MATHEMATICAL SPECTRUM

A MAGAZINE FOR STUDENTS AND TEACHERS OF MATHEMATICS AT SCHOOLS, COLLEGES AND UNIVERSITIES

Volume 23 1990/91 Number 2

Mathematical Spectrum is a magazine for students and teachers 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 23 of *Mathematical Spectrum* will consist of four issues, of which this is the second. The first was published in September 1990, and the third and fourth will appear in February and May 1991 respectively.

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.

The Editorial Committee welcomes the submission of suitable material, including correspondence, queries and solutions to problems, for publication in *Mathematical Spectrum*. Students are encouraged to send in contributions. All correspondence about the contents should be sent to:

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

Mathematical Spectrum Awards for Volume 22

Prizes have been awarded to the following student readers for contributions published in Volume 22:

Chris Nash for his article 'Musings on an Interesting Sequence' (pages 19–22);

Gregory Economides for problems, solutions to problems, and reviews.

Additional prizes have been awarded to the following:

K. Prakash for his article 'A Sequence Free from Powers' (pages 92–93);

Amites Sarkar for various contributions.

Aimée Davison, who is younger than the majority of our contributors, has been awarded a special prize for a solution to a problem.

You Will Get Your Pi, Eventually

P. GLAISTER, *University of Reading*

The author is currently a lecturer in mathematics at the University of Reading. His research interests include computational fluid dynamics, numerical analysis, perturbation methods and mathematical education. He is a keen cricket enthusiast, preferring to enjoy the delights of the game from the comfort of his armchair!

Many A-level mathematics syllabuses now contain a reference to iterative methods, e.g. Newton–Raphson. Whilst such techniques prove immensely useful in undergraduate studies, their application at school level is somewhat limited. The following example, therefore, may prove of general interest to readers as well as to those currently learning about or teaching 'iterative methods'.

Consider the equation

$$\sin x = 0 \quad (x \in \mathbb{R}) \tag{1}$$

whose solutions are

$$x = n\pi \quad (n \in \mathbb{Z}). \tag{2}$$

A straightforward application of the Newton–Raphson procedure

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \tag{3}$$

where x_0 is an initial guess, to equation (1) with $f(x) = \sin x$ gives the iteration

$$x_{n+1} = x_n - \tan x_n. \tag{4}$$

In view of the solutions (2), we expect (when convergence is obtained) the iteration (4) to converge to an integral multiple of π; although the zero multiple is clearly not of great interest! Thus we have a simple example of an iteration that leads to a concrete solution, i.e. π. I believe that this is more interesting than the frequently used examples, e.g. finding the roots of $x^3 - 2x + 1 = 0$, for the following reasons.

Firstly, as we all know, π has a certain fascination amongst students of mathematics, whatever their age, and its frequent appearance in problems involving calculus makes it a very familiar number. Secondly, there are many other ways of approximating π and, as we shall see, (4) competes well with these. Finally, the gradual appearance of the first few significant digits of π in an iteration has that kind of magical quality not to be missed.

We note in passing that (4) could more easily have been found by employing the technique used to derive the general iteration (3), namely set

$$\begin{aligned} 0 = \sin x_{n+1} &= \sin(x_n + h) \\ &= \sin x_n \cos h + \sin h \cos x_n \\ &\simeq \sin x_n + h \cos x_n \end{aligned}$$

(where $\cos h \simeq 1$ and $\sin h \simeq h$, for small h), so that, to linear terms in h,

$$x_{n+1} = x_n + h = x_n - \tan x_n,$$

which is a useful alternative for those not familiar with Newton–Raphson.

Two questions now present themselves.

(i) For what values of x_0 does the iteration (4) converge to π?

(ii) How quickly does the iteration (4) converge?

The first question is the more technically demanding of the two, and after some analysis it can be shown that, for $x_0 \in [1.98, 4.73]$ the iterates of (4) converge to π, i.e. $x_n \to \pi$ as $n \to \infty$. (For readers who are interested, we need (4) to be a contraction, i.e. $|x_{n+1} - \pi| < |x_n - \pi|$, which will give the required interval.)

The second question is most easily answered by a demonstration, and the results in table 1 represent the iteration (4) using different starting values x_0:

Table 1

$x_0 = 1.98$	$x_0 = 2.0$	$x_0 = 2.5$	$x_0 = 3.0$
4.285 822 24	4.185 039 86	3.247 022 30	3.142 546 54
2.085 467 40	2.467 893 68	4.141 200 28	3.141 592 65
3.853 790 87	3.266 186 28	3.141 592 65	3.141 592 65
2.990 432 73	3.140 943 91	3.141 592 65	
3.142 754 58	3.141 592 65		
3.141 592 65	3.141 592 65		
3.141 592 65			
$x_0 = 3.1$	$x_0 = 3.5$	$x_0 = 4.0$	$x_0 = 4.73$
3.141 616 65	3.125 414 36	2.842 178 72	2.014 152 12
3.141 592 65	3.141 594 07	3.150 872 94	4.119 918 09
3.141 592 65	3.141 592 65	3.141 592 39	2.634 343 44
	3.141 592 65	3.141 592 65	3.190 096 24
		3.141 592 65	3.141 554 58
			3.141 592 65
			3.141 592 65

It is clear that the convergence is remarkably rapid, and this provokes the question as to why this is so. However, if we define $e_n = x_n - \pi$ for all $n \geqslant 0$, then, from (4),

$$e_{n+1} + \pi = e_n + \pi - \tan(\pi + e_n),$$

i.e.

$$\begin{aligned} e_{n+1} &= e_n - \tan(\pi + e_n) \\ &= e_n - \tan e_n \\ &= e_n - (e_n + \tfrac{1}{3}e_n^3 + \dots) \\ &\simeq -\tfrac{1}{3}e_n^3, \end{aligned}$$

using a series expansion for $\tan e_n$, i.e. a cubic rate of convergence, so that the convergence rate as observed is perhaps not quite so surprising.

Finally, I should like to point out that all the details above are appropriate for school work, and are held together by one of the most intriguing of numbers, π. Readers may like to devise other iterations for finding π, although the speed of convergence may be hard to beat!

Nice Polynomials of Degree 4

CHRIS K. CALDWELL, *University of Tennessee at Martin*

The author currently enjoys teaching mathematics at the University of Tennessee at Martin. He lives on a small farm with his wife, four children, and numerous animals.

1. Introduction

An old exercise in calculus classes was to draw the graph of a polynomial using its roots and its local extrema (located at the roots of the derivative). Today's calculus students are often asked to draw the graphs of polynomials using the roots of the first and second derivative (to find the intervals of increase and decrease, and the intervals of upward and downward concavity.) In both cases the arithmetic is simplified if all the roots involved are integers, that is, we want polynomials $P(X)$ for which both $P(X)$ and the derivative $P'(X)$ have distinct integer roots. We call these *nice polynomials* (short for nice polynomials to put on calculus tests).

Nice polynomials of degree 2 are easy to find (just make the sum of the roots of $P(X)$ even). Those of degree 3 have been studied (see reference 7), and the interest continues (reference 1). It has been asked if nice polynomials exist for every degree (references 3 and 4), but this query was answered positively only for degree 3 (reference 6). Here we answer yes for degree 4, finding all of the symmetric nice quartics and infinitely many non-symmetric nice quartics.

2. Nice quartics

Let $P(X)$ be any quartic nice polynomial,

$$P(X) = aX^4 + bX^3 + cX^2 + dX + e.$$

Dividing $P(X)$ by its leading coefficient does not alter its roots, so we may assume the polynomial is monic ($a = 1$). Recall that monic polynomials with integral roots have integral coefficients, so the coefficients of $\frac{1}{4}P'(X)$ are integral, and

$$\tfrac{1}{4}P'(X) = X^3 + \tfrac{3}{4}bX^2 + \tfrac{1}{2}cX + \tfrac{1}{4}d.$$

In particular, 4 divides b. The substitution $X = X' - \frac{1}{4}b$ translates the roots of $P(X)$ by the integer $\frac{1}{4}b$, and the coefficient of X'^3 is 0 (the coefficient of X^3 is the negative of the sum of the four roots of $P(X)$, see reference 5). This translation also changes the coefficients of the lower-degree terms, but they remain integers. This shows that it is sufficient to consider polynomials with the following special form:

$$P(X) = X^4 + cX^2 + dX + e.$$

A computer search found hundreds of examples of these nice quartics, but in every case d was zero. That is, all the small nice quartics are symmetric about the origin. In section 3 we find all symmetric nice quartics. In the final section we return to the non-symmetric case, where $d \neq 0$.

3. Symmetric nice quartics

Symmetric polynomials have $d = 0$, so that

$$P(X) = X^4 + cX^2 + e,$$

where c and e are integers. The derivative, $P'(X) = 2X(2X^2+c)$, also has integral roots, so c must equal $-2T^2$ for some integer T. Using the quadratic formula (by treating $P(X)$ as a quadratic in X^2) we find that

$$X^2 = T^2 \pm \sqrt{T^4 - e}.$$

This shows that $T^4 - e$ must be the square of some integer S. Let the roots of $P(X)$ be $\pm N$ and $\pm M$. Then

$$N^2 = T^2 + S, \qquad M^2 = T^2 - S.$$

Adding these two equations we find that

$$2T^2 = N^2 + M^2.$$

This is the quadratic diophantine equation that must be solved to find three squares in arithmetical progression ($M^2 + r = T^2 = N^2 - r$). Dickson (reference 2) devotes an entire chapter to this subject, mentioning that Diophantus, Fermat, Vieta and Euler each worked with this equation. Dickson gives the general solution as

$$T = v^2 + u^2, \qquad N = v^2 - 2uv - u^2, \qquad M = v^2 + 2uv - u^2,$$

where u and v are any integers. To find the primitive solutions (those for which N, M and T are relatively prime) we must take u and v to be relatively prime, with one of them even and the other odd.

The four roots of our nice quartic $P(X)$ are now given by

$$\pm(v^2 - 2uv - u^2), \qquad \pm(v^2 + 2uv - u^2),$$

and the critical numbers (roots of $P'(X)$) are

$$0, \qquad \pm(v^2 + u^2).$$

The first few of these are listed in table 1.

Can the second derivative of these polynomials,

$$P''(X) = 12X^2 - 4(v^2 + u^2)^2,$$

have integral roots? To see that the answer is no, notice that the restrictions on u and v do not allow $P''(X)$ to have 0 as a root, so we may

Table 1. Primitive symmetric nice quartics

u	v	$P(X)$	$P'(X)$
1	2	$(X+7)(X+1)(X-1)(X-7)$	$4X(X+5)(X-5)$
2	3	$(X+17)(X+7)(X-7)(X-17)$	$4X(X+13)(X-13)$
1	4	$(X+23)(X+7)(X-7)(X-23)$	$4X(X+17)(X-17)$
3	4	$(X+31)(X+17)(X-17)(X-31)$	$4X(X+25)(X-25)$
2	5	$(X+41)(X+1)(X-1)(X-41)$	$4X(X+29)(X-29)$
1	6	$(X+47)(X+23)(X-23)(X-47)$	$4X(X+37)(X-37)$

rewrite the equation $P''(X) = 0$ as follows:

$$3 = \frac{(v^2+u^2)^2}{X^2}.$$

The square root of 3 is not rational, so this equation cannot have a solution in integers u, v and X.

4. General nice quartics

After failing to prove that there were no non-symmetric nice quartics, I greatly expanded my computer search and found the five polynomials listed in table 2. Of course, these can be made into infinite families of nice quartics using any combination of the following: reflecting the polynomials about the y-axis, multiplying the polynomials by a constant, multiplying all the roots by an integral constant and/or translating the roots by any integral constant. Thus if $P(X)$ is a nice quartic and a, b and c are integers, then

$$a^4cP\left(\frac{X+b}{a}\right)$$

is also a nice polynomial.

Table 2. Non-symmetric nice quartics

$$\begin{aligned}
P(X) &= (X+9817)(X+1307)(X-2741)(X-8383)\\
P'(X) &= 4(X+6931)(X-648)(X-6283)\\[1ex]
P(X) &= (X+17\,479)(X-3227)(X-6571)(X-7681)\\
P'(X) &= 4(X+11\,707)(X-4536)(X-7171)\\[1ex]
P(X) &= (X+29\,537)(X+6553)(X-10\,793)(X-25\,297)\\
P'(X) &= 4(X+21\,253)(X-1728)(X-19\,525)\\[1ex]
P(X) &= (X+46\,451)(X-1079)(X-10\,973)(X-34\,399)\\
P'(X) &= 4(X+31\,901)(X-5832)(X-26\,069)\\[1ex]
P(X) &= (X+77\,959)(X-15\,313)(X-26\,687)(X-35\,959)\\
P'(X) &= 4(X+52\,205)(X-20\,160)(X-32\,045)
\end{aligned}$$

References
1. P. A. Batnik, Problem E3221, *American Mathematical Monthly*, **94** (1987), 681.
2. L. E. Dickson, *History of the Theory of Numbers, Vol. II: Diophantine Analysis* (Chelsea, New York, 1966).
3. F. Schmidt, Problem 87–9, *Mathematical Intelligencer*, **2** (8) (1986), 48.
4. F. Schmidt, Problem 87–9 [reposed], *Mathematical Intelligencer*, **3** (9) (1987), 40.
5. J. V. Uspensky, *The Theory of Equations* (McGraw–Hill, New York, 1948), 43–44.
6. S. H. Weintraub, Partial solution to problem 87–9, *Mathematical Intelligencer*, **2** (10) (1988), 55. (Degree 3 only.)
7. K. Zuser, Uber eine gewisse Klasse von ganzen rationalen Funktionen 3. Grades, *Elemente der Mathematik*, **18** (1963), 101–104.

An Interesting Dull Real Number

MICHAEL DAVIES, *Westminster School*

In *Mathematical Spectrum*, Volume 22 Number 3, Nick Mackinnon proves that there are uncountably many dull real numbers by demonstrating that there can be at most a countable number of interesting ones, defined to be those which can be described by a finite English phrase. His method involves assigning to each finite English phrase a unique natural number, which we shall call its 'Mackinnon number': the set of interesting real numbers is thus put into one-to-one correspondence with a subset of the natural numbers, and so is at most countable.

There is a problem with this, however. Suppose we list the (countable) set of interesting real numbers, in increasing order of Mackinnon number, and then construct a further real number thus: start with a decimal point; the first digit after the decimal point should be one greater (modulo 10, so that $9+1=0$) than the first decimal digit of the first interesting real in our list; the second digit after the decimal point should be one greater (modulo 10) than the second decimal digit of the second interesting real in our list; the third digit after the decimal point should be one greater (modulo 10) than the third decimal digit of the third interesting real in our list and so on. This is a well defined process, and gives us a real number (between 0 and 1) which is interesting, because I have just described it in a finite English phrase (the second sentence of this paragraph). Unfortunately, this real number differs from the first number in our list of interesting reals, because they have different first decimal digits; it differs from the second number in our list of interesting reals, because they have different second decimal digits; it differs from the third number in our list of interesting reals, because they have different third decimal digits and so on: in fact it is not in the list of interesting reals at all. The assumption that a list can be made of the interesting reals (as defined by Mackinnon) thus leads to a contradiction.

A frequent response to a first meeting with an argument like this (it is called a *diagonal argument*) is: 'Why not just add this extra interesting number to the list, then? There are two answers to this: the first is that our argument does not just demonstrate that a particular list of interesting reals, claimed to be complete, has a particular interesting real missing, but that any list of interesting reals will always lead inexorably to an interesting real which is not in the list, and thus that a complete list of interesting reals is impossible. This would hold in any application of the diagonal argument, but in this case, there is a second answer which gives an added twist because if you add the newly constructed interesting real to the incomplete list the finite English phrase which constitutes the second sentence of the last paragraph will then define a different real number (if you add the new number in as the nth in the list, the real number defined by the phrase must differ from it in the nth decimal place) and so the real number you have just inserted is no longer defined by a finite English phrase and so needs to be removed—it has become dull again: in fact we can say that this number is interesting and so should be in the list of interesting numbers if and only if it is not in the list of interesting numbers, which brings out the resemblance of this conundrum to the famous 'Barber Paradox'.

It is not clear whether Mackinnon intends to call a real interesting if and only if it can be described by a finite English phrase, but in fact, his proof that there are uncountably many dull reals only requires us to assume that each interesting number can be described by a finite English phrase, not that each number described by a finite English phrase is interesting. If we decide that he did not intend us to make this converse assumption, then the argument of the second paragraph in fact produces a dull real number (describable in a finite English phrase), contradicting the second part of his proposition, that no dull real number can be exhibited—and this of course makes such a dull number rather interesting ....

Sharp readers may have noticed a slight problem with my original argument: it is not necessarily true that two numbers whose decimal representations differ in the nth place are different; for instance, 0.1000... and 0.0999.... To cope with this, construct my new number using the rule 'make its nth decimal digit equal to one more than the nth digit of the nth interesting real, except that if the nth digit is eight, replace it with seven, and if the nth digit is nine, replace it with eight', so that this number will contain no zeros or nines in its decimal expansion, and hence not be affected by the minor ambiguity in decimal notation.

Perception of Speed on the Motorway

RICHARD STONE, *CSIRO Division of Mathematics and Statistics, Melbourne*

After toiling away in laboratories to graduate in science, the author underwent a conversion to become a statistician. He currently does statistical research and consulting for the Australian government, and can now share in the excitement of scientific discovery without personally messing up any experiments. His interests include causal inference and industrial statistics.

When driving along a motorway, have you ever considered whether your speed is typical of the rest of the traffic, or if you are travelling unusually quickly or slowly? If you travel at 60 miles per hour on an uncongested motorway, you may well have the impression that a surprising proportion of the vehicles are travelling very much faster than you. Conversely, if you race along at 85 miles per hour, you may be surprised by how many vehicles are travelling much slower.

Using a simple model for how we perceive the speed of other vehicles, together with some basic assumptions about traffic flow, these sorts of observations can be given a mathematical explanation.

The model for perception is that the observer notices the speeds of those vehicles which overtake or are overtaken by the observer. These are the vehicles that enter or leave our field of vision. The main assumption about traffic flow is that the presence of the observer's vehicle does not affect the speed of nearby vehicles. For an uncongested motorway, this may not be unrealistic. We suppose that vehicles pass a fixed point on the motorway at a total rate λ per second and that the 'true' probability density of vehicles travelling at speed v miles per second is $f_0(v)$. This means that vehicles in the speed range $(v, v+\mathrm{d}v)$ pass a fixed point at a rate $\lambda f_0(v)\,\mathrm{d}v$ per second.

Now the observer's vehicle is not stationary at a fixed point, but is travelling at some speed u miles per second. We can determine the rate at which the observer is overtaken by vehicles in the speed range $(v, v+\mathrm{d}v)$ by considering the spatial density of these vehicles, which is $\lambda f_0(v)\,\mathrm{d}v/v$ per mile. The observer's vehicle moves at a speed $v-u$ relative to this stream of vehicles and so is overtaken by them at a rate $(v-u)\lambda f_0(v)\,\mathrm{d}v/v$ per second. Similarly, if the observer's speed u is greater than v, these vehicles would be overtaken at a rate $(u-v)\lambda f_0(v)\,\mathrm{d}v/v$ per second. Overall, the observer overtakes, or is overtaken by, vehicles in the speed range $(v, v+\mathrm{d}v)$ at a rate

$$\frac{|v-u|\lambda f_0(v)\,\mathrm{d}v}{v} = \left|1-\frac{u}{v}\right|\lambda f_0(v)\,\mathrm{d}v \quad \text{per second.}$$

The perceived probability density of speeds is therefore given by

$$f_u(v) = k_u \left| 1 - \frac{u}{v} \right| f_0(v), \tag{1}$$

where k_u is a normalising constant.

Note that the density drops to zero for vehicles travelling at the same speed as the observer, since these vehicles never overtake, nor are overtaken.

The effects of the distortion given by equation (1) may be illustrated for a motorway, where average speeds of 70 miles per hour are quite plausible. The true speed distribution is taken as Gaussian (normal) with mean 70 miles per hour and standard deviation 10 miles per hour and is plotted, together with the normalised perceived distribution, in figure 1. If all cars travelling faster than 80 miles per hour are classified as being 'very fast', and the observer travels at 60 miles per hour, then the proportion of very fast cars jumps from 16% to 30%.

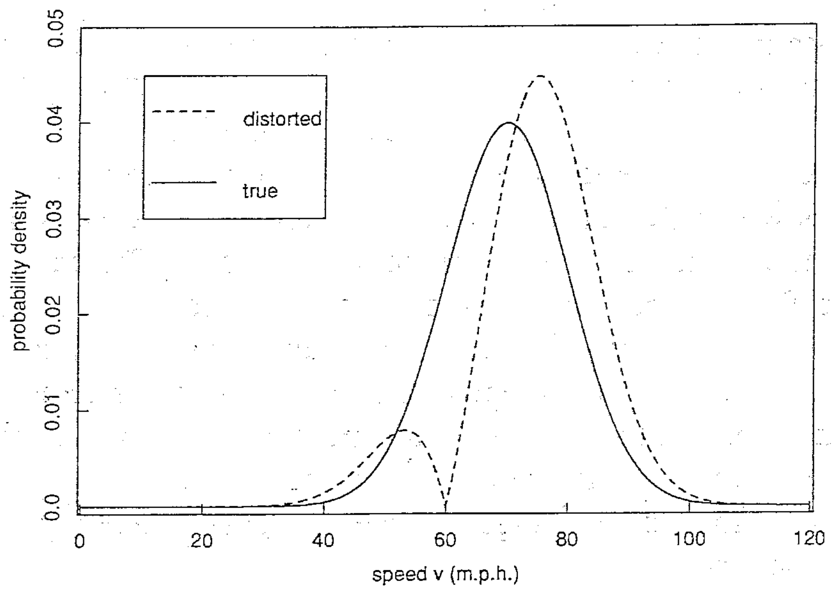


Figure 1 A 'true' and a distorted speed distribution when the observer's speed is 60 miles per hour

As may be expected, the upper tails of the speed distribution can be distorted either upwards or downwards, depending on what the observer's speed is. In figure 2 the perceived proportions exceeding various critical speeds are plotted as a function of the observer's speed, assuming the same 'true' speed distribution as above. Because the distribution given by equation (1) cannot be integrated analytically, the results have been obtained by numerical integration.

The distortion effects do not exhibit any simple sort of pattern that might allow us to make mental adjustments to the perceived speed

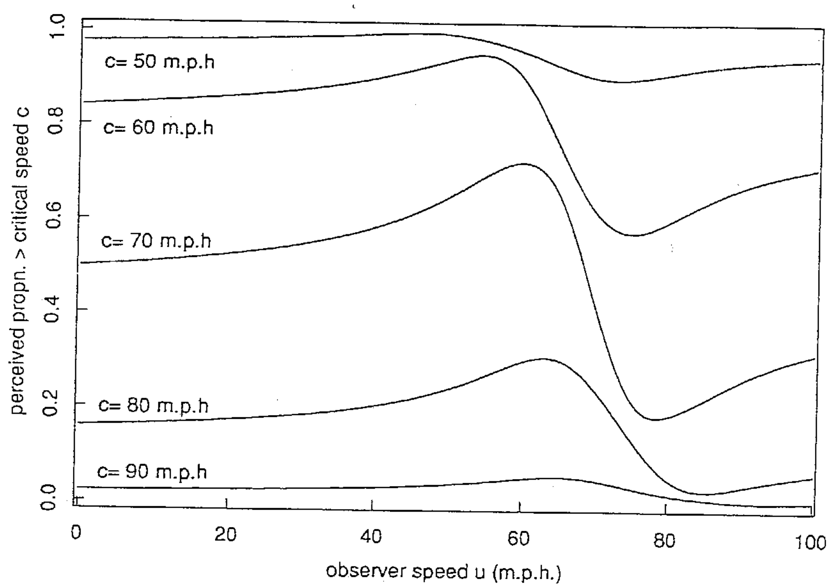


Figure 2 Distortions as a function of the observer's speed

distribution while driving. This absence of simple rules for adjustment perhaps explains how it is possible to be surprised by our perceptions.

It may be noted that, for a given critical speed c miles per second, the perceived proportion of vehicles travelling faster than c is undistorted at $u = 0$ and at one other value which may be higher or lower than c. Also, as u increases, the term $|1-(u/v)|$ tends towards $|u/v|$. In the limit (which is well beyond the legal speed limit!), the perceived distribution becomes

$$f_u(v) = k\frac{f_0(v)}{v}, \tag{2}$$

which is independent of u. Thus the upper tails of the distribution become thinner, and the lower tails become fatter.

These results have more than just curiosity value. For example, they may reassure the economically disadvantaged who drive old cars that they are not quite as slow as they thought, and so improve their self-esteem. Similarly, it may serve to remind those who speed at 85 miles per hour that they are not quite as exclusive as they may like to believe.

If

$$u_{n+1} = \frac{2u_n - 2}{u_n}, \qquad u_1 = 3,$$

can you find u_n for all values of n?

ALAN FEARNEHOUGH
Portsmouth Sixth Form College

I. Infinity and Enumeration

JOSEPH ROSENBLATT, *Ohio State University*

Professor Rosenblatt is currently on the faculty at the Ohio State University, Columbus. He was an undergraduate student at Reed College (1964–68) and a graduate student at the University of Washington (1968–72) where he wrote his Ph.D. thesis under the direction of Professor Isaac Namioka. His major research interests are in harmonic analysis and ergodic theory.

1. The philosophy of infinity

The infinite has been part of our thinking from our earliest years as a species. Anyone looking up at a night's moonless sky and seeing the Milky Way has surely shared the feelings of our earliest ancestors, that there in the heavens was created before our eyes an infinity of worlds, worlds more numerous than the number of grains of sand on all the beaches on all the shores. But even though we might be in awe of a very large number, it does not make this number a true representative of the infinite.

The symbol ∞ for infinity was first employed by Wallis in 1655 in his *Arithmetica Infinitorum*; it does not occur again until 1713, when it is used in James Bernoulli's *Ars Conjectandi.* The sign was sometimes employed by the Romans to denote the number 1000, and it has been conjectured that this led to its being applied to represent any very large number. This is at best only an improvement in kind over the concept of number in some primitive cultures where 1 and 2 were recognized as numbers, but anything else would be called 'many'. But surely even a googol, $G = 10^{100}$, or worse still 10^G, is not really infinite. No matter how large a real quantity is, or how seemingly impossible to enumerate, a collection of objects in the natural world can only contain a finite number of elements.

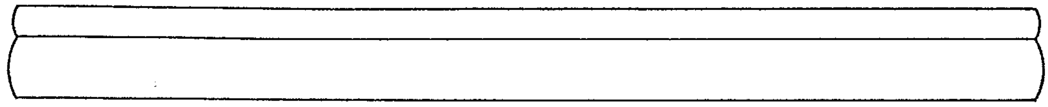
The pursuit of a large number as a representative for infinity is a hopeless one. As Norton Juster puts it in the chapter This Way to Infinity from *The Phantom Tollbooth*:

> 'Yes, please', said Milo. 'Can you show me the biggest number there is?'
>
> 'I'd be delighted,' [the Mathemagician] replied, opening one of the closet doors. 'We keep it right here. It took four miners just to dig it out.'
>
> Inside was the biggest
>
> # 3
>
> Milo had ever seen. It was fully twice as high as the Mathemagician.

'No, that's not what I mean,' objected Milo. 'Can you show me the longest number there is?'

'Surely,' said the Mathemagician, opening another door. 'Here it is. It took three carts to carry it here.'

Inside this closet was the longest



imaginable. It was just about as wide as the 3 was high.

'No, no, no, that's not what I mean either," he said, looking helplessly at Tock.

'I think what you would like to see,' said the dog, scratching himself just under half-past four, 'is the number of greatest possible magnitude.'

'Well, why didn't you say so?' said the Mathemagician who was busily measuring the edge of a raindrop. 'What's the greatest number *you* can think of?'

'Nine trillion, nine hundred ninety-nine billion, nine hundred ninety-nine million, nine hundred ninety-nine thousand, nine hundred ninety-nine,' recited Milo breathlessly.

'Very good,' said the Mathemagician. 'Now add one to it. Now add one again,' he repeated when Milo had added the previous one. 'Now add one again. Now add one again. Now add one again. Now add one again. Now add one again. Now add one again. Now add . . . .'

'But when can I stop?' pleaded Milo.

'Never,' said the Mathemagician with a little smile, 'for the number you want is always at least one more than the number you've got, and it's so large that if you started saying it yesterday you wouldn't finish tomorrow.'

Quoted with permission from *The Phantom Tollbooth* by Norton Juster, Copyright © 1961 by Norton Juster, published by Random House, 1964.

Historically, our earliest understanding of the infinite was probably religious in nature. God as the all-powerful supreme Creator is sometimes thought of as the embodiment of the infinite. Indeed, some of the early opposition to discussing the concept of infinity may have been motivated by the fear of being sacrilegious. But if we are too logical about this representation of the infinite, we do get into conceptual problems. If God is infinitely powerful, He could create an immovable rock. If the rock is immovable, God could not move it. If God could not move the immovable rock, then

God is not infinitely powerful. In such a paradoxical logical situation, one feels that the idea of infinity has somehow been misused. Another version of this same type of paradox of the infinite occurs in set theory. A set is a collection of things or members. For instance, the set of all people contains as members both men and women. As such, a set is a thing too! Some sets, like the set of dogs, are not members of themselves. But the set of all sets seems to be a set and so it is a member of itself. Consider momentarily the particular set G which is the set of all sets that are not members of themselves. This set G is a set too, I think. Is it a member of itself? Well, if it is, then by definition of G, it is not a member of itself. And if it isn't, then by definition it is! What a mess! Bertrand Russell was able to abolish this particular paradox of set theory by developing a theory of types, members of a set being of one type and sets being of a higher type, disallowing that a set can ever be a member of itself. In particular, there is no such set as 'the set of all sets'. With this understanding, the 'set' G here is the 'set' of all things and as such is not appropriately called a set at all. So here the problem that generated the paradox was that all things together comprise too much to be reasonably called a set.

In any case, when someone says that infinity is an arbitrarily large number or an arbitrarily large collection of things, they mean not any specific number, but a variable that could increase without bound. That is, in this sense infinity is a potential infinity, not an actual infinity. For instance, if the world population is going to increase by 10% each year without fail, then as time goes on this population will increase to infinity. Rather, as time tends to infinity, then the population size tends to infinity. Similarly, when someone says that they have in mind an arbitrarily small number, they don't mean any specific small number. They are just entertaining the possibility that the number they are thinking about may be smaller than you realize. The story is told of the famous mathematical team of G. H. Hardy and J. E. Littlewood receiving back from the printer the galley proofs of an article they had written for publication in a mathematical journal. They noticed on one line, in the middle of the line, a very small black speck that seemed to be intentional. After examining it through a magnifying glass, they discovered it was the traditional symbol ϵ (the Greek letter epsilon) used by mathematicians for an arbitrarily small quantity. When they asked the publisher why it had been put there, the publisher said, "The text read, 'Let ϵ be as small as possible." Clearly, the printer missed the point. Hardy commented later that the symbol they had used was so small that they must have scoured all of London to find it.

One fundamental belief that was held for many thousands of years about infinity was that there was nothing that could be thought of as an actual object in itself which was infinite. It was believed that there was not an actual infinity. The term infinity, or the expression, 'tending to infinity', was seen only as a way of speaking about more concrete objects which were in themselves finite. The only clear aspect of infinity that was agreed upon was that it was distinguishable from the finite. But there was certainly at most only one infinity and there could be no structure to the 'sizes' of infinite things. It was considered paradoxical that if there are infinitely many balls in a box, and one ball is removed, then there are still infinitely many balls in the box. Indeed, the basic word that is used to describe the impossibly large is a negation of something presumably more real; we say something is infinite if it is not finite (just as a thought is incomprehensible when it is not comprehensible). This bit of language reflects an ancient and fundamental antipathy to infinity. After all, we say plural for the opposite of singular, not inplural for singular or insingular for plural. The finite is seen as human, something that can be listed or comprehended. The infinite is seen as something inhuman, something that cannot be listed or is incomprehensible.

Not until the time of Georg Cantor in the nineteenth century did mathematicians begin to realize that it was both possible and reasonable to distinguish between different sizes of infinity and that it was possible to manipulate infinite numbers in a type of arithmetic bearing a vague resemblance to the arithmetic learned in school. This was a great step forward in the development of modern mathematics, in many ways as fundamental as the development of calculus some 200 years earlier. Cantor believed that opposition to the use of infinity in theology, philosophy, and mathematics was based upon a fundamental and pervasive error. Scholars had assumed that, if infinity existed as a 'number' itself, then it would have to act like finite numbers. So the fact that positive real numbers a and b satisfy $a+b > a$ and $a+b > b$ was well understood. But $\infty + a = \infty$ was not understood and was taken as evidence that infinity could not actually be conceptualized and that infinite sets could not be manipulated. Cantor's ability to separate himself from the traditional thinking of the past, and to entertain a new world of sets in which infinite sets were as reasonable to use as were finite sets, was the intellectual gift which allowed him to make his amazing discoveries in the late nineteenth century. To paraphrase David Hilbert's comment on Cantor's work, Cantor led us into the paradise of infinite sets from which we shall never be dislodged.

2. Infinity and enumeration

One of the failings of the mathematics of the Greeks was that they studiously avoided referring to the infinite or the infinitesimal. With the Renaissance in Europe, the notion of infinity was one of the first to be attacked. Galileo addressed the infinite in his *Dialogue Concerning the New Sciences,* which appeared in 1632. He asked us to consider all squares arranged in order, 1, 4, 9, 16, ... With this in mind, it is possible to count the squares by assigning a number to each; 1 is the first square, 4 is the second, 9 is the third, and so on. No matter how many numbers we may have, we can always find its square. The conclusion, in Galileo's words, is this, 'So far as I see, we can only infer that the number of squares is infinite and the number of their roots is infinite; neither is the number of squares less than the totality of all numbers, nor the latter greater than the former; and finally the attributes 'equal', 'greater', and 'less' are not applicable to infinite, but only to finite quantities. When, therefore, Simplicio introduces several lines of different lengths and asks how is it possible that the longer ones do not contain more points than the shorter, I answer him that the one line does not contain more, or less, or just as many, points as another, but that each line contains an infinite number.'

This passage from Galileo's work shows that he understood the idea of corresponding two sets element by element or term by term in order to compare them. However, Galileo still saw all infinite collections as being the same size: infinite. The idea of corresponding elements of sets was used by Cantor more systematically and he reached some fascinating conclusions as a result. We need some basic language from set theory to discuss this.

First, a function f always brings with it a pair of distinguished sets, one being its domain A and the other being its range B. We write $f: A \to B$ to denote that f associates to elements in A some elements in B. We also write $b = f(a)$ for the element in B corresponding to a in A, and assume that the function somehow unambiguously defines b when a is fixed. For instance, each person in a school has a height at a specific point in time, and this gives us a height function. But we have to fix a specific time to measure height, otherwise different heights might be appropriate for the same person (as they age for instance). When we say 'consider the function $f(x) = x^2$', we are usually thinking of the domain A consisting of real numbers and a range B of real numbers (although B could just be the non-negative real numbers).

A function f is said to be *one-to-one* (sometimes written 1–1) if two different elements in A cannot be mapped to the same point; formally, if $f(a_1) = f(a_2)$ then $a_1 = a_2$. The height function is not generally 1–1 because different people could have the same height. However, the function which assigns to a given person their right thumb print is 1–1, because thumb prints uniquely determine the person. Also, a function $f: A \to B$ is said to be *onto* if every element in B is the image under f of some point in A. That is, if b is in B, then there is some a in A with $f(a) = b$. The height function is certainly not onto. The function $f(x) = x^2$ is not onto if B is all real numbers. If B is just all non-negative real numbers then it is onto. Indeed, given any $b \geqslant 0$, there is a unique $a \geqslant 0$ such that $a^2 = b$.

Cantor's first contribution to the idea of infinity was to realize how to compare the sizes of sets that were possibly not finite. He used a notion that others before him had used, but he used it more extensively and systematically. Specifically, we say that A and B have the same 'size' (mathematicians say 'cardinality' without at first defining what cardinality is) if there is a 1–1 onto function f from A to B. Galileo had realized then that the whole numbers and the numbers that are squares have the same cardinality. But Cantor went further and pointed out the surprising fact that the set of whole numbers $\mathbb{N}$ and the set of rational numbers $\mathbb{Q}$ even have the same cardinality. It is not hard to see that it is enough to show that $\mathbb{N}$ and $\mathbb{Q}^+$, the positive rational numbers, have the same cardinality. But each positive rational number $x = p/q$ for some whole numbers p and q. To get a mapping $f{:}\mathbb{N} \rightarrow \mathbb{Q}^+$ that is 1–1 and onto, we use Cantor's diagonalization method. Consider the array of all the positive rational numbers (figure 1). Every positive rational number appears in the array above many times. But we can list the rational numbers by following the arrows along the diagonals and only list a given rational number once in the process. This will give a 1–1 onto mapping of $\mathbb{N}$ to $\mathbb{Q}^+$. This correspondence starts like this:

$$\begin{array}{l} \\ \downarrow \\ \end{array} \begin{array}{ccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & \ldots \\ \\ 1 & 2 & \frac{1}{2} & 3 & \frac{1}{3} & 4 & \frac{3}{2} & \frac{2}{3} & \frac{1}{4} & 5 & \frac{1}{5} & 6 & \ldots \end{array}$$

Is this mapping specific enough for you? The formula or pattern is clear, but it might be pretty hard to figure out which number is the 101st in the list! Can you come up with a simpler formula or function that does the same thing?

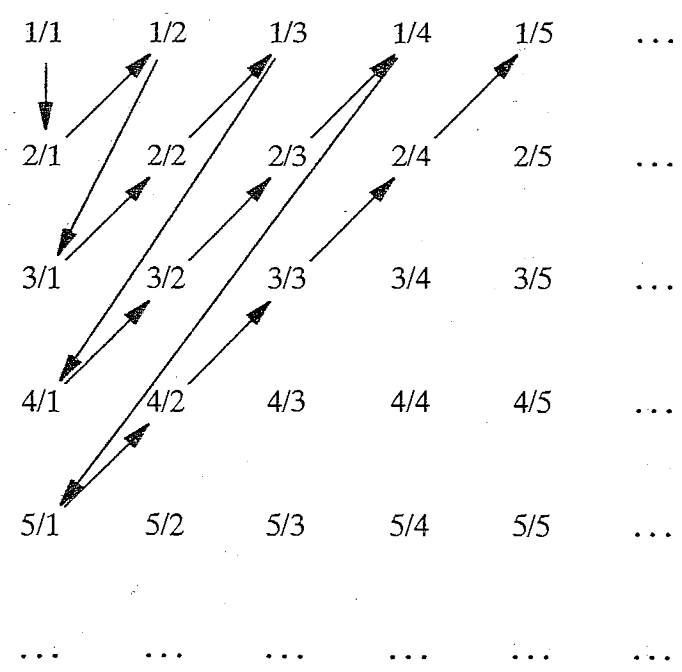


Figure 1

Cantor formalized the idea of cardinality by saying that a cardinal number is the class of all sets with the same cardinality. Also, if there is a 1–1 mapping of A into B, then B is said to have cardinality at least as large as A. The cardinality of $\mathbb{N}$ is the first infinite cardinal (there are many other infinite cardinals, as we shall see!) because any infinite set has cardinality at least as large as $\mathbb{N}$. Cantor chose a striking symbol for this smallest cardinal number. He denoted it by the first letter in the Hebrew alphabet, aleph. Since there are other sizes of infinity, it is denoted $\aleph_0$ (read 'aleph zero'). We also use the following terminology. A set is *finite* if it can be put in 1–1 onto correspondence with some set $\{1,2,\ldots,n\}$, where n is a whole number. A set is called *countable* if it is finite or has the cardinal value $\aleph_0$. It is called *countably infinite* if it can be counted as above by the whole numbers, i.e. it has the cardinality of $\mathbb{N}$.

In December of 1873, Cantor realized how to show that not every infinite set has the same cardinality as $\mathbb{N}$. He had written to another mathematician, Richard Dedekind, and conjectured that the set of all real numbers is not of the same cardinality as $\mathbb{N}$, but it took him a while to prove this. The proof given here is not Cantor's original one, but it is nevertheless a beautiful proof that embodies much of the style of modern analysis. To understand the proof, we first need to discuss what real numbers are and how we can represent them in decimals to the base 10.

There are many ways of writing numbers. The *integers* 0, 1, −1, 2, −2, ... (or as is sometimes written ... −3, −2, −1, 0, 1, 2, 3, ...) can be used to create many other numbers with a few operations. For instance, ratios p/q, where p and q are integers ($q \neq 0$), give us all the *rational numbers*. But, using square roots and other radical operations, we can create more numbers. For example, $\sqrt{2}$ is another number, approximately 1.414, which is not itself a rational number. Such a number is called an *irrational number*. Historically, irrational numbers took a great deal of time to gain acceptance as real numbers. There were even arguments about what a real number is. The use of decimals to base 10 to denote numbers essentially resolved this problem both technically and theoretically. It is not hard to see that each whole number $w = 1,2,3,\ldots$ can be written in a unique way as a finite sum

$$a_{n+1}10^n + \ldots + a_3 10^2 + a_2 10 + a_1,$$

where $a_1, a_2, \ldots, a_{n+1}$ are themselves restricted to be whole numbers between 0 and 9, excepting that a_{n+1} is not zero. For short, we write $w = a_{n+1}a_n \ldots a_2 a_1$. So, for example, $3\times10^2+1\times10+2$ is represented by 312. By the same token, we can write fractions as follows. We write $x = 0.b_1 b_2 b_3 \ldots b_n$ to represent the number $b_1/10 + b_2/10^2 + \ldots + b_n/10^n$, where again the numbers b_n are taken to be whole numbers between 0 and 9.

You can see then that the numbers of the form $a_{n+1}a_n \ldots a_1.b_1b_2b_3 \ldots b_n$ give us a great quantity of different numbers (an infinite number), but not all of the real numbers because numbers of this form are always rational. To get all real numbers we continue the decimal expansion to infinity $x = 0.b_1b_2b_3\ldots$ and combine this with the expression for w. This gives us all the real numbers. There are some problems with this representation. It is not easy to define sums and products of real numbers in this form and to see that the operations are completely well-defined because there can be infinitely many b_n which are not zero. Also, there is a non-uniqueness in this expansion. The number 1 can be represented as 1 or as 0.999999 etc...; we write this as $0.\bar{9}$ to indicate that the 9 repeats indefinitely. I recall one student who was very bothered by representing $1 = 0.\bar{9}$. We argued about it for a while and concluded this. He agreed that if $0.\bar{9}$ meant anything, then it had to be 1, and I agreed to let him think that $0.\bar{9}$ didn't have a well-defined meaning. Clearly, he didn't want to believe in infinite sequences of terms!

We are ready to understand Cantor's theorem that the real numbers are not countably infinite. The argument proceeds by contradiction. Suppose that the real numbers are countable. Then it is easy to see that the numbers between 0 and 1 are countable. Such an enumeration (a fancy word for 'list') is:

$$\begin{aligned} x_1 &= 0.a_{11}a_{12}a_{13}a_{14}\ldots \\ x_2 &= 0.a_{21}a_{22}a_{23}a_{24}\ldots \\ x_3 &= 0.a_{31}a_{32}a_{33}a_{34}\ldots \\ &\ldots \quad \ldots \quad \ldots \end{aligned}$$

The digits in the expansions here have to have double indexing because the digit depends on the decimal place as well as which number one has in the list. You can assume that the decimal expansions above do not end in a string of an infinite number of 9's in order to guarantee uniqueness of the expansion. We shall now construct a real number between 0 and 1 which is not in the list, thus contradicting the fact that this was supposed to be an enumeration of all the numbers between 0 and 1. Each digit is only one of 10 possible numbers. Choose $b_1 \neq a_{11}$, $b_2 \neq a_{22}$, $b_3 \neq a_{33}$, etc., making sure that we do not eventually keep choosing 9 by choosing all $b_n \neq 9$. Let $x = 0.b_1b_2b_3\ldots$. This number is not in the list. It is not x_1 since it differs from x_1 in the first decimal place. It is not x_2 because x and x_2 differ in the second decimal place. And so on! This contradiction shows why the real numbers in total form an infinite set which must be of a different order of infinity than the rational numbers. This was considered an amazing fact by all those who came to understand it. The argument itself shows the great power of the use of non-terminating decimal expansion to represent real numbers.

Cantor saw how to use this basic result of his to prove a very interesting theorem that was just becoming widely accepted in his generation. First, his argument clearly shows that there are real numbers that are not rational numbers. But we already knew this by a simpler argument. He went further. An *algebraic number* is a number x which is the solution of an equation of the form $a_0+a_1x+\ldots+a_nx^n = 0$, where the numbers a_n are integers, not all of which are zero. A *transcendental number* is one that is not algebraic. Cantor showed (it is not too hard to see) that the set of algebraic numbers (which includes all the rational numbers) is countably infinite. But then they cannot comprise all of the real numbers! So there must be transcendental numbers. Today we know by technical arguments that numbers like π and e are transcendental. But Cantor showed that in fact, in some sense, most real numbers are transcendental. Even though the algebraic numbers are dense in the real numbers (in fact, between any two distinct real numbers, there is a rational number), the rest of the numbers that lie between them are overwhelmingly more in the sense of cardinality.

Another use that Cantor made of 1–1 onto mappings to define cardinality was at first thought by him to destroy all meaning of dimension. He showed that the unit segment in the real numbers has the same cardinality as the unit square in the plane! We can represent the unit square U by Cartesian coordinates (x,y) where x and y are between 0 and 1. Cantor's original, slightly flawed argument to see the correspondence was this. Define a mapping $f\colon [0,1] \to U$ as follows. Write $w = 0.b_1b_2b_3b_4\ldots$ and then let $f(w)$ be the point (x,y) where $x = 0.b_1b_3b_5\ldots$ and $y = 0.b_2b_4b_6\ldots$. He claimed that his mapping is 1–1 and onto. To see that f is onto, write $x = 0.x_1x_2x_3x_4\ldots$ and $y = 0.y_1y_2y_3y_4\ldots$ and let w be the decimal $w = 0.x_1y_1x_2y_2x_3y_3x_4y_4\ldots$ Then $f(w) = (x,y)$. However, f is not 1–1; for example, $0.0090909\ldots$ and $0.1000\ldots$ have the same image under f. Cantor was able to correct his argument, however. He thought at first that this means that there is no real meaning in the notion of dimension, that the square is two-dimensional and that the unit segment is one-dimensional. However, Dedekind pointed out that the corrected version of Cantor's correspondence f above is not continuous. In fact, it is not hard to show the most basic of such results: there is no continuous mapping of $[0,1]$ into U which is 1–1, onto, and has a continuous inverse mapping, even though, as Cantor showed, there is such a mapping that is just 1–1 and onto.

After many more years of study, Cantor came to realize that what he had begun with the real line, the plane, the rationals and the irrationals, could be applied abstractly to all sets. He denoted the cardinality of the real numbers (often called the cardinality of the continuum) by $\mathfrak{c}$. Thus, we showed above that the rationals have smaller cardinality than the real numbers, i.e. $\aleph_0 < \mathfrak{c}$. Cantor also defined many other different types of infinite cardinals. You might think it would be very hard to see that there must be infinitely many infinite cardinals, but actually it is very easy. We only need to look briefly at what is sometimes called the *power set operation*, the operation of constructing the set of all subsets of a given set. For example, let $S = \{1, 2, 3\}$, which has three elements. The possible subsets are $\varnothing$ (the symbol for the empty set), $\{1\}$, $\{2\}$, $\{3\}$, $\{1,2\}$, $\{1,3\}$, $\{2,3\}$ and $\{1,2,3\}$. That gives eight subsets in all. You should be able to see that, if $S = \{1, 2, 3, \ldots n\}$, then the power set has 2^n elements. If S is a set, the power set is usually denoted by 2^S partly for this reason.

Now every point s in S corresponds to $\{s\}$ in 2^S, so 2^S has cardinality at least as large as S. But why can't they have the same cardinality? To see this, suppose that there is a 1–1 onto mapping $f\colon S \to 2^S$. We shall construct a set which is not in the range of f, contradicting the fact that f is onto. Let S_0 be the set of all elements s in S such that s is not in $f(s)$. This sounds familiar, doesn't it? Suppose s_0 is in S and $f(s_0) = S_0$. If s_0 is not in S_0, then, by definition of S_0, it is in S_0. But if s_0 is not in S_0, then by definition it is in S_0. This impossible situation shows that there cannot be such an s_0! So 2^S has a greater cardinality than S. Thus, if we take the set of all subsets of the real numbers, then we have an infinite set which is truly bigger than the real numbers themselves. But we can continue taking the power set of the power set of the power set, indefinitely. So we can create in this fashion an infinite number of ever greater cardinal numbers.

Another good use that Cantor made of the notion of 1–1 onto mappings was to see how to define infinite sets directly. The pigeon-hole principle, as it is colourfully called, is that if you have more objects than boxes and put all the objects into the boxes, then some box contains more than one object. That is, if F_1 is finite and F_2 is a proper subset of F_1, then, for any mapping $f\colon F_1 \to F_2$, some two different x and y in F_1 have $f(x) = f(y)$. So, there is no 1–1 mapping of a finite set onto a proper subset of itself. On the other hand, Cantor observed that any infinite set S does have such a mapping. Indeed, if s is in S, then there is a 1–1 mapping of S onto $S \setminus \{s\}$. Because of this, one can define a set to be infinite if it has the same cardinality as a proper subset of itself, while a set is finite if it is not infinite. This certainly turns the tables on the usual negative definition of infinite as non-finite.

One of the great problems of set theory that Cantor was not able to solve (and he sorely wished that he could have solved it) was to determine whether $\mathfrak{c}$ was the next cardinal greater than $\aleph_0$. He was able to show that $\mathfrak{c}$ has the cardinality of the power set of the rationals. But could there be another infinite cardinal between $\aleph_0$ and $\mathfrak{c}$? To say that there is not became known as the *continuum hypothesis*. Not until the work of Kurt Gödel was this at least partly resolved. He showed that it is consistent with the set theory of the time to assume that the continuum hypothesis is true. But then later, Paul Cohen showed that it is consistent with set theory to assume that there are other cardinals between $\aleph_0$ and $\mathfrak{c}$. This shows that the continuum hypothesis is really an axiom, to be assumed or not assumed depending on your needs. It does not follow, as Cantor had thought, from simpler and more fundamental principles.

We have seen now that there can be structure to infinity. We have seen some of what Cantor did to show that there is a pantheon (some would have it called a zoo) of different infinities. Cantor showed that you can in a suitable sense add and multiply these infinite cardinals as if they are numbers. The arithmetic you get is a little bit strange, but it is consistent and understandable after some reflection on the basic principles. But one of Cantor's most fundamental contributions to mathematics, philosophy, and science was to change our preconceptions from what they had been, to force us to include a richer world of sets and a greater set theory than anyone had thought possible up until that time.

Reference

J. Dauben, *George Cantor, His Mathematics and Philosophy of the Infinite* (Harvard, University Press, Cambridge, 1979).

Computer Column

MIKE PIFF

Permutations

A permutation can be considered as a finite 'function' defined only on the numbers 1 to n, taking the same values once and once only. Thus, it can be specified as a list of values, such as

$$\phi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 1 & 3 & 9 & 6 & 5 & 8 & 10 & 4 & 7 \end{pmatrix},$$

where $\phi(1) = 2$, $\phi(4) = 9$, and so on.

Any such permutation cyclically permutes certain subsets of the numbers. Thus, that above cyclically permutes 1 and 2; 7, 8 and 10; and so on. These subsets are called the *cycles* of the permutation. Also, all permutations are either *even* or *odd* depending on the lengths of these cycles, by a formula which can be deduced from the following program. The *sign* is $+1$ or -1 to indicate even and odd.

When the above permutation is input, the program indicates that it is even, and produces a list of its cycles.

```
MODULE Cycles;
FROM InOut IMPORT OpenInput, ReadInt,
  WriteInt, CloseInput, WriteLn,
  WriteString, Write;
CONST
  minelt=1; maxelt=100; mone=-1; space=' ';
TYPE
  elts=INTEGER[minelt..maxelt];
  eltarrays=ARRAY elts OF INTEGER;
  permutations=RECORD
    nrfelts:elts;
    elt:eltarrays;
    sgn:[-1..1];
  END;
VAR
  p:permutations;
PROCEDURE ReadPerm(VAR p:permutations);
VAR
  i:elts;
  j:INTEGER;
  PROCEDURE Initialise(VAR p:permutations);
  VAR
    i:elts;
  BEGIN
    WITH p DO
      FOR i:=minelt TO nrfelts DO
        elt[i]:=i;
      END;
      sgn:=1;
    END;
  END Initialise;
  VAR
    ok:ARRAY elts OF BOOLEAN;
BEGIN
  WITH p DO
    WriteString('Give number of elements: ');
    ReadInt(j);nrfelts:=elts(j);WriteLn;
    WriteString
      ('Type the bottom line of permutation,');
    WriteString(' followed by return:');WriteLn;
    Initialise(p);
    FOR i:=minelt TO maxelt DO
      ok[i]:=TRUE;
    END;
    FOR i:=minelt TO nrfelts DO
      ReadInt(j);
      IF ok[j] THEN
        IF j>nrfelts THEN
          WriteString('Too big!');HALT;
        END;
        elt[i]:=j;
        ok[j]:=FALSE;
      ELSE
        WriteString('Repeated element!');WriteLn;
        HALT;
      END;
    END;
  END;
END ReadPerm;
PROCEDURE Tag(VAR p:permutations);
VAR
  i:elts;
  j,k:INTEGER;
BEGIN
  WITH p DO
    FOR i:=minelt TO nrfelts DO
      j:=elt[i];
      WHILE j>i DO
        k:=elt[j];
        elt[j]:=-k;
        j:=k;
        sgn:=-sgn;
      END;
      elt[i]:=-elt[i];
    END;
  END;
END Tag;
PROCEDURE WritePerm
  (VAR p:permutations);
VAR
  i:elts;
BEGIN
  WITH p DO
    FOR i:=minelt TO nrfelts DO
      WriteInt(i,4);
    END;
    WriteLn;
    FOR i:=minelt TO nrfelts DO
      WriteInt(elt[i],4);
    END;
    WriteLn;
    WriteString('Sign is ');
    WriteInt(sgn,1);WriteLn;
  END;
END WritePerm;
PROCEDURE WriteCycles
  (VAR p:permutations);
VAR
  i:elts;
  j:INTEGER;
BEGIN
  WITH p DO
    FOR i:=minelt TO nrfelts DO
      j:=elt[i];
      IF j<0 THEN
        j:=ABS(j);
        WriteString('(');WriteInt(i,1);
        WHILE j#i DO
          Write(space);WriteInt(j,1);
          j:=ABS(elt[j]);
        END;
        WriteString(')');
      END;
    END;
    WriteLn;
  END;
END WriteCycles;
BEGIN
  ReadPerm(p);WriteLn;
  WriteString('Permutation is');WriteLn;
  WritePerm(p);WriteLn;
  Tag(p);
  WriteString('Cycle structure is:');
  WriteCycles(p);
END Cycles.
```

Binomial Identities by Leibniz's Theorem

AMITES SARKAR, *Winchester College*

The author wrote this article in the first year of the sixth form. He is now at Trinity College, Cambridge, reading mathematics.

Not so long ago, I found myself differentiating $w = x^{m+n}$ using Leibniz's rule. The problem was to find the ath derivative of w for $1 \leqslant a \leqslant n$, $1 \leqslant a \leqslant m$, and I was feeling less than astute. I put $y = x^m$ and $z = x^n$, so that, from Leibniz's theorem,

$$\frac{\mathrm{d}^a w}{\mathrm{d}x^a} = \sum_{r=0}^{a} \binom{a}{r} \frac{\mathrm{d}^r y}{\mathrm{d}x^r} \frac{\mathrm{d}^{a-r} z}{\mathrm{d}x^{a-r}} = \sum_{r=0}^{a} \binom{a}{r} \frac{m!\,x^{m-r}}{(m-r)!} \frac{n!\,x^{n-a+r}}{(n-a+r)!}$$

$$= \sum_{r=0}^{a} \binom{m}{r}\binom{n}{a-r} a!\,x^{m+n-a}.$$

This is all very well, except that everyone knows that the 'correct' answer is given by

$$\frac{\mathrm{d}^a w}{\mathrm{d}x^a} = \frac{(m+n)!\,x^{m+n-a}}{(m+n-a)!}.$$

Comparing results, we see that

$$\binom{m+n}{a} = \sum_{r=0}^{a} \binom{m}{r}\binom{n}{a-r}.$$

Of course, there are several other ways of getting this identity. If we expand both sides of $(1+x)^{m+n} = (1+x)^m(1+x)^n$ and compare coefficients of x^a, it will do the trick, as will counting the number of subsets containing a elements of a set with $m+n$ elements in two ways. It is a matter of taste. If we put $m = n = a$, we obtain

$$\binom{2n}{n} = \sum_{r=0}^{n} \binom{n}{r}^2.$$

Readers might like to try using the above method with $w = x^{m-n}$ for $m > n$, writing it as $x^m \times x^{-n}$.

Other, more obscure, identities can be similarly generated. For example, let $w = x^n\sqrt{x}$, and write $y = x^n$ and $z = \sqrt{x}$. Then

$$\frac{\mathrm{d}^n w}{\mathrm{d}x^n} = (n+\tfrac{1}{2})[(n-1)+\tfrac{1}{2}][(n-2)+\tfrac{1}{2}]\ldots(1+\tfrac{1}{2})\sqrt{x} = \frac{(2n+2)!\,\sqrt{x}}{2^{2n+1}(n+1)!}.$$

By Leibniz's theorem,

$$\begin{aligned}
\frac{\mathrm{d}^n w}{\mathrm{d}x^n} &= \sum_{r=0}^{n} \binom{n}{r} \frac{\mathrm{d}^r y}{\mathrm{d}x^r} \frac{\mathrm{d}^{n-r} z}{\mathrm{d}x^{n-r}} \\
&= \sum_{r=0}^{n-1} \binom{n}{r} \frac{n!\,x^{n-r}}{(n-r)!} \frac{(-1)^{n-r-1}(2n-2r-2)!\,x^{r-n+\frac{1}{2}}}{2^{2n-2r-1}(n-r-1)!} + n!\sqrt{x} \\
&= \sqrt{x} \sum_{r=0}^{n-1} n! \binom{n}{r} \binom{2n-2r-2}{n-r-1} \frac{(-1)^{n-r-1}}{(n-r)2^{2n-2r-1}} + n!\sqrt{x}.
\end{aligned}$$

Therefore

$$\frac{(2n+2)!\sqrt{x}}{2^{2n+1}(n+1)!} = \sqrt{x} \sum_{r=0}^{n-1} n! \binom{n}{r} \binom{2n-2r-2}{n-r-1} \frac{(-1)^{n-r-1}}{(n-r)2^{2n-2r-1}} + n!\sqrt{x},$$

so that

$$\binom{2n+2}{n+1}(n+1) = \sum_{r=0}^{n-1} \binom{n}{r} \binom{2n-2r-2}{n-r-1} \frac{(-1)^{n-r-1}2^{2r+2}}{(n-r)} + 2^{2n+1}.$$

The binomial theorem can be recovered as follows:

$$\begin{aligned}
(a+b)^n &= \mathrm{e}^{-(a+b)x} \frac{\mathrm{d}^n(\mathrm{e}^{ax}\mathrm{e}^{bx})}{\mathrm{d}x^n} \\
&= \mathrm{e}^{-(a+b)x} \sum_{r=0}^{n} \binom{n}{r} a^r \mathrm{e}^{ax} b^{n-r} \mathrm{e}^{bx} = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}.
\end{aligned}$$

Finally, consider $w = x^n \ln x$. This time Leibniz's theorem sounds like a good idea. Put $y = x^n$ and $z = \ln x$. Then

$$\frac{\mathrm{d}^r y}{\mathrm{d}x^r} = \frac{n!}{(n-r)!} x^{n-r} \quad (\text{for } r \geqslant 0) \quad \text{and} \quad \frac{\mathrm{d}^r z}{\mathrm{d}x^r} = \frac{(-1)^{r-1}(r-1)!}{x^r} \quad (\text{for } r \geqslant 1).$$

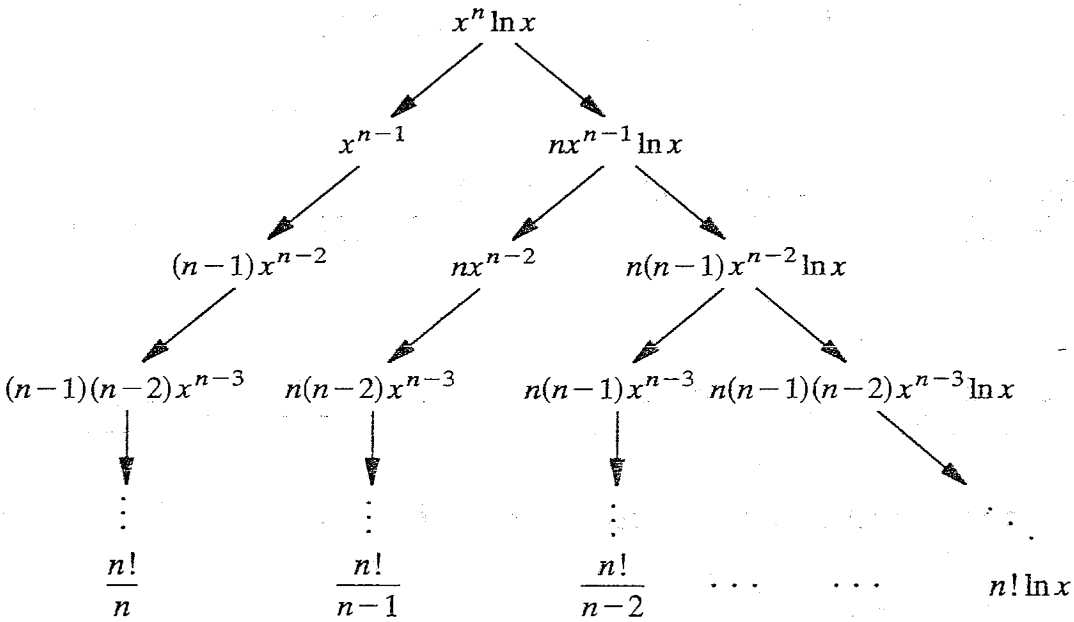
Therefore

$$\begin{aligned}
\frac{\mathrm{d}^n w}{\mathrm{d}x^n} &= \sum_{r=0}^{n-1} \binom{n}{r} \frac{\mathrm{d}^r y}{\mathrm{d}x^r} \frac{\mathrm{d}^{n-r} z}{\mathrm{d}x^{n-r}} + n!\ln x \\
&= \sum_{r=0}^{n-1} \binom{n}{r} \frac{n!\,x^{n-r}}{(n-r)!} \frac{(-1)^{n-r-1}(n-r-1)!}{x^{n-r}} + n!\ln x \\
&= n! \sum_{r=1}^{n} \frac{(-1)^{r-1}}{r} \binom{n}{r} + n!\ln x
\end{aligned}$$

(replacing $n-r$ by r in the summation). However, it is easily proved by induction that

$$\frac{\mathrm{d}^r w}{\mathrm{d}x^r} = \frac{n!\,x^{n-r}}{(n-r)!}\sum_{s=n-r+1}^{n}\frac{1}{s} + \frac{n!}{(n-r)!}x^{n-r}\ln x \quad (\text{for } 1 \leqslant r \leqslant n).$$

Alternatively study the figure, in which a downwards arrow represents differentiation.



Therefore

$$\frac{\mathrm{d}^n w}{\mathrm{d}x^n} = \sum_{r=1}^{n}\frac{n!}{r} + n!\ln x.$$

The result is that

$$\sum_{r=1}^{n}\frac{(-1)^{r-1}}{r}\binom{n}{r} = \sum_{r=1}^{n}\frac{1}{r}.$$

This seems familiar; see Problem 21.11 of *Mathematical Spectrum*.

I can think of one way of generalizing this idea, involving extending Leibniz's theorem to derivatives of products such as $w = f(x)g(x)h(x)$. Readers may be able to see other possibilities.

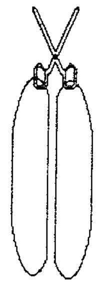
Letters to the Editor

Dear Editor,

Dirac's scissors

In Volume 22, Number 3 of *Mathematical Spectrum*, Martin Gardner's *Riddles of the Sphinx* was reviewed. A highlight of the book was said to be the topological puzzle 'Dirac's Scissors'. However, the solution given by Martin Gardner is incomplete, and does not get one out of the tangle of twisted string as promised!

The puzzle is: Tie a loop of string through each of the handles of some scissors. Stand on the other ends of the two untwisted loops, putting your feet through them. Hold the scissors vertically, pointing them towards the ceiling, then give the scissors two full turns (720 degrees) clockwise (as viewed from the ceiling). Now, return the cords to their untwisted state without rotating the scissors.



Gardner's 'solution' is: 'Hold the scissors in your right hand. With your left hand take the centre of the twisted strands, carry the string up on the far side of the scissors, pass the loop over the scissors and allow it to fall on your right arm. Take the scissors in your left hand. Release the scissors with your right hand, allowing the loop of the string to fall. Raise the scissors. You are back where you started. The tangles have vanished'.

However, the tangles have not vanished! In order to complete the solution, one must repeat the above operation, but change 'right hand' to 'left hand', 'left hand' to 'right hand' and 'far side' to 'near side'. NOW, the tangles have vanished!

(Solution due to Simon Chatterjee, Sixth Form, Queen Elizabeth Grammar School, Blackburn.)

Yours sincerely,
DAVID YATES
(67 Wilbraham Street,
Preston, Lancs PR1 5NN)

Dear Editor,

White to move and mate in two

In his article in Volume 22 Number 3 Keith Austin gives the solution to a two-move chess problem. However, there is another solution!

1. N → N5 mates next move by either 2. Q takes King pawn, or 2. Q → N3, depending on Black's reply.

Yours sincerely,
DAVID YATES

This does not affect the thrust of Keith Austin's article. Editor.

Dear Editor,

Expressing powers of odd integers as sums of consecutive integers

In regard to the request of Mr Dutta on page 78 of Volume 22, Number 3 of *Mathematical Spectrum*, the general rule already appears in my Letter to the Editor entitled 'On sums of consecutive integers' on page 57 of Volume 18, Number 2.

Another representation of *odd* integer powers as sums of consecutive integers is given by

$$x^n = \sum_{i=-\frac{1}{2}(x-1)}^{\frac{1}{2}(x-1)} (x^{n-1}+i) \quad (x \text{ odd}, n \geqslant 2),$$

i.e. the sum of x consecutive integers evenly split across x^{n-1}, and thus equal to $x \times x^{n-1} = x^n$. Of course this cannot be done if x is even. Also, the numbers involved are very large even for reasonable n, so my earlier method is better.

Yours sincerely,
JOSEPH MCLEAN
(9 Larch Road,
Glasgow G41 5DA)

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

Problems

23.4 (Submitted by Jeremy Bygott, Queens' College, Cambridge)
Find all positive integers x such that the decimal representation of $x+2$, when written backwards, equals $\frac{1}{2}x$.

23.5 (Submitted by G. N. Thwaites, Oakham School)
ABC is a triangle and ABC', BCA' and CAB' are equilateral triangles drawn on the sides AB, BC and CA, respectively of ABC, exterior to ABC. Prove that AA', BB' and CC' are concurrent and of equal length.

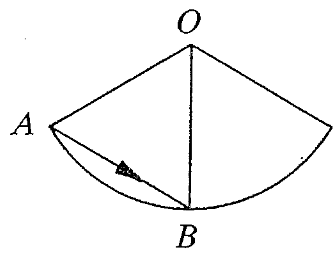
23.6 (Submitted by Krzysztof Hryniewiecki, a student at Grammar School No. 1, Białystok, Poland)
Does there exist a set S of points in three-dimensional space which has a finite non-empty intersection with every plane?

Solutions to Problems in Volume 22 Number 3

22.9. An anthill is in the shape of a circular cone. The diameter of the base of the cone is 20 mm and its side is 30 mm. An ant wishes to go (above ground) from a point A on the base of the hill to the point B diametrically opposite. If the ant can move at the rate of 1 mm/s, how long would the journey take if it took the fastest route?

Solution by Pierino Gattei (Queen Elizabeth's Grammar School, Blackburn)

If we cut the cone along OA, where O is its apex, and flatten it into a plane we obtain a sector of a circle radius OA. The fastest route from A to B is now the straight line AB. The arc AB has length $\frac{1}{2}\pi\times 20 = 10\pi$ mm so that $\angle AOB = 10\pi/30 = \frac{1}{3}\pi$ radians. Hence $\triangle AOB$ is an equilateral triangle, so that $AB = 30$ mm. Hence the journey takes 30 s.



Also solved by Amites Sarkar (Winchester College).

22.10. Prove that

$$\sum_{k=0}^{n}(-1)^{n-k}2^{2k}\binom{n+k}{2k} = 2n+1.$$

Solution by Amites Sarkar

Let

$$A_n = \sum_{k=0}^{n}(-1)^{n-k}2^{2k}\binom{n+k}{2k}, \qquad B_n = \sum_{k=0}^{n-1}(-1)^{n-k+1}2^{2k}\binom{n+k}{2k+1}.$$

Using Pascal's identity for binomial coefficients, we see that

$$A_n - B_n = B_{n+1}, \qquad 4B_{n+1} - A_n = A_{n+1},$$

with $A_1 = 3$ and $B_1 = 1$. Assume inductively that $A_n = 2n+1$ and $B_n = n$. Then

$$B_{n+1} = 2n+1-n = n+1, \qquad A_{n+1} = 4(n+1)-2n-1 = 2(n+1)+1,$$

completing the inductive step, so that $A_n = 2n+1$ for all natural numbers n.

Also solved by Peik Bremer (University of Hanover).

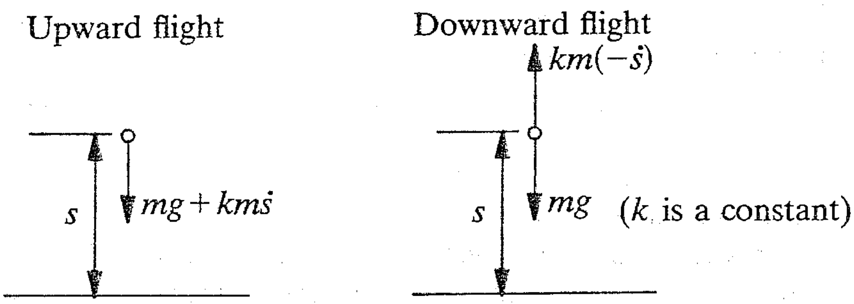
22.11. A ball is thrown vertically upwards in a uniform gravitational field and experiences a force of resistance per unit mass of magnitude proportional to its speed. How does the time taken by the ball to reach the highest point compare with the time taken for it to return from the highest point back to the point of projection?

Solution

In both the upward and downward flight, the equation of motion is

$$mg + km\dot{s} = -m\ddot{s}.$$

Note that $\dot{s} > 0$ on the upward flight, but $\dot{s} < 0$ on the downward flight. We now reverse the upward flight and measure time from the highest point. Although this time reversal changes the direction of the velocity, it does not change the acceleration. Now, at each moment of time, the acceleration is greater on the reversed upward flight than on the downward flight, so that the time taken for the upward flight is less than the time taken for the downward flight.



22.12. Let α be an $(n+1)$-digit number $(n \geqslant 1)$ with digits $x_0, x_1, \ldots, x_n$, so that

$$\alpha = x_n \times 10^n + x_{n-1} \times 10^{n-1} + \ldots + x_1 \times 10 + x_0.$$

Prove that

$$\alpha \geqslant \sum_{i=0}^{n} x_i + \prod_{i=0}^{n} x_i$$

and determine for which numbers there is equality.

Solution by Amites Sarkar

$$\alpha - \sum_{i=0}^{n} x_i \geqslant x_n(10^n - 1) \geqslant 9x_n 10^{n-1} \geqslant 9^n x_n \geqslant \prod_{i=0}^{n} x_i,$$

with equality if and only if $n = 1$ and $x_0 = 9$, so that the required numbers are 19, 29, 39, 49, 59, 69, 79, 89 and 99.

Also solved by Pierino Gattei.

Reviews

Mathematics Focus. Editor DAVID BURGHES. Centre for Innovation in Mathematics Teaching, School of Education, University of Exeter. £20 per year (three issues) starting in June 1989. (ISSN 0955-6257).

This is the first issue of 'a magazine of resources aimed at enhancing mathematics teaching in secondary schools'. The unusual feature of this magazine is that it consists of 20 worksheets. Each worksheet is produced on a separate glossy A4 card and contains an investigation or activity on one side with teacher's notes on the other side. The materials can be freely copied within individual schools or colleges. The stated aim of each issue is to include a cross-section of problems, applications, project ideas, investigations, games, puzzles, algorithms and computer programs.

I was a little disappointed by the content of the pack since only about a quarter of it was new material to me, the rest being a rehash of material already available. I was impressed by the material that was new, however. There was a worksheet on Bode's Law, an unexplained empirical law about the distances of the planets from the Sun, in which students were invited to make conjectures about the asteroids and the origin of Neptune and Pluto. This is an interesting mathematical application suitable as extension material at GCSE. Of the other worksheets there was interesting material dealing with planar circuits, horse racing, the games of Bridge It and Zig Zag (which were new to me), and testing physical fitness using the Gallagher and Braihe Test.

The remainder of the worksheets dealt with better-known topics such as magic squares, dissections, reaction times, nets for polyhedra, a sumsearch, Russian multiplication and a 'Fence it Off' investigation. All of these topics are already available in some other form.

The material overall was very well presented but I felt that there was not a sufficient number of original ideas that I would personally use for GCSE mathematics work. I should need to look at subsequent issues before wholeheartedly recommending this magazine at a price of £20 for three issues.

Portsmouth Sixth Form College L. A. FEARNEHOUGH

Innumeracy. By JOHN ALLEN PAULOS. Viking Penguin, London, 1989. Pp. 135. £12.95. (ISBN 0-670-83008-9).

'Innumeracy, an inability to deal comfortably with the fundamental notions of number and chance, plagues far too many otherwise knowledgeable citizens. The same people who cringe when words such as "imply" and "infer" are confused react without a trace of embarrassment to even the most egregious of numerical solecisms'. Thus John Allen Paulos begins his introduction to *Innumeracy*.

The contents are arranged in five chapters as follows: Introduction; 1: Examples and Principles; 2: Probability and Coincidence; 3: Pseudoscience; 4: Whence Innumeracy? 5: Statistics, Trade-Offs, and Society, Close.

The examples are often funny, and they tease and provoke us into considering serious issues from a more numerate standpoint. Two of the many questions raised as examples are:

1. What are the chances you just inhaled a molecule which Caesar exhaled in his last breath?
2. Judy is thirty-three, unmarried, and quite assertive. She graduated in political science and was deeply involved in campus social affairs, especially anti-discrimination and anti-nuclear matters. Which statement is more probable?
   (a) Judy works as a bank clerk.
   (b) Judy works as a bank clerk and is active in the feminist movement.

John Allen Paulos demonstrates 'the ease with which mathematical certainty can be invoked to bludgeon the innumerate into dumb acquiescence' and he inveighs against the 'pseudosciences', astrology in particular: 'the gravitational pull of the delivering obstetrician far outweighs that of the planet or planets involved ... does this mean that fat obstetricians deliver babies that have one set of personal characteristics, and skinny ones deliver babies that have quite different characteristics?'

The author ranges over various aspects of popular culture, including elections, sports, records, sex discrimination, insurance, lotteries and drug testing. The result is an immensely exciting and a brilliantly entertaining study of a very serious subject. I recommend it unreservedly to sixth-form students, to their teachers and to anyone who would like to develop a more quantitative way of looking at the world.

Medical School, GREGORY D. ECONOMIDES
University of Newcastle upon Tyne

Paradoxes in Probability Theory and Mathematical Statistics. By GÁBOR J. SZÉKELY. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1987. Pp. 250. £41.50 (ISBN 90-277-1899-7).

'This books aims to show how the mathematical methods of random phenomena have developed from paradoxes. It tries to show those exciting moments that preceded or followed the solution of some outstanding paradoxical problems'. Thus Gábor Székely, in his introduction, encapsulates the theme of his book.

The contents are arranged in five chapters as follows: 1, classical paradoxes of probability theory; 2, paradoxes in mathematical statistics; 3, paradoxes of random processes; 4, paradoxes in the foundations of probability theory; 5, paradoxology.

The book contains 42 paradoxes and 47 'quickies', as the author calls them, which do not fit into the main line of the book. The former include: the paradox of dice (which, in spite of its simplicity, several great mathematicians, including Leibniz and d'Alembert, failed to solve); the paradox of horse kickings; the St. Petersburg paradox; Bertrand's paradox and the paradox of zero probability.

Each chapter begins with one or more quotations and finishes with quickies. Each paradox is discussed in five parts: history, formulation, explanation, remarks, and references. The quotations grasp the reader's interest, and the quickies leave him or her delighted and eager for more. The formulation and explanation of the paradoxes, some of which have not been published before, are excellent, and the references (as good as a gold mine) should enable the interested reader to see these tantalising snapshots in context. The history and remarks (for me these are real gems) provide us with a summary of the historical and philosophical backgrounds, and highlight the road ahead for those involved in the study of probability theory and mathematical statistics.

The level of sophistication needed by the reader is that of a good A-level in further mathematics.

This English edition is a revised and updated version of the Hungarian *Paradoxonok a véletlen matematikában*. The translation is well done. The book is on the expensive side, but it will definitely enrich any bookshelf. If only workers in other fields in mathematics would publish gems such as this book.

Sixth Form, Royal Grammar School, GREGORY D. ECONOMIDES
Newcastle Upon Tyne

We have received a further request from a university student in Iran for help in finding a penfriend in Britain. Anyone interested is invited to contact:

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

CONTENTS

ISSN 0025-5653

PRICES *(postage included)*

Prices for Volume 23 (Issues Nos. 1, 2, 3 and 4)

Subscribers in	*Price per volume*	
North, Central and South America	US$13.00 or £8.00	(Note: These overseas prices apply even if the order is placed by an agent in Britain.)
Australia	$A16.00 or £8.00	
Britain, Europe and all other countries	£6.00	

A discount of 10% will be allowed on all orders for five or more copies of Volume 23 sent to the same address.

Details of reduced prices for two- and three-year subscriptions available on request.

Back issues

All back issues except Volume 1 are available; information concerning prices and a list of the articles published may be obtained from the Editor.

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

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

Published by the Applied Probability Trust

Typeset by the University of Nottingham

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