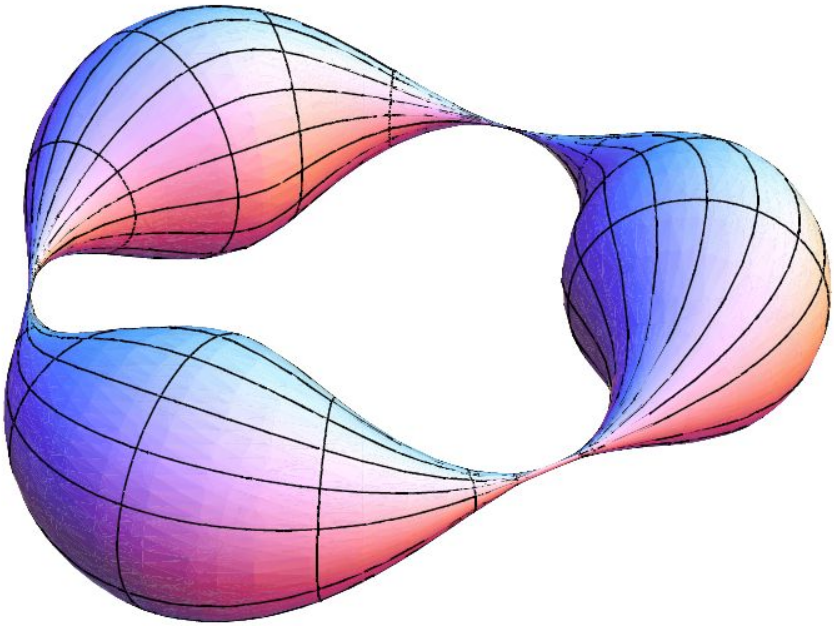
*

ISSN 1350-8539

M500

M500 176



The M500 Society and Officers

The M500 Society is a mathematical society for students, staff and friends of the Open University. By publishing M500 and 'MOUTHS', and by organizing residential weekends, the Society aims to promote a better understanding of mathematics, its applications and its teaching.

The magazine M500 is published by the M500 Society six times a year. It provides a forum for its readers' mathematical interests. Neither the editors nor the Open University necessarily agree with the contents.

MOUTHS is 'Mathematics Open University Telephone Help Scheme', a directory of M500 members who are willing to provide mathematical assistance to other members.

The September Weekend is a residential Friday to Sunday event held each September for revision and exam preparation. Details available from March onwards. Send SAE to Jeremy Humphries, below.

The Winter Weekend is a residential Friday to Sunday event held each January for mathematical recreation. Send SAE for details to Norma Rosier, below.

Advice to authors. We welcome contributions to M500 on virtually anything related to mathematics and at any level from trivia to serious research. Please send material for publication to Tony Forbes, above. We prefer an informal style and we usually edit articles for clarity and mathematical presentation. If you use a computer, please also send the file on a PC diskette or via e-mail. Camera-ready copy can be accepted if it follows the general format of the magazine.

Cryptography – keyless and deniable encryption

John Bull

Cryptography is a subject backed by thousands of papers, hundreds of books, and hundreds of research programmes. This note presents just a few results of topical interest, dispels a myth, and offers an unsolved problem.

Cryptography is used to hide a secret message in transmission. Alice encrypts a message such as 'He is a spy', sends it to Bob, and Bob decrypts and reads it. Normally the encryption and decryption algorithms are public but Alice and Bob hold secret encryption and decryption keys. Without the decryption key, anyone intercepting the message would be unable read it.

This discussion assumes asymmetric encryption and decryption algorithms with symmetric keys, known as private key cryptography. In this cryptographic system, Alice and Bob know the same secret key, k. Public key cryptography uses symmetric algorithms with asymmetric but related keys, k^+ and k^-.

A character of text is encoded as a number in a computer byte (8 bits), so the message above would be the 22 hexadecimal numbers '48 65 20 69 73 20 61 20 73 70 79'. Each byte would be encrypted by a function $c = E(p, k)$ and be decrypted by a function $p = D(c, k)$, where E and D operate in a finite arithmetic field of 0 to 255; that is, all the arithmetic is modulo 256.

For example, where all numbers are expressed in hexadecimal arithmetic, we may have $c = E(p, k) = pk + 25 \bmod 100$. With $k = 11$, and thus $c = 11p + 25 \bmod 100$, the encrypted message would be 'ED DA 45 1E C8 45 96 45 C8 95 2E'. The decryption function would then be $p = D(c, k) = \mathrm{F1}c + 2\mathrm{B} \bmod 100$. This is a very poor cipher and would be broken by a cryptanalyst, with little difficulty. But it illustrates the basic principle.

In practice, a function would operate on an array of bytes in the input stream, so that an encryption of one byte would depend on the values of other bytes. Also a function would be applied many times in a number of 'rounds', so that the final encrypted output would be a thorough confusion of all the input data. The encryption and decryption functions can be extremely complex provided they are conservative; that is, provided no information (in the Shannon theory sense) is added or lost when applying logic or arithmetic functions. Further discussion of all of this can be found in [6].

Cryptography is built around the idea of a one-way function; that is, a

function that is very easy to compute one way, but where it would not be practically feasible to compute the inverse. To be more precise, a function f is required such that:

- Given y, it is not feasible to compute x such that $y = f(x)$
- Given x, it is not feasible to compute $v \neq x$ such that $f(v) = f(x)$
- It is not computationally feasible to find a pair of numbers (v, x) such that $f(v) = f(x)$.

A function that thoroughly scrambles an array of data, being more a computational function than a mathematical one, is known as a one-way hash function, $H(x)$. If such a function is used it leads to two further requirements:

- $H(array)$ can be applied to a block of any size
- $H(array)$ produces a fixed length output.

A conservative logic operator frequently used in cryptography is the 'exclusive or', which will be represented by the symbol $\oplus$. This combines bits such that

$$0 \oplus 0 \to 0, \quad 0 \oplus 1 \to 1, \quad 1 \oplus 0 \to 1, \quad 1 \oplus 1 \to 0.$$

This operator is important because it is conservative and reversible; that is, for all x, y and z,

$$x \oplus y \to z \;\Rightarrow\; y \oplus x \to z,\; y \oplus z \to x,\; z \oplus y \to x,\; x \oplus z \to y,\; z \oplus x \to y.$$

Bearing in mind that functions in cryptography apply in finite arithmetic, where a number of bytes taken together determine the size of a finite field, there are two number theoretic problems that are commonly used as one-way functions: the factoring problem, and the discrete logarithm problem. In brief, the two problems are, respectively:

1. It is easy to compute the product of two large primes, but given a number that is known to be such a product it is not computationally feasible to find the factors.

2. In a function $y = g^x \bmod n$, where g and n have certain properties, and x and n are large, it is easy to compute y given x, but not computationally feasible to compute x given y.

For a source of one-way function in this article we will use the discrete logarithm. For a given prime p in the equation $y = g^x \bmod p$, some values of g, known as primitive roots of p, will, for each value of x from 1 to $p-1$, generate distinct values for y between 1 and $p-1$. For example, $g = 2$, 3,

10, 13, 14, 15 are all primitive roots of $p = 19$. In this example, primitive roots $g = 3$ and $g = 10$ will generate values for $g^x \bmod p$ as follows:

$x =$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$3^x =$	3	9	8	5	15	7	2	6	18	16	10	11	14	4	12	17	13	1
$10^x =$	10	5	12	6	3	11	15	17	18	9	14	7	13	16	8	4	2	1

If we now consider the equation $y = g^x \bmod p$, we can see that provided g is a primitive root of p, each x will arise from a different value of y. Borrowing the concept of logarithm from the familiar, non-discrete arithmetic, we have $x = \log(\text{base } g, \bmod p)y$. This is normally written as $x = \text{ind}_{g,p}y$, where ind stands for index. Other analogous results and operations also follow, such as $\text{ind}_{g,p}1 = 0$, $\text{ind}_{g,p}g = 1$, $\text{ind}_{g,p}xz = \text{ind}_{g,p}x + \text{ind}_{g,p}z$.

Given g, x and p, it is relatively easy to compute $y = g^x \bmod p$ [1], pages 78–79. Given g, y and p, the time it would take to compute x using the asymptotically fastest known algorithm would be of the order of $\exp(((\log p)^{1/3} \log\log p)^{2/3})$ [2]. For a large prime p (of a magnitude greater than, say, 2^{1024}) this would be too long a time to be feasible.

Once values of g and p are chosen they can be made public and adopted in standards. Hence there has been much debate about which values of g and p to choose, and various additional criteria are imposed. For example, $(p-1)/2$ should also be a prime. Choices can be found in the literature, particularly in standards for public key cryptography [3].

It isn't strictly necessary for n in $y = g^x \bmod n$ to be prime, since if g is a primitive root of n, all values of x from 1 to $\phi(n)$ will generate distinct values for y. For each integer $n \geq 1$, $\phi(n)$ is defined as the number of positive integers not exceeding n which are relatively prime to n (M381, Unit 5, Section 3, pp 17–24). But using a composite n rather than a prime does not appear to offer any advantage in this application.

Suppose we have correspondents Alice and Bob who each hold a secret key, k, and Alice wishes to send a secret message, m, to Bob. Apparently a simple way would be to 'exclusive or' the message with the key. So Alice forms $c = m \oplus k$, sends c to Bob, and Bob decrypts the message using $m = c \oplus k$. Without knowing k, an eavesdropper would be unable to find m. As an isolated transaction this would be secure, but if the same key k were used over and over again for different messages, m_1, m_2, m_3, etc., it would not be too difficult for a cryptanalyst to break this cipher.

This protocol might be strengthened by adding some randomness into each message. For example, suppose Alice generates a random number r, and sends $c = m \oplus k \oplus r$ to Bob. This would make the encryption function

different for each transaction. Unfortunately, in order for Bob to decrypt the message, he also has to know the random number r, and the only way he can know this is if Alice sends it to him. It would then also be available to an eavesdropper. A solution is to embed r in a one-way function so that even if an eavesdropper knows r, it would be of no value to him without also knowing k.

Alice generates a random number r and encrypts the message as $c = m \oplus g^{k \oplus r} \bmod p$. Alice sends both c and r to Bob. Bob deciphers the message using $m = c \oplus g^{k \oplus r} \bmod p$. This protocol is secure, although it is open to replay whereby an attacker re-uses c and r some time later to fool Bob (assuming the attacker knows what Bob might do with the message). There are many further issues to be dealt with before this theory can be turned into a practical implementation, such as matching fields sizes, and fragmenting and padding messages to suit, but assuming these are handled properly the encryption protocol is sound and secure.

The disadvantage of this protocol, compared with a straightforward commercial encryption package, is that the amount of data that needs to be sent in the form of c and r is twice the size of the original message. However, this is not such a high price to pay for strong encryption now that communications bandwidth is becoming less of a concern. Computing the one-way function needs a fair amount of processing power, even as a forward computation, and this is also of concern, but becoming less so as machines become more powerful.

A number of secure hash functions are freely available and these can also be used to build an encryption protocol. Unlike commercial encryption products, many hash functions are not protected by patent or licence. The best is the Secure Hash Algorithm [3, 4 pp 442–445]. The one-way function given above can be used in an encryption protocol without keys. In this case there are three message exchanges:

- Alice chooses an integer $a < p$ that is relatively prime to $p - 1$ and sends $c = m^a \bmod p$ to Bob.
- Bob chooses an integer $b < p$ that is relatively prime to $p - 1$ and sends $d = c^b \bmod p$ to Alice.
- Alice computes the integer u such that $au = 1 \bmod p - 1$ and sends $e = d^u \bmod p$ to Bob.
- Bob computes the integer v such that $bv = 1 \bmod p - 1$ and computes $f = e^v \bmod p$.
- The result f is equal to the original message m.

We need to prove that the arithmetic works and, to turn this into a practical proposition, that integers u and v can easily be found.

The integers u and v are well defined, since a and b were taken to be coprime with $p-1$, and hence are just the inverses of a and b respectively in the integer group from 1 to $p-1$. Integers u and v can be found using Euclid's algorithm (M381, Unit 1, Section 5.3, pp 31–32). Euclid's algorithm will produce i and j such that $ia+j(p-1)=1$, from which i can be selected to be u. Similarly, a value for v can be found.

We now need to prove that f is in fact the original message m.

• By definitions, $f = e^v \bmod p = (d^u)^v \bmod p = d^u v \bmod p = c^b uv \bmod p = c^{(bv)u} \bmod p$.

• But $bv = 1 \bmod p-1$, so that (for some t), $c^b v \bmod p = c^{1+t(p-1)} \bmod p = c^{(p-1)t} c \bmod p$.

• By Fermat's Little Theorem (M381, Unit 4, Section 1, pp 5–8), $c^{p-1} = 1 \mod p$. Thus we have $f = c^u \bmod p$.

• But $c = m^a \bmod p$, so that $f = m^a u \bmod p$.

• Again, since $au = 1 \bmod p-1$, Fermat's Little Theorem gives the desired result, $f = m$.

This protocol suffers from the disadvantage of a three fold increase in time, since there are three message exchanges. The cryptographic community have been aware of the protocol for many years (originally attributed to Adi Shamir), but probably because of performance doubts it was never published. With modern communications, compared to 25 years ago, this is not such a serious drawback. The novelty is that no key is involved, so it would be impossible for anyone (including national security or law enforcement agencies) to demand that a key be handed over to enable them to decipher encrypted communications.

A related problem is that of 'deniable encryption'. Messages are encrypted such that decryption with one key will produce the truly secret message, but decryption with a second key will produce an innocuous message. If anyone who intercepts the ciphertext demands that a key be handed over, it would be possible to hand over the key to produce the harmless message and deny that any other message were encrypted. It is known that such a proposition is feasible [5] but so far no-one has devised a suitable encryption function nor practical proposal. This is an unsolved problem. Just as public key cryptography was known to be feasible long before Diffie-Hellman, and Rivest, Shamir and Alderman proposed functions and practical implementations, no doubt someone will discover a way to achieve deniable encryption

in a few years.

The above theory illustrates that anyone willing to sacrifice communications bandwidth or performance could engineer strong encryption with little difficulty. Furthermore they could do it in a way that would confound currently proposed Investigatory Powers legislation. Given rapid advances in theory, it should not be too long before deniable encryption is a practical proposition, and when this happens it will lead to encryption products that will make present Investigatory Powers obsolete.

References

[1] R. J. B. T. Allenby and E. J. Redfern, *Introduction to Number Theory with Computing*, Edward Arnold, ISBN 0–7131–3661–8.

[2] T . Beth, M. Frisch and G. Simmons, *Public Key Cryptography: State of the Art and Future Directions*, Springer, New York, 1991.

[3] 'Digital Signature Standard', *National Institute for Standards and Technology*, NIST FIPS PUB 186, US Department of Commerce, May 1994.

[4] Bruce Schneier, *Applied Cryptography*, Second edition, ISBN 0–471–12845–7.

[5] RanCanetti, CynthiaDwork, MoniNaor and RafailOstrovsky, 'Deniable encryption', Research paper supported by BSF grant 32–00032.

[6] John Bull, 'Recreational Cryptography'. Unpublished draft available to M500 members. Fifteen pages including program examples.

Solution 174.5 – Root 11

Consider the fractional part f and the integer part I of the number $(\sqrt{11}+3)^{2n+1}$, where n is a positive integer. Prove:

i. the fractional part is given by$f = (\sqrt{11}-3)^{2n+1}$;

ii. the integer part of $(\sqrt{11}+3)^{2n+1}$, I, is divisible by 2^{n+1};

iii. that $f = \frac{1}{2}\left(\sqrt{I^2+2^{2n+3}} - I\right)$.

David Kerr

Write $A = \sqrt{11}+3$, $B = \sqrt{11}-3$,

$$I_n \;=\; A^{2n+1} - B^{2n+1}, \qquad J_n \;=\; \frac{1}{\sqrt{11}}\left(A^{2n+1} + B^{2n+1}\right).$$

It is clear that I_n and J_n are integers, for when you expand the binomial expressions A^{2n+1} and B^{2n+1} and gather the terms together, everything

involving $\sqrt{11}$ cancels out and only integers remain. Also $0 < B^{2n+1} < 1$ and $A^{2n+1} = I_n + B^{2n+1}$; therefore $f = B^{2n+1}$, the fractional part of A^{2n+1}.

For (ii) we use induction. First, $I_0 = A - B = 6$ and $J_0 = (A+B)/\sqrt{11} = 2$ are both multiples of 2. Now suppose I_n and J_n are both multiples of 2^{n+1}. Consider I_{n+1} and J_{n+1}. We have

$$\begin{aligned} I_{n+1} &= A^{2n+1}A^2 - B^{2n+1}B^2 \\ &= 20\left(A^{2n+1} - B^{2n+1}\right) + 6\sqrt{11}\left(A^{2n+1} + B^{2n+1}\right) \\ &= 20I_n + 66J_n. \end{aligned}$$

Hence I_{n+1} is a multiple of 2^{n+2}. Similarly,

$$\begin{aligned} J_{n+1} &= \frac{1}{\sqrt{11}}\left(A^{2n+1}A^2 + B^{2n+1}B^2\right) \\ &= \frac{20}{\sqrt{11}}\left(A^{2n+1} + B^{2n+1}\right) + 6\left(A^{2n+1} - B^{2n+1}\right) \\ &= 20J_n + 6I_n \end{aligned}$$

and therefore J_{n+1} is a multiple of 2^{n+2}.

To prove (iii), we have

$$\begin{aligned} 11J_n^2 - I_n^2 &= A^{4n+2} + 2(AB)^{2n+1} + B^{4n+2} \\ &\qquad - \left(A^{4n+2} - 2(AB)^{2n+1} + B^{4n+2}\right) \\ &= 4(AB)^{2n+1} = 2^{2n+3}, \end{aligned}$$

using the fact that $(\sqrt{11}+3)(\sqrt{11}-3) = 2$. Rearranging and taking square roots,

$$\sqrt{I_n^2 + 2^{2n+3}} = \sqrt{11}J_n = A^{2n+1} + B^{2n+1} = I_n + 2B^{2n+1}.$$

Hence

$$f = B^{2n+1} = \frac{1}{2}\left(\sqrt{I_n^2 + 2^{2n+3}} - I_n\right),$$

as required.

Similar solutions were received from **Barry Lewis** (the originator of the problem), **Peter Fletcher** and **Sue Bromley**.

Viète's infinite irrational product

Jim James

It was our old friend Archimedes who started it all, when he introduced his polygon approximations to the circumference of a circle. This was in about 250 B.C., nearly 1800 years before Viète was even born.

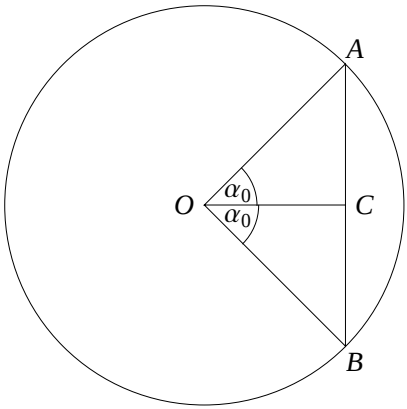
In modern terminology, the Archimedes polygon approximations relate to the fact that the circumference of a circle is bounded below by the perimeter of any n-gon inscribed within it and bounded above by the perimeter of any n-gon circumscribed about it; the greater the value of n, the closer the perimeters approach the circle circumference.

From his study of inscribed and circumscribed 96-gons and using the simplest of mathematical tools, Archimedes was able to deduce that π was greater than $223/71$ and less than $22/7$, (or $3.1408 < \pi < 3.1428$, approximately). Note that this result, giving π correct to better than 0.1%, was accomplished over 2000 years ago, without access to the calculus, trigonometry, algebra, logarithms, or even decimal place notation.

The polygon approximations, and variants thereof, remained the principal technique for π-students and researchers until well into the 17th century. They were used in the late 1500s by François Viète, a lawyer by profession but true to the spirit of the age a gifted amateur mathematician too, to prove his fascinating infinite irrational product. Here it is:

$$\frac{2}{\pi} = \sqrt{\frac{1}{2}} \times \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}} \times \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}}} \times \cdots.$$

Viète's adaptation of the polygon approximations, like those of many before him, was probably wholly geometric and, as a result, somewhat complicated. With a little elementary trigonometry, however, and a smattering of real analysis, we can derive the same result quite easily.



Consider a circle of unit radius, and a chord, AB, representing one side of an inscribed regular n_0-gon ($n_0 \geq 3$). Let α_0 be the angle subtended by half the chord at the centre of the circle, as in the diagram.

Let P_0 be the perimeter of the n_0-gon. Clearly,

$$P_0 \;=\; 2n_0 \sin\alpha_0. \tag{1}$$

If now $n_1 = 2n_0$, we can write

$$P_1 \;=\; 2\cdot 2n_0 \cdot \sin\frac{\alpha_0}{2} \;=\; 2n_0 \cdot \frac{\sin\alpha_0}{\cos(\alpha_0/2)} \;=\; \frac{P_0}{\cos(\alpha_0/2)}.$$

A further iteration, with $n_2 = 2n_1 = 4n_0$, gives

$$P_2 \;=\; \frac{P_1}{\cos(\alpha_1/2)} \;=\; \frac{P_0}{\cos(\alpha_0/2)\cdot\cos(\alpha_1/2)} \;=\; \frac{P_0}{\cos(\alpha_0/2)\cdot\cos(\alpha_0/4)}$$

and in general,

$$P_k \;=\; \frac{P_0}{\cos(\alpha_0/2)\cdot\cos(\alpha_0/4)\cdot\cdots\cdot\cos(\alpha_0/2^k)}. \tag{2}$$

Equations (1) and (2), together, provide a general solution for the perimeter of any inscribed regular $2^k n_0$-gon, formed from an original n_0-gon by k iterations of successive side doublings.

As k increases, so (2) defines an infinite sequence, $\{P_k\}$. We leave a formal proof that $\{P_k\}$ converges to the circle circumference as an exercise for those who actually enjoy such things. For those who don't we provide a semi-intuitive presentation in the appendix; this shows that $\lim_{k\to\infty} P_k = 2\pi$ which, in turn, gives

$$\pi \;=\; \lim_{k\to\infty}\left(\frac{n_0\sin\alpha_0}{\cos(\alpha_0/2)\cdot\cos(\alpha_0/4)\cdot\cdots\cdot\cos(\alpha_0/2^k)}\right). \tag{3}$$

To derive Viète's product we take $n_0 = 4$, so $\alpha_0 = 45^\circ$ and

$$\sin\alpha_0 \;=\; \cos\alpha_0 \;=\; \sqrt{\frac{1}{2}}.$$

And now, since $\cos\alpha_0 = 2\cos^2(\alpha_0/2) - 1$, we have

$$\cos\frac{\alpha_0}{2} \;=\; \sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}},$$

$$\cos\frac{\alpha_0}{4} \;=\; \sqrt{\frac{1}{2}+\frac{1}{2}\cos\frac{\alpha_0}{2}} \;=\; \sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}},$$

and so on. Substituting in Equation (3) we get

$$\pi \;=\; \frac{4\sqrt{\frac{1}{2}}}{\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}\times\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}}\times\ldots}$$

and hence

$$\frac{2}{\pi} \;=\; \sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}\times\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}}\times\ldots,$$

as required.

But this is not all. Equation (3) can be used to derive similar infinite products for other starting values of n_0. Try $n_0 = 3$ for example. This gives $\alpha_0 = 60°$, so $\sin\alpha_0 = \sqrt{3/4}$, $\cos\alpha_0 = 1/2$ and then

$$\cos\frac{\alpha_0}{2} \;=\; \sqrt{\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}} \;=\; \sqrt{\frac{3}{4}},$$

$$\cos\frac{\alpha_0}{4} \;=\; \sqrt{\frac{1}{2}+\frac{1}{2}\cos\frac{\alpha_0}{2}} \;=\; \sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{3}{4}}},$$

and so on. So now

$$\pi \;=\; \frac{3\sqrt{\frac{3}{4}}}{\sqrt{\frac{3}{4}}\times\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{3}{4}}}\times\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{3}{4}}}}\times\ldots},$$

which gives

$$\frac{3}{\pi} \;=\; \sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{3}{4}}}\times\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{3}{4}}}}\times\ldots.$$

Other curious infinite products can also be derived in this way, but the complications rapidly become horrendous (try $n_0 = 5$, for example). Together with the fact that the rate of convergence of such sequences is so low that they are of no practical use, the higher order infinite products are probably best left to the imagination!

What is important in all of this is that Viète's infinite irrational product of 1593 achieved two major advances.

1. It was the first representation of π as the limit of an infinite sequence of algebraic operations. This provided the motivation for later researchers to discover other, much more useful sequences, enabling π to be calculated to an accuracy greater than anyone can ever need.

2. It was also the first published infinite product of any kind, thereby introducing an extremely valuable tool for application in many diverse mathematical fields.

Appendix To show that $\lim_{k\to\infty} P_k = 2\pi$. Observe that since $P_k = P_{k-1}/\cos(\alpha_0/2^k)$ and $0 < \cos(\alpha_0/2^k) < 1$ for all $k \geq 0$, $\{P_k\}$ is strictly monotone increasing. Also, since all regular polygons are convex in shape and the length of a chord of a circle is less than that of the corresponding minor arc, we can write $P_k < 2\pi$ for all $k \geq 0$, that is, $\{P_k\}$ is bounded above by the circle circumference. Hence, by a theorem in real analysis, the sequence must converge.

Consider an arbitrary point, Q, lying within the circle, but not on its circumference, distance d from the centre. Clearly $0 \leq d < 1$. Now the shortest distance from the circle centre to any side of the inscribed $2^k n_0$-gon is $\cos(\alpha_0/2^k)$, (e.g., the length of the line segment OC in the diagram above, in which $k = 0$). It follows that if $d \leq \cos(\alpha_0/2^k)$, then Q lies either within the $2^k n_0$-gon or on its perimeter.

But $\cos(\alpha_0/2^k)$ converges to 1 as k increases, so however close Q is to the circle circumference (however close d is to 1), we can always select finite values of k such that $d \leq \cos(\alpha_0/2^k) < 1$. It follows that as k tends to infinity, every point lying within the circle, but not on its circumference, will lie either within the inscribed $2^k n_0$-gon or on its perimeter. Since no part of the inscribed $2^k n_0$-gon can lie outside the circle, the only way for this to occur is for the perimeter to approach 'infinitesimally close' to the circle circumference; that is as required.

Plugging Pascal's triangle

Elsie Page

When one side of Pascal's triangle is blocked, a modified version is obtained.

Pascal's triangle	becomes	
1	**1**	0
1 1	1 0	1
1 2 1	1 **1** 0	2
1 3 3 1	1 2 0 0	3
1 4 6 4 1	1 3 **2** 0 0	4
1 5 10 10 5 1	1 4 5 0 0 0	5
1 6 15 20 15 6 1	1 5 9 **5** 0 0 0	6
1 7 21 35 35 21 7 1	1 6 14 14 0 0 0 0	7
1 8 28 56 70 56 28 8 1	1 7 20 28 **14** 0 0 0 0	8

This new arrangement yields an interesting comparison. The numbers in the central column—those in bold—are called *Catalan numbers*, two appearances of which are given below. They can each be simply expressed as an exact fraction of the corresponding entry in Pascal's triangle.

$$14 = \frac{70}{5}, \quad 5 = \frac{20}{4}, \quad 2 = \frac{6}{3}, \quad 1 = \frac{2}{2} \quad \text{and} \quad 1 = \frac{1}{1}.$$

So for even numbered rows, the central value is ${}^{2n}C_n/(n+1)$, where $n = k/2$. What is more, every non-zero entry in the restricted diagram is also a simple fraction of its Pascal counterpart. These can be proved by induction using the step(s)

$$\frac{k+1-2r}{k+1-r}\,{}^kC_r + \frac{k+1-2(r+1)}{k+1-(r+1)}\,{}^kC_{r+1} \;=\; \frac{k-2r}{k+1-r}\,{}^{k+1}C_{r+1},$$

$r = 0, 1, 2, \ldots, k/2$ or $(k-1)/2$.

The total for each line is ${}^{2k}C_k$ and, again, this can be proved inductively.

Here are two examples which generate the Catalan numbers.

First, a game consists of tossing a coin where for each toss, heads means a win for you and tails a loss. The game lasts for n trials—a trial being two tosses of the coin.

Suppose you are interested in knowing for what proportion of the terms you can expect to be continuously either in the lead or drawn.

For one trial the possible results are

HT		UE	U = up
HH	which can appear as	UU	D = down
TH		DE	E = even
TT		DD	

The successes are 2 out of 4. For two trials, HHHH → UUUU, HTHH → UEUU, etc. Successes here are 6 out of 16.

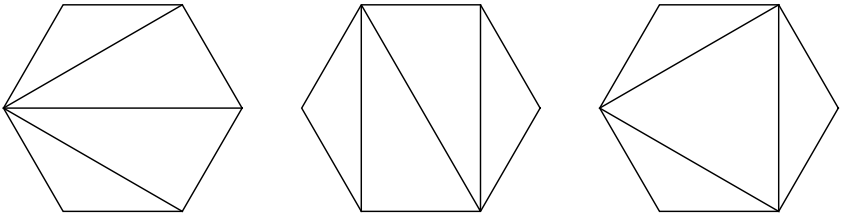
In some cases the 'successful' sequences will end in a draw, and it is the total numbers of these which are Catalan.

For one trial	1 result	HT → UE
For two trials	2 ways	HTHT → UEUE, HTHT → UEUE
For three trials	5 ways	
For four trials	14 ways	

This example can be modelled directly by the restricted triangle shown at the beginning.

The second example is of a different nature. Consider the number of ways that a convex polygon can be divided into triangles by straight lines joining the vertices. The lines must not intersect and each vertex is distinct.

For a triangle	1 way
For a quadrilateral	2 ways
For a pentagon	5 ways
For a hexagon	14 ways, as illustrated below



It looks as though the Catalan numbers are being generated but what is the connection?

Do you know your infinity times tables?

Martin Cooke

The function x^x is interesting and begs the question of the value of 0^0 (which is addressed by Kaplan in [1]). Since $x^y = x^{y+0} = x^y x^0$ it may seem that 0^0 can be any value which when multiplied by zero can yield zero. In particular situations this will be $0^0 = 1$, but the general, intrinsic case seems to be $0^0 \in F$, where F is a field with additive identity 0 (for example, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$); explaining why different paths in $\mathbb{C}^2$ can tend to different limits for z^z at $z = 0$, as described in [1].

Manipulating this general case with set structures (e.g. $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$) is cumbersome, however, and the division by zero of the extended number line (e.g. the most basic geometry, projective, adds a point at $\infty \equiv 1/0$ to $\mathbb{R}$, giving $\mathbb{R} \cup \{\infty\}$, see [2]) also sits uneasily with the standard set-theoretical reductionism. So, regarding maths as the science of patterns (of numbers), I propose a different sort of collection as the most basic, or natural, for our number systems, by showing that it is the neatest way to extend the complex numbers—much as i is introduced by extending $\sqrt{}$ to apply to -1.

Since the natural numbers can be regarded as abstracted from collections of similar things (alongside noun conception) I will collect only dissimilar numbers, so $(1, 2, 2) = (1, 2)$ but, unlike sets, I will have no substructures, so $(2) = 2$, although set properties can of course be added. For example, let $X = (x|x \in \mathbb{R} \cup \{\infty\})$, so X has a subcollection $\mathbb{R}$, which is a field (and hence, technically, a set). These 'extended sets' can be potentially infinite and so sit easily alongside a geometry where lines are not made of points but may contain $\mathbb{R}$ and $\mathbb{R} \cup \{\infty\}$ equally naturally. But I shall now get back to 0^0.

For a field F with identities 0 and 1, let $\infty = 1/0$. Also consider the extended set $X = (x|x \in F \cup \{\infty\})$. Since $-\infty = \infty$ and $0\infty = 0^0$ I also have $\infty \in 0^0$ (where $\in$ applies to extended sets or sets). Hence $0^0 = X$ (and the lack of a subset structure to extended sets makes restrictions, e.g. to $0^0 = 1$, much easier) and also $\infty + \infty = X$ since, e.g. $\infty + 1 = \infty$ so $1 \in (\infty + \infty)$. This means that the extended set does not have distributivity (since $2\infty = \infty$) although its substructure F does, of course.

Now to the point of all this, the following structures for multiplication and division (of top row by left column) of $X = (0, 1, \infty)$, where replacing the unit by a general non-zero rational ($F = \mathbb{Q}$) or real ($F = \mathbb{R}$) does not change the structure. Compare with the rather patternless field substructures.

$\times$	0	1	∞	X
0	0	0	X	X
1	0	1	∞	X
∞	X	∞	∞	X
X	X	X	X	X

$\div$	0	1	∞	X
0	X	∞	∞	X
1	0	1	∞	X
∞	0	0	X	X
X	X	X	X	X

$\times$	0	1
0	0	0
1	0	1

$\div$	0	1
1	0	1

Finally, the extension of $\mathbb{C}$ shows the naturalness of the pattern, with two units (1 and i) and negatives included; notice how the blocks of ∞s and 0s correspond with the elements of the blocks of four units which are either (i) of different sign to the other three, or (ii) the opposite sign to the element 1 in that block.

Incidentally, extended sets may seem vague, but they follow from the way we list roots of unity, for example. So $0^0 = X$ indicates an extended set of potential values, and we may be interested in one, several or all of these without prejudice. So if 0 is the additive identity of a field, then I think that $0^0 = X$, where X is the extension of that field. Also, insofar as 0 is the number zero and naturally embedded in $\mathbb{C}$, 0^0 is the extension of $\mathbb{C}$.

$\times$	X	∞	$-i$	-1	0	1	i	∞	X
X	X	X	X	X	X	X	X	X	X
∞	X	∞	∞	∞	X	∞	∞	∞	X
$-i$	X	∞	-1	i	0	$-i$	1	∞	X
-1	X	∞	i	1	0	-1	$-i$	∞	X
0	X	X	0	0	0	0	0	X	X
1	X	∞	$-i$	-1	0	1	i	∞	X
i	X	∞	1	$-i$	0	i	-1	∞	X
∞	X	∞	∞	∞	X	∞	∞	∞	X
X	X	X	X	X	X	X	X	X	X

$\div$	X	∞	$-i$	-1	0	1	i	∞	X
X	X	X	X	X	X	X	X	X	X
∞	X	X	0	0	0	0	0	X	X
$-i$	X	∞	1	$-i$	0	i	-1	∞	X
-1	X	∞	i	1	0	-1	$-i$	∞	X
0	X	∞	∞	∞	X	∞	∞	∞	X
1	X	∞	$-i$	-1	0	1	i	∞	X
i	X	∞	-1	i	0	$-i$	1	∞	X
∞	X	X	0	0	0	0	0	X	X
X	X	X	X	X	X	X	X	X	X

References

[1] Kaplan, *The Nothing that is*, Penguin, Harmsworth, 1999.

[2] Brennan, Esplen & Gray, *Geometry*, CUP, 1999.

Square roots

Pete Charlton

In M500 **74**, October 1981, Colin Davies asked about finding square roots as taught at school in the old days—now it's press a button. This note explains the thinking behind the method. My collection of M500s is not complete, so I may have missed the answer, but there must new (modern?) students who have not heard of this method.

The identity $(p+x)^2 = p^2+2px+x^2$ is used at each stage of the process, modified as below, which is best explained by an example:

Square root of 537.4

Step 1. Starting at the decimal point, divide the number into pairs of digits in both directions.

	2 3. 1 8	
	5 37.40 00 00 ...	step 1
(see note)	4	step 2
	1 37	
$(43 \times 3 =)$	1 29	step 3
	8 40	
$(461 \times 1 =)$	4 61	step 4
	3 79 00	
$(4628 \times 8 =)$	3 70 24	
	8 76 00	

Step 2. What is the largest (whole) number whose square is less than 5? Answer 2, with square 4. Write 4 under the 5, subtract and bring down the next two figures (37) to give 137 as the next dividend. Write 2 above the 5. Ignore all decimal points in this process, so that, after step 4, $p = 231$ not 23.1, and treat the dividends as whole numbers.

Step 3. Let $p = 2$ (the answer to date, the p^2 in the identity). Now we want the remainder. But we are looking for a digit which is ten times less significant than the answer to date, so use $10p$ for p in the modified identity, i.e., $2 \cdot 10px + x^2 = 20px + x^2$. Then we have to solve $20px + x^2 < 137$ to find x, the next digit; $x = 2$ gives 84, $x = 3$ gives 129. Subtract 129 from 137 to give 8 remainder and bring down the 40 to make 840 as the next dividend, and write 3 above the 37.

Step 4. Now p equals 23, use $20px + x^2 < 840$; $x = 1$ gives 461, $x = 2$ gives 924. Therefore write 1 above the 40 and subtract 461 from 840 to give 379, bring down the next two figures (00) to make 37900 the next dividend.

Repeat as required.

Note: At school I was taught, after step 2, to double the 2 (making 4) and to write 4 to the left of 137, then find a number such that the 4 becomes forty-something times the something so that the answer was less than 137, as shown on the left above. Or in algebraic terms, $(2\cdot 10p+x)x = 20px+x^2$, as above.

To find cube roots, use the same procedure. The number is divided into groups of three either side of the decimal point. Find the largest number whose cube is less than the first group then use the identity $(p+x)^3 = p^3 + 3p^2x+3px^2+x^3$ in the form $3(10p)^2x+3\cdot 10px^2+x^3 = 300p^2x+30px^2+x^3 <$ dividend.

The system will work for cubes and should work for higher roots, but it looks very laborious, and I have not tried it.

Solution 173.2 – Nine darts

Chris Pile

There are 22 ways to score 501 with nine darts in a standard game (ignoring order of scoring on the first eight darts). There are six possible scores for the last dart, which are the five doubles 24, 30, 34, 36, 40, and the bull, which is 50, and so there are six corresponding scores for the first eight darts, which are 477 (one way), 471 (three ways), 467 (one way), 465 (seven ways), 461 (three ways), and 451 (seven ways).

Finish	8-dart score	Last eight darts
24	477	$(7\cdot 60+57)$
30	471	$(7\cdot 60+51)$ $(6\cdot 60+57+54)$ $(5\cdot 60+3\cdot 57)$
34	467	$(6\cdot 60+57+50)$
36	465	$(7\cdot 60+45)$ $(6\cdot 60+57+48)$ $(6\cdot 60+54+51)$ $(5\cdot 60+2\cdot 57+51)$ $(5\cdot 60+57+2\cdot 54)$ $(4\cdot 60+3\cdot 57+54)$ $(3\cdot 60+5\cdot 57)$
40	461	$(6\cdot 60+51+50)$ $(5\cdot 60+57+54+50)$ $(4\cdot 60+3\cdot 57+50)$
50	451	$(6\cdot 60+57+34)$ $(6\cdot 60+51+40)$ $(5\cdot 60+57+54+40)$ $(5\cdot 60+51+2\cdot 50)$ $(4\cdot 60+3\cdot 57+40)$ $(4\cdot 60+57+54+2\cdot 50)$ $(3\cdot 60+3\cdot 57+2\cdot 50)$

Solution 174.3 – Eight wires

> I am on the third floor with one end of an 8-way cable (all eight wires are identical) and I know the other end is in the basement. Given a continuity meter, what is the least number of trips to the basement I must make to identify each wire?

David Kerr

One trip is all that is needed.

Label the wires on the third floor A to H and join A to B, C to D and E to F.

Go to the basement and find three pairs which show a completed circuit. Label these {1,2}, {3,4} and {5,6}. Label the other two 7 and 8. Thus $\{G, H\} = \{7, 8\}$ and $\{\{A, B\}, \{C, D\}, \{E, F\}\} = \{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$.

Join 1 to 7, 2 to 3 and 4 to 5.

Climb to the third floor and undo the connections there.

One of G or H will now form a circuit with one of A to F. Assume arbitrarily it is G to A. We now know that G joins to 7 and A to 1, and hence B to 2 and H to 8. Next, B will form a circuit with C, D, E or F. Assume it is C. This means that C joins to 3 and D to 4. Now D will form a circuit with E or F. Assume it is E. This gives $E = 5$ and $F = 6$.

ADF writes—I also had responses from **Malcolm Maclenan** (who posed the problem in the first place), **R. M. Boardman** and **Colin Davies**. However, all three solutions require a second trip to the basement.

Malcolm asks about a general method for n wires. We have it. David's solution for eight wires extends to any number $n \geq 4$ (with a slight adjustment if n is odd). I needed convincing, so I tried it out. I made a 26-way cable out of some unused wire which I found in my attic. I don't have a third floor; I used the ground floor instead, and I placed the other end of the cable in an imaginary basement. I labelled the ground-floor ends of the wires A, B, ..., Z, and I adapted David's procedure in the obvious manner. It worked!

For the remaining values of n: Clearly, $n = 3$ can be solved with one trip to the basement and $n = 1$ needs no trips at all. That leaves the most useful case. Think of how often you might want to run a bell cable from a battery in your house to a polarity-sensitive device in a shed some distance away—and bell cable has no markings to distinguish between the two wires. Well, if you can find a solution to the problem for $n = 2$, we would very much like to see it!

Solution 174.2 – Incredible identity

Show that

$$\sqrt{11+2\sqrt{29}}+\sqrt{16-2\sqrt{29}+2\sqrt{55-10\sqrt{29}}} \;=\; \sqrt{5}+\sqrt{22+2\sqrt{5}}$$

exactly.

Sue Bromley

This appears to be a question of first spotting that $16 = 11 + 5$ and then remembering (eventually, in my case) that $(\sqrt{a}+\sqrt{b})^2 = a + 2\sqrt{a}\sqrt{b} + b$.

$$\begin{aligned}
&\sqrt{11+2\sqrt{29}}+\sqrt{16-2\sqrt{29}+2\sqrt{55-10\sqrt{29}}}\\
&\quad= \sqrt{11+2\sqrt{29}}+\sqrt{11-2\sqrt{29}+2\sqrt{5}\sqrt{11-2\sqrt{29}}+5}\\
&\quad= \sqrt{11+2\sqrt{29}}+\sqrt{\left(\sqrt{11-2\sqrt{29}}+\sqrt{5}\right)^2}\\
&\quad= \sqrt{11+2\sqrt{29}}+\sqrt{11-2\sqrt{29}}+\sqrt{5}\\
&\quad= \sqrt{5}+\sqrt{\left(\sqrt{11+2\sqrt{29}}+\sqrt{11-2\sqrt{29}}\right)^2}\\
&\quad= \sqrt{5}+\sqrt{(11+2\sqrt{29})+2\sqrt{(11+2\sqrt{29})(11-2\sqrt{29})}+(11-2\sqrt{29})}\\
&\quad= \sqrt{5}+\sqrt{22+2\sqrt{121-116}} \;=\; \sqrt{5}+\sqrt{22+2\sqrt{5}}.
\end{aligned}$$

Also solved in a similar manner by **David Kerr**, **Peter Fletcher** and **John Bull**.

Gerald Whitrow

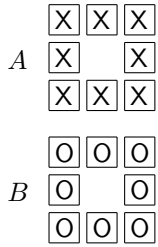
EK

Professor Whitrow died on June 2, aged 87. All his life he was obsessed by time, considering that it had been unjustly neglected in the work of his contemporaries. His last book was *Time in History*, 1988, but he wrote numerous papers and articles on the philosophy of time, many appearing in popular media. When he provided three talks for Radio 3 it was unfortunate that they played them in the wrong order.

He was fond of telling how he had once at a party told a fellow guest that he was to see Popper the next day; the reply was 'Good heavens, is your father still alive.' He was happily married for 53 years. His wife survives him.

Roll me over

Eight cubes have a pair of opposite faces marked with 'X' and 'O'. They are placed in a box in configuration A. Get them into configuration B. The only legal move is to roll a cube into the vacant space (leaving a hole behind it for the next move). No sliding, no lifting out and putting back.



Tony Forbes

I came across this problem while I was browsing old M500s to see if there was anything suitable for the '25 years ago' department. The problem was stated in M500 **28** (under the same title as this article), and in February 1976 a 38-move solution, described as 'probably minimal', appeared in M500s **30** and **31**. I became interested when it occurred to me (in August 2000) that settling the question of minimality should be easy—with a little help from a personal computer.

We use the letters L, R, U, D to denote, respectively, roll a cube to the left, to the right, up, down. We can assume that the first three moves are LUR, so let S_n be the number of n-move sequences that start with LUR. Of these S_n sequences, suppose M_n leave the hole in the middle of the array and C_n leave the hole in a corner. Then we have a simple recursion:

$$S_3 = M_4 = C_4 = 1,$$
$$S_{2n} = M_{2n} + C_{2n}, \qquad S_{2n+1} = 3M_{2n} + C_{2n},$$
$$M_{2n+2} = C_{2n}, \qquad C_{2n+2} = 6M_{2n} + C_{2n}.$$

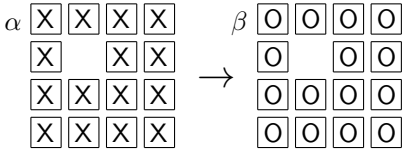
Observe, by the way, that $M_{2n+1} = C_{2n+1} = 0$ and $S_{2n+2} = 2S_{2n+1}$.

One can now calculate that there are 232,426,077 sequences of up to 36 moves. It is not much trouble to test them all. The result: just two 36-move solutions,

LURD LDRU RDLU RULL DRDL URUL DRDL UURD DRUL

and its reverse. Hence **36 is best possible**. The problem is solved.

The $\mathbf{4 \times 4}$ problem, too, can be completely solved with a reasonable amount of computation. This time we have to get fifteen cubes from configuration α into configuration β.



It is easier if we split the task into two stages: First get the cubes in a pair of adjacent edges correct. Then solve the remaining eight cubes (which occupy a 3×3 array) without disturbing the first seven. In this way I managed to find a number of 48-move solutions; the first that came out of

the computer was

LLUR RRDL DLLU RULD RUUR DDRD
LLUU ULDR URDL URRD LLUR DRDL.

To prove minimality it suffices to examine every possible sequence of up to 46 moves. Not 47, because an odd number of moves will displace the hole from its original position. We can halve the work by considering only sequences beginning with L or R. Furthermore, it speeds things up greatly if we note that for testing sequences of length at most n we can abandon an entire branch of the search tree whenever $s + 2x + y > n$, where s is the number of moves executed, x is the number of cubes showing 'X' and y is the number of cubes showing neither 'X' nor 'O'. The idea, of course, is that these cubes require at least two moves and one move respectively to put right.

There were 63,318,023,509 sequences to test. No solutions were found; hence **48 is best possible**.

The $\mathbf{5 \times 5}$ problem is more difficult. Doing it in two stages did not seem to work very well, so I had to rely on inspired guesswork and a certain amount of patience to obtain this 68-move solution:

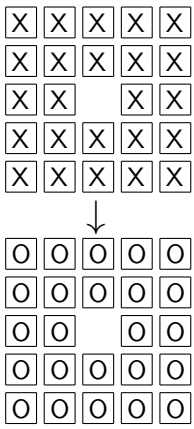
LURR DDLL LUUR URRD RDLU URDD DLDL ULUR DRRD
LLLL URUU LURR RDDR DLLD LULU URRR DDLU.

It turns out that **68 is best possible**. We prove this by systematically testing sequences of up to 66 moves—not a trivial exercise.

It is sufficient to consider sequences beginning with LLU or LU and it helps if we refine the inequality stated in the 4×4 case to

$$s + 2x + y + 2e + \min\{4,\ f + g + h\} > n.$$

Here, n, s, x and y are as before, e is the number of cubes in the border of the array that have the 'X' face pointing outwards, f is the number of cubes that have the 'X' face pointing towards a cube showing 'O', g is the number of adjacent pairs of cubes that have their 'X' faces pointing towards each other, and h is the number of occurrences of three cubes in a line, ABC, say, where A and C have their 'X' faces pointing towards each other and B does not show 'O' and does not have its 'X' face pointing towards A or C. The mysterious parameter 4 is an unwelcome complication. Replacing it by infinity slows the program down too much.



Re: Problem 171.1 – Cylinder

Colin Davies.

I also spent time in bed at 3 a.m. thinking about Gordon Alabaster's analysis of the fall of a cylinder [M500 **174** 16].

His conclusion is apparently that a cylinder will behave in the same way as the rectangular box into which it fits. I think this is only true in practice if it is dropped vertically from a fixed position so it has no rotation or sideways motion as it falls. This seem unlikely in any real situation. A tossed coin usually follows a parabolic path, so goes a bit sideways, and spins as well.

Unless the falling cylinder lands exactly flat on its circular end, or exactly flat along a straight edge of its curved side, both situations being highly unlikely, it must initially touch the floor with the rim of a circular end. If the cylinder has any sideways motion with a component along the tangent to the rim along the floor, the cylinder will tend to roll along its rim, following the tangents along the floor instant by instant, so it will tend to roll on the floor in a curve at the same time as it rotates around its cylindrical axis. It won't just fall over as Gordon A's cube does. If the cylinder spins while falling and then hits the floor, the effect will be similar, which is presumably why a spun coin, having slowed and fallen over, then runs (and bounces a bit) around its rim at a low angle to the ground before finally ending on its flat face.

As the cylinder is rotating, the axis of rotation has to be in a vertical plane at right angles to, and moving with, the succession of tangents followed by the rim. There will be a coriolis force along that vertical plane, which also goes through the centre of mass of the cylinder. This will not make much difference to the final outcome unless the centre of mass is very nearly above the point of contact of the rim and the floor, when it may flip the centre of mass over the point of balance.

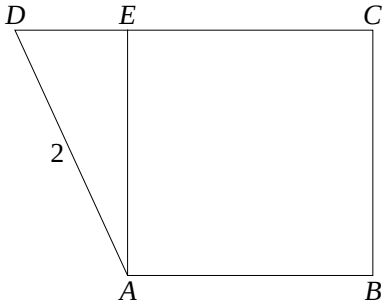
If the cylinder rotates clockwise, the rolling direction will be to the right, and if anticlockwise, to the left. Either way I think the coriolis force will be in the same direction. I don't have a gyroscope to experiment with, but I have a gut feeling that the coriolis force will push the cylinder towards falling on its cylindrical surface.

So I suspect that in practice a cylinder of radius to height ratio of 1 to 2 is more likely to fall onto its curved edge. I am sure that the way the cylinder is tossed (how much lateral component), and whether or not it is spun as well, will have a very complicated effect on the probability of a curved or flat landing surface.

'URGENT!! Did you receive this letter? If not, please read it now and act as quickly as possible, as time is running out.'—Insurance company mailing. [Sent by Tom Barker.]

Quadrilateral

Ken Greatrix



Re Harold Moulsen's quadrilateral, M500 **173**, page 29. From the data given I construct the figure shown. Here, $CD > AB$ but we could also have $CD = AB$ or $CD < AB$. We require the length of AB which makes the area of the quadrilateral $ABCD$ equal to 4.

Let AE meet CD (extended if necessary) at 90° so that $ABCE$ is a square and AED is a right-angled triangle.

Denote the length of DE by x. Then $AB = \sqrt{4 - x^2}$ and the area of $ABCD$ is

$$4 - x^2 + \frac{1}{2}x\sqrt{4 - x^2}, \qquad \text{(I)}$$

and this has the value 4. Hence $x = 0$, which gives the 'obvious' solution $AB = 2$, or $x = 2/\sqrt{5}$, $AB = 2\sqrt{4 - 4/5} = \sqrt{3.2} \approx 1.78885$.

Expression (I) can be used to obtain the maximum area. Differentiating and equating to zero, we have

$$-2x - \frac{1}{2} \cdot \frac{x^2}{\sqrt{4 - x^2}} + \frac{1}{2} \cdot \sqrt{4 - x^2} \;=\; 0.$$

Simplifying,

$$x^2 - 2 \;=\; -2x\sqrt{4 - x^2},$$

which on squaring becomes

$$5x^4 - 20x^2 + 4 \;=\; 0.$$

The solutions are $x = \sqrt{2 \pm 4/\sqrt{5}}$. The maximum occurs at $x = \sqrt{2 - 4/\sqrt{5}} \approx 0.45951$, giving

$$AB \;=\; \sqrt{4 - x^2} \;=\; \sqrt{2 + \frac{4}{\sqrt{5}}} \approx 1.94650,$$

and the area of $ABCD$ is $\sqrt{5} + 2 \approx 4.23607$.

From a Radio 4 discussion on dangerous sports: 'If they do make bungee jumping illegal, they'll just drive it underground.' [Spotted by EK.]

Solution 174.4 – 32 pounds

> You start with £32 and bet against your opponent on the toss of a coin. On each turn you stake half your capital, and your opponent matches your stake. You play six times, and you win half of the plays. What is your capital now?

Tony Huntington

I reckon JRH should pick his betting partners more carefully. After winning half of the six plays he will have only £13.50p left of his £32.

After each play your total capital is either 50% or 150% of the amount before the play. So to take a fairly simple problem and complicate it with mathematical symbolism:

$$T_n = \begin{cases} \frac{1}{2}\,T_{n-1} & \text{for a lose} \\ \frac{3}{2}\,T_{n-1} & \text{for a win} \end{cases}$$

$$T_n = \left(\frac{1}{2}\right)^{n-w}\left(\frac{3}{2}\right)^{w} = \frac{3^w}{2^n}\,T_0,$$

where T_n is your total capital after n plays, T_0 is your initial capital, and w is the number of wins in n plays. Sticking in the numbers and turning the handle gives you

$$T_6 = \frac{3^3}{2^6}\,32 = 13.5.$$

Also solved by **Arthur Quigley** (see page 26) and **R. M. Boardman**.

Problem 176.1 – Two cyclists

Keith Drever

Two cyclists were travelling towards each other, one travelling at 10 m.p.h., the other at 20 m.p.h. When the riders were 180 miles apart, a fly left the handlebar of one cycle, and travelled towards the other cyclist. When it reached the latter, it instantly reversed direction and flew back to the first cyclist, and continued winging back and forth between them until the two cyclists eventually met.

If the fly's speed was 100 m.p.h., what was the total distance that the fly had covered?

Solution 174.1 – Four people

> A, B, C, and D have to get across a bridge at night. The bridge cannot take more than two at a time. They have one torch, and no crossing can be made without the torch. A can cross the bridge in one minute, B in two minutes, C in five and D in ten. When two people cross, they travel at the speed of the slower person. What is the shortest time for all four to get across?

Gail Volans

At last, a problem I can solve ...

The quickest possible time is 17 minutes.

A and B go across. A comes back. $2 + 1$ minutes. C and D go across. B comes back. $10 + 2$ minutes. A and B go again, 2 minutes.

Kiwi fruit

ADF

I thought that the kiwi-fruit situation (where a supermarket sells kiwi fruit at 15p each, ten for a pound [M500 **174**, 25]) was a one-off, never to be repeated. Well, I was in Sainsbury's a week or two later and I saw this special offer:

> Nectarines: 35p. Buy 4 for £1.

So I presented myself at the check-out with *three* nectarines (and a bag of potatoes); £1.05 came up on the cash register for the nectarines. I protested, of course, but there was no room for debate. That was the amount to pay. The reduced price only applied to four. However, they did agree to sell me one twice!

POTATOES	0.86
NECTARINE	
3 @ £0.35	1.05
4 ITEMS PURCHASED	
BALANCE DUE	1.91
NECTARINE	0.35
MULTIBUY	
BUY 4 FOR 100P	-0.40
5 ITEMS PURCHASED	
BALANCE DUE	1.86

Problem 176.2 – Population

In a given population 2/3 of the men are married and 3/5 of the women are married. What fraction of the population are married?

Letters to the Editors

32 pounds

Sir,

I believe it is £13.50 which if it is correct is highly counter-intuitive. I mean the answer to JRH's problem 174.4.

At first common sense leads one to doubt that there is a definite answer— since losses at the start and losses at the end don't appear to be equivalent. But if some mathematical insight allows one to believe there is an answer then common sense tells one, once again, that the answer must be £32.00p. After all surely this is the only way to balance the equal wins and losses.

Besides isn't the opponent in exactly the same position—three wins and three losses—so he must also be left with £13.50p, if this is the answer. So where on earth has the remainder of £37.00p gone to? After all they chipped in £32.00p each, and twice this minus twice £13.50p is £37.00p, I think.

And another thing—why are the figures so quirky, when the problem is based on halving and doubling a number which is a power of two. One hardly expects primes and fractions to come into the puzzle.

May I suggest a worthy task and a challenge to literate folk amongst your mathematical readers, would be to explain the result in ordinary English, so that it makes simple sense, with whatever metaphors and analogies are appropriate. I believe Michael Faraday in a different context, faced with some fearsome looking field equations, asked the same from Maxwell.

I don't know what his answer was though—perhaps someone else does. Lots of mathematics, but statistics in particular, is replete with surprising consequences, which frequently remain surprising even after the mathematical expert has explained the result. For example, I found the following in Ch. 5 on p. 53 in Martin Gardner, *Mathematical Puzzles & Diversions* (Penguin, 1991).

Ask the question 'What is the probability that the other child is a boy?'—the answer is one half. Ask the question again (to someone else)—the answer is one third. In the first case the information is 'that Smith has two children—the oldest is a boy' and in the second 'that Smith has two children—at least one of whom is a boy'.

Gardner explains. In the first case the possibilities are BB & BG. In the second, BB, BG & GB. According to Gardner, in both cases, all the possibilities listed are equally probable.

He also points out that information on age is not a crucial ingredient in the puzzle—any information on 'ordering' would do to alter the answer in the way described. The question might mention that the heaviest was a boy, or the tallest, he says. Presumably if one were to say 'the one with the

longest fingernails was a boy', that would alter the answer as well.

Now this is all very surprising, if I've understood it properly. One could draw up a probability function for a variable, which varied with any other variable whatever—in spite of the lack of any obvious dependency. I wonder what your readers versed in statistics make of it.

The book has been through many editions and publishers—so is widely available. I'd be curious to get any thoughts on it.

Best wishes,

Arthur Quigley

Chords

Dear Tony,

Further to the chords problem [M500 **169**, page 19: if we have a regular n-gon inscribed in a unit circle, what is the product of the $n-1$ chords from a given vertex?], which has generated a lot of interesting contributions, I have come across in the *USSR Olympiad Problem Book* a most interesting formula.

Though $\Pi(\text{chords}) = n$ and $\Sigma(\text{chords})^2 = 2n$, it seemed to me for a long time that there was no formula for $\Sigma(\text{chords})$; i.e.

$$\sum_{r=1}^{n-1} 2\sin\frac{r\pi}{n} \;=\; 2\left(\sin\frac{\pi}{n} + \sin\frac{2\pi}{n} + \cdots + \sin\frac{(n-1)\pi}{n}\right).$$

The formula is $2\cot\dfrac{\pi}{2n}$, a result derived from a more general formula

$$\sin\phi + \sin(\phi+\alpha) + \sin(\phi+2\alpha) + \cdots + \sin(\phi+n\alpha)$$

$$= \frac{\sin\left(\dfrac{n+1}{2}\,\alpha\right)\sin\left(\dfrac{n}{2}\,\alpha+\phi\right)}{\sin\dfrac{\alpha}{2}}$$

with $\phi = 0$ and $\alpha = 2\pi/n$. Proof is by complex numbers—have a go!

All the best,

Sebastian Hayes

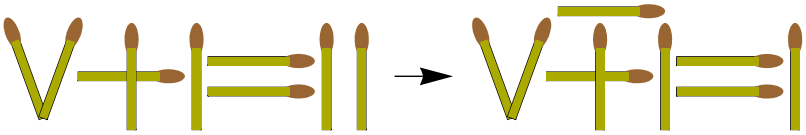
Problem 176.3 – Tricubic

Find the positive real root of $x^9 + 768x^6 = 768156$.

Nine matches

Dear Tony,

Re: M500 **174**, p. 23 [Move one match to make this correct: ].



seems better than anything else!

David Singmaster

[*Correction*: Of course, Martin Cooke's second solution in M500 **174** should have been  .]

Complex complex complex

Dear Jeremy,

With regard to your (complex)^3 challenge, the mention of a pub sign (M500 **173**, page 28) reminds me of the more contrived sign for a shop with two proprietors, 'HAMMAND and ANDERSON'. The discussion concerned the relative spacing of 'HAMM' and 'AND', and 'AND' and 'and', and 'and' and 'AND', and 'AND' and 'ERSON'.

Chris Pile

Problem 176.4 – Factorial squares

John Reade

When is $n!+1$ a square? For example, $4!+1 = 25 = 5^2$, $5!+1 = 121 = 11^2$, $7!+1 = 5041 = 71^2$.

What is the next example? Are there infinitely many examples?

[If that's too difficult, what about $n!-1$ a square?—eds.]

Problem 176.5 – Construct a square

R. M. Boardman

Given a unit length line segment, construct a square of side one unit, using only a pair of compasses.

Twenty-five years ago

From M500 27

Roger Bridgman—Get your brains round this one, logic freaks. I thought of it while washing my hair; it's a bit thin perhaps, but then so is the hair:

THEOREM *No theorem is always true.*

PROOF If the theorem were untrue, a counterexample would exist. Which would show that the theorem was not always true. But then it would be an example, not a counterexample. The contradiction proves the theorem.

But it's still not always true.

New ideas are born out of a state of confusion—A. N. Whitehead.

Lytton Jarman—I saw an OU degree certificate which is extremely disappointing. It doesn't list the courses and is the plainest of documents. Certainly it is the poorest of any literature produced by the OU. Has anyone complained? You get a better certificate for a bronze medal in Ballroom Dancing.

Winter Week-end

Norma Rosier

The twentieth M500 Society WINTER WEEK-END will be held at **Nottingham University** from **Friday 5 to Sunday 7 January, 2001**.

This is an annual residential Weekend to dispel the withdrawal symptoms due to courses finishing in October and not starting again until February. It is an opportunity to get together with friends, old and new, and do some interesting mathematics. It promises to be as much fun as ever!

Ian Harrison is running it and the theme will be the **History of Mathematics**. Anyone investigating mathematics for themselves is often recreating paths trodden by others before. The Week-end is an opportunity to match your skills against mathematicians of the past, and share understandings of problems that have enticed people for centuries. You may be surprised at what some mathematicians in the past got up to, and find that this casts fruitful fresh light on your mathematical activity and understanding today.

Cost: £130 for M500 members, £140 for non-members. This includes accommodation and all meals from dinner on Friday to lunch on Sunday. Please send a stamped, addressed envelope for booking form to Norma Rosier.

Contents M500 176 – October 2000