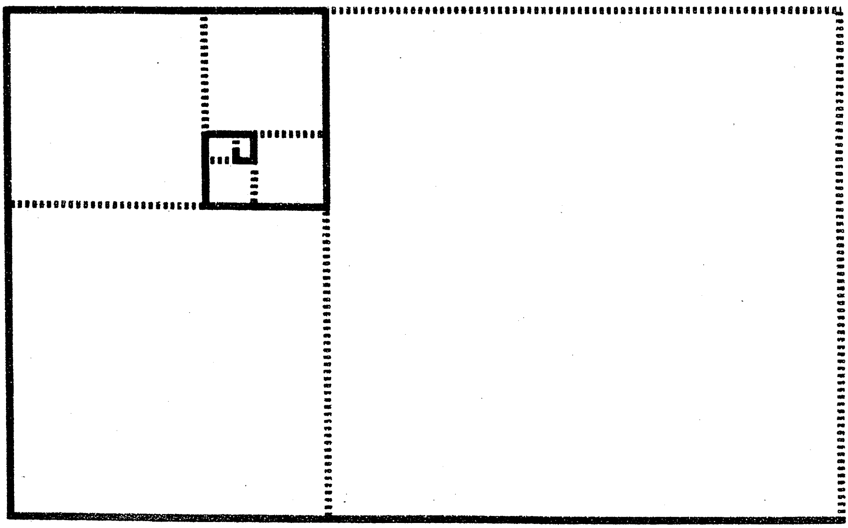
Function

A School Mathematics Journal

Volume 21 Part 1 February 1997



Department of Mathematics and Statistics - Monash University

Reg. by Aust. Post Publ. No. PP338685/0015

Function is a refereed mathematics journal produced by the Department of Mathematics and Statistics at Monash University. The journal was founded in 1977 by Prof G B Preston. *Function* is addressed principally to students in the upper years of secondary schools, and more generally to anyone who is interested in mathematics.

Function deals with mathematics in all its aspects: pure mathematics, statistics, mathematics in computing, applications of mathematics to the natural and social sciences, history of mathematics, mathematical games, careers in mathematics, and mathematics in society. The items that appear in each issue of *Function* include articles on a broad range of mathematical topics, news items on recent mathematical advances, book reviews, problems, letters, anecdotes and cartoons.

* * * * *

Articles, correspondence, problems (with or without solutions) and other material for publication are invited. Address them to:

The Editors, *Function*
Department of Mathematics and Statistics
Monash University
900 Dandenong Rd
Caulfield East VIC 3145, Australia
Fax: +61 (03) 9903 2227
e-mail: function@maths.monash.edu.au

Function is published five times a year, appearing in February, April, June, August, and October. Price for five issues (including postage): \$20.00*; single issues \$5.00. Payments should be sent to: The Business Manager, *Function*, Department of Mathematics and Statistics, Monash University, Clayton VIC 3168, AUSTRALIA; cheques and money orders should be made payable to Monash University.

For more information about *Function* see the journal home page at `http://www.maths.monash.edu.au/~cristina/function.html`.

*\$10 for *bona fide* secondary or tertiary students.

EDITORIAL

We welcome new and old readers alike with our 21st volume of *Function*. We hope you find in it many interesting and enjoyable items.

The number of articles written about the golden ratio seems to be endless, but there is always something interesting to say about it. The front cover of this issue depicts the rectangular spiral obtained with a sequence of inscribed golden rectangles. Bert Bolton calculates the length of the spiral and finds its point of convergence.

We include two feature articles in this issue of *Function*. For those who are likely to forget the combination of a lock, Peter Grossman presents in his article the mathematics which could help you in the task of cracking the combination. The second article, by Michael Deakin and Otto Steinmeyer, takes us to the first century BC to see how mathematics can be used to make sense of some lines in a poem by Vergil.

The *History of Mathematics* column presents an update on an earlier article which gave biographies of saints or candidates for sainthood who also made a significant contribution to mathematics. In the *Computers and Computing* section you will be introduced to the mathematics behind drawing tools used by graphic designers.

We thank the many readers who send solutions to our problems. We publish all those that reach us before the publication date. There are a few new problems in the *Problem Corner* for your entertainment. We also challenge you with the problems set for the participants of the 1996 AMOC Senior Mathematics Contest.

Happy reading!

* * * * *

THE FRONT COVER

The Golden Rectangle and its Rectangular Spiral

Bert Bolton, University of Melbourne

The spiral on the front cover is based on the golden rectangle. The golden rectangle, known to the early Greek mathematicians as the most pleasurable rectangle to look at, is drawn in Figure 1. The shorter sides are taken to have unit length and the longer sides have length r. The rectangle is said to be golden if the smaller rectangle created by completing the square of side unity has the same proportions as the larger rectangle[1]. Equating the ratios of the longer to the shorter sides gives

$$\frac{r}{1} = \frac{1}{r-1}, \tag{1}$$

yielding the quadratic equation

$$r^2 - r - 1 = 0 \tag{2}$$

and the two roots are $r = (1 \pm \sqrt{5})/2$. A measured length of a side is positive and the root with the positive sign gives the physical solution $r = 1.618034\ldots$. The other way that the Greeks looked at the rectangle was that the ratio of the longer to the shorter side is the same as the sum of the two sides to the longer side. Hence

$$\frac{r}{1} = \frac{r+1}{r} \tag{3}$$

which yields the same quadratic equation (2).

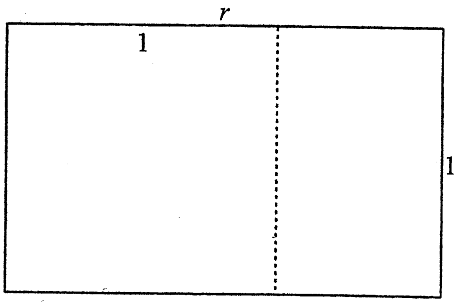


Figure 1

[1]See also *Function, Vol* 16 *Part* 5.

The shapes of paintings are often close to this golden rectangle when hung either horizontally, as in the larger rectangle in Figure 1, or vertically, as in the smaller rectangle.

In Figure 2, x and y axes have been added. Figure 2 has been drawn for the particular case of the golden rectangle, but the argument is expressed for any value of r. The reduction in size from one rectangle to the next is called the scale factor $s = 1/r$.

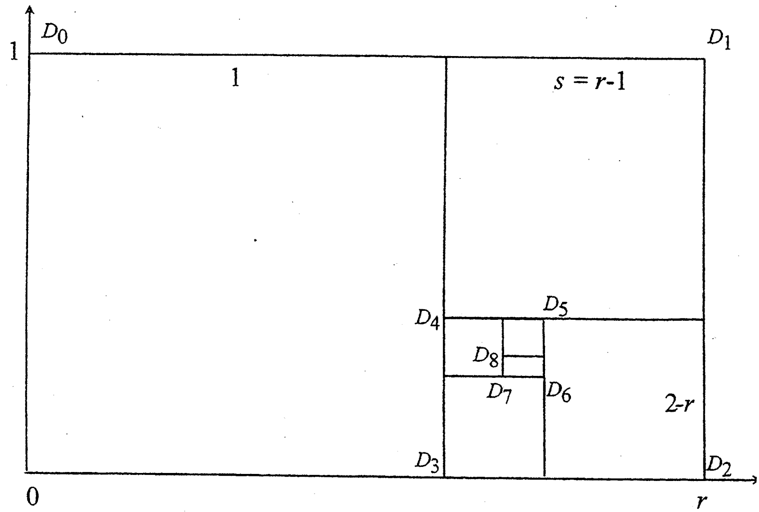


Figure 2

Consider the sequence of points D_0, with co-ordinates $(0, 1)$ at the beginning of the long side of the original rectangle, and the equivalent points D_1, D_2, D_3, etc. shown in Figures 2 and 3. The sides D_0D_1, D_1D_2, D_2D_3 etc. decrease in length by the scale factor s, and form a rectangular spiral, which is shown on the front cover. If the sequence continues to infinity, this rectangular spiral converges to a point within the original rectangle. This point will be called the *asymptotic point* of the rectangle and has co-ordinates (X, Y).

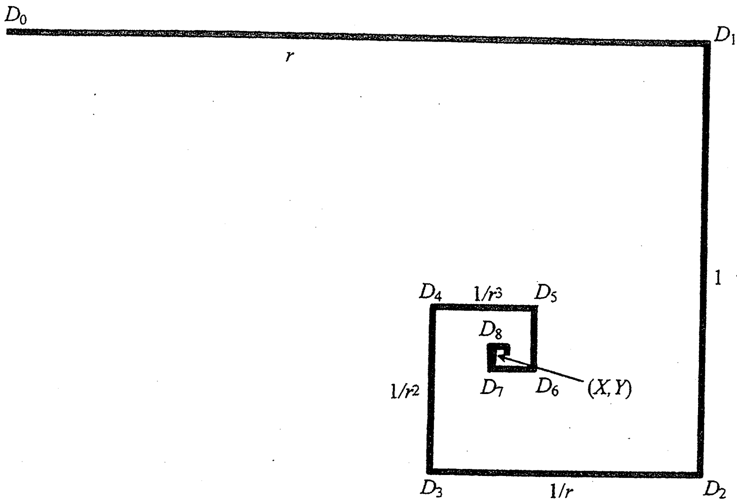


Figure 3

These co-ordinates can be found by starting with the co-ordinates of D_0 which are $(0,1)$, and adding or subtracting the values of lengths of the sides of successive rectangles. From Figure 3 we have

$$X = D_0D_1 - D_2D_3 + D_4D_5 - D_6D_7 + \ldots$$

or

$$X = r - rs^2 + rs^4 - rs^6 + \ldots$$

which is written as $X = rA(s)$ where $A(s)$ is the infinite series

$$A(s) = 1 - s^2 + s^4 - s^6 + s^8 - \ldots$$

The expression on the right hand side is known as a geometric progression or geometric series with a constant ratio $-s^2$ between any two consecutive terms. Such a series can be summed as follows. Multiply $A(s)$ by s^2 to get

$$s^2A(s) = s^2 - s^4 + s^6 - s^8 + \ldots$$

Adding this to $A(s)$ gives

$$(1+s^2)A(s) = 1,$$

because all the terms cancel except the first. Thus $A(s)$ reduces to

$$A(s) = \frac{1}{1+s^2}.$$

The value of X for the asymptotic point is

$$X = \frac{r}{1+s^2}$$

and using $s = 1/r$, $X = r^3/(r^2+1)$. So for the golden rectangle

$$X = \frac{(1+\sqrt{5})^3}{4(5+\sqrt{5})} = 1.17082\ldots .$$

The value of the second component of the asymptotic point, Y, follows in a similar manner. Referring to Figure 3,

$$Y = D_3D_4 - D_5D_6 + D_7D_8 - \ldots$$

or

$$Y = rs^3 - rs^5 + rs^7 - \ldots$$

which is $Y = rs^3A(s)$, where $A(s) = 1/(1+s^2)$, already calculated.

Putting $s = 1/r$ as before, gives

$$Y = \frac{1}{r^2+1} = \frac{2}{5+\sqrt{5}} = 0.276393\ldots .$$

The pattern of each rectangle is at right angles to its predecessor. On Figure 2 this means that the diagonals D_0D_2 and D_1D_3 are at right angles and all other diagonals lie on one or other of these lines. Their intersection is the point (X, Y).

The spiral has an infinite number of sides but a finite length. The length is

$$D_0D_1 + D_1D_2 + D_2D_3 + \ldots = r + rs + rs^2 + rs^3 + \ldots$$

which is $r(1+s+s^2+s^3+\ldots)$, revealing another geometric progression with a ratio s. The length of the spiral can thus be written as

$$\frac{r}{1-s} = \frac{r^2}{r-1} = \frac{(1+\sqrt{5})^2}{2(-1+\sqrt{5})} = 4.23607\ldots .$$

A spiral can also be made by going from any point in the first rectangle to the equivalent point in the second rectangle and so on. For example, start from the centre of the first square in Figure 2, go to the centre of the second square, and so on. The sides of the spiral are not parallel to the x and y axes and the algebra is more complicated than that given above, but its asymptotic point is the same as (X, Y) calculated above. Sketches of other spirals suggest that the asymptotic point is the same for all and that it is a special point for the golden rectangle. Did the Greeks know of this asymptotic point?

* * * * *

The traditional mathematics professor of the popular legend is absent-minded. He usually appears in public with a lost umbrella in each hand. He prefers to face the blackboard and to turn his back to the class. He writes a, he says b, he means c; but it should be d. Some of his sayings are handed down from generation to generation.

"In order to solve this differential equation you look at it till a solution occurs to you."

"This principle is so perfectly general that no particular application of it is possible."

"Geometry is the science of correct reasoning on incorrect figures."

"My method to overcome a difficulty is to go round it."

"What is the difference between method and device? A method is a device which you used twice."

– George Polya in *How to Solve it*
Princenton University Press, 1945

* * * * *

If we possessed a thorough knowledge of all the parts of the seed of any animal (e.g. man), we could from that alone, by reasons entirely mathematical and certain, deduce the whole conformation and figure of each of its members, and, conversely if we knew several peculiarities of this conformation, we would from those deduce the nature of its seed.

– René Descartes in *Discours de la Méthode*

* * * * *

CRACKING THE COMBINATION THE GOOD WAY!

Peter Grossman

Imagine you want to enter a room with a combination lock on the door, but you have forgotten the combination. The lock has 10 buttons, labelled with the digits 0 to 9, and to open the door you need to press the 3-digit combination in the right order. The correct sequence always opens the door immediately, irrespective of which buttons were pressed previously. You decide to try every combination until you find the one that works, but of course you want to minimise the time and effort required. You can do this by allowing the 3-digit sequences to overlap; for example, if you enter the sequence 1112 the door will open if either 111 or 112 is the right combination. What strategy should you use?

Let's begin by looking at a simpler version of the problem. Suppose that instead of the ten decimal digits, the lock uses only the two binary digits 0 and 1. The problem now is to find the shortest sequence of binary digits (or "bits") containing every 3-bit string. If possible, we would like the sequence to contain every such string exactly once.

After some trial and error, we might devise the following sequence:

$$0001011100 \tag{1}$$

The eight different 3-bit strings are 000, 001, 010, 011, 100, 101, 110 and 111, and you can easily check that each of them occurs exactly once in the sequence (1).

Notice also that the last two bits of the sequence (1) are the same as the first two. Hence we can remove the last two bits, provided we agree that the resulting sequence, 00010111, is to be treated as a cycle, i.e. the 3-bit strings 110 and 100 are obtained respectively from the 7th, 8th and 1st bits, and the 8th, 1st and 2nd bits.

The sequence 00010111 is an example of a (binary) *de Bruijn sequence* of order 3. The general definition is as follows.

Definition. A *de Bruijn*[1] *sequence* of order n is a sequence of bits within which every possible n-bit string occurs exactly once, where the bits in the sequence are taken in cyclic order.

A little thought reveals that a de Bruijn sequence of order n must contain 2^n bits. As a simple first exercise, you could write down a de Bruijn sequence of order 2.

De Bruijn sequences arise in a number of areas of applied mathematics, including communication theory and coding. As an example to illustrate how these sequences might be used in a practical setting, imagine a horizontally mounted disk which can be rotated into any one of eight positions spaced at equal angles. The disk is divided into eight sectors, each with one bit (0 or 1) marked on it. A read-head fixed in position over the disk can read the bits marked on the three consecutive sectors that fall directly under it. If the eight bits on the disk follow the de Bruijn sequence of order 3 given above, then the information read by the read-head is sufficient to allow the position of the disk to be uniquely determined.

There are many questions we could ask about de Bruijn sequences. In this article, we will confine our attention to just one: Can a de Bruijn sequence of order n be found for each natural number value of n?

In order to answer this question, it is helpful to depict the problem graphically. This can be done using a *Good*[2] *diagram*. The Good diagram for de Bruijn sequences of order 3 is shown in Figure 1. It takes the form of a *directed graph*, or *digraph*: a diagram in which points, called *vertices*, are connected by *directed edges*, depicted as curves with arrows. In a Good diagram of order n, each of the 2^{n-1} vertices is labelled with a different string of $n-1$ bits. If two vertices, u and v, are labelled with the bit strings $b_1b_2\ldots b_{n-1}$ and $b_2b_3\ldots b_{n-1}b_n$ respectively, so that the last $n-2$ bits of the label for u are the same as the first $n-2$ bits of the label for v, then there is a directed edge from u to v labelled $b_1b_2\ldots b_{n-1}b_n$. (Here, each b_i is either 0 or 1.)

[1] N G de Bruijn was a Dutch researcher who investigated these sequences in 1946 in the context of a problem in telephone engineering. He did not invent the concept; as early as 1882, the French telegraph engineer Émile Baudot had used a de Bruijn sequence of order 5 to solve a problem arising in connection with transmitting the 5-bit codes for the 32 characters used in telegraphy.

[2] I J Good was a British mathematician who in 1946 developed the method described in this article in order to solve a problem involving recurring decimals.

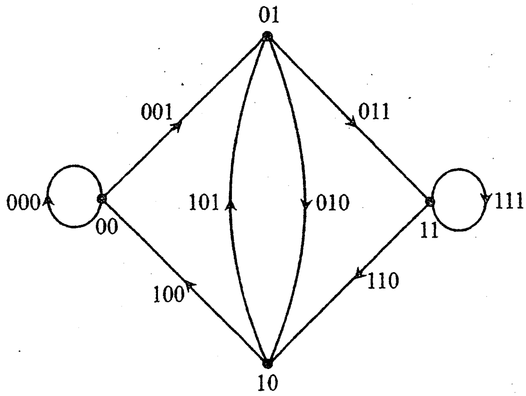


Figure 1

If you examine Figure 1 carefully, you will see that it satisfies the definition of a Good diagram with $n = 3$. For example, since the last bit of 00 is the same as the first bit of 01, there is a directed edge labelled 001 from the vertex 00 to the vertex 01. In particular, note that there are directed loops attached to the vertices labelled 00 and 11 (but not to the other two vertices).

We will now make the connection between de Bruijn sequences and Good diagrams. Starting at the vertex labelled 00 in Figure 1, follow the directed edge 000 which returns to the vertex 00. Now follow the directed edge 001 to the vertex 01. Note that 000 and 001 are the first two 3-bit strings in the de Bruijn sequence we found earlier. They are guaranteed to overlap correctly, the last two bits of 000 coinciding with the first two bits of 001, because the common bits 00 form the label of the vertex we pass through when we follow the directed edge 000 with the directed edge 001.

If we continue in this way, following the directed edges in the Good diagram in the order in which the 3-bit strings appear in the de Bruijn sequence, we find ourselves tracing a path that uses each directed edge in the diagram and returns to the starting vertex. In fact, since each 3-bit string occurs once in the de Bruijn sequence, the path in the Good diagram must traverse each directed edge *exactly once*.

You may have come across a problem like this before. An old and very well-known puzzle, the *Königsberg bridge problem*, asked whether it was possible for a person to walk around the 18th-century town of Königsberg along a route that crossed each of the seven bridges in the town exactly once. We won't discuss the problem here, but you can find it in any introductory textbook on graph theory[3]. Our present problem is very similar, except that in our case the edges, which correspond to the bridges in the Königsberg bridge problem, can be traversed only in the direction indicated by the arrows.

As a second exercise, see if you can draw the Good diagram of order 4. You should discover that it can be drawn in a way that has a neat symmetry with no edges crossing. Then use the diagram to find a de Bruijn sequence of order 4.

The question of whether there exist de Bruijn sequences of all orders now becomes: Is it always possible to trace a path in a Good diagram, returning to the starting vertex and using every edge exactly once? We will show that the answer is yes, omitting some rather technical details at one point.

We start by noting that every vertex in a Good diagram must have two edges entering it and two edges leaving it. (Can you see why?) So if we trace a path without repeating edges, we can enter each vertex at most twice. Furthermore, each time we enter a vertex we can also leave it, unless it is the vertex we started from. Consequently, if we trace a path through the diagram without repeating edges, continuing until we can't go any further, we must end where we started, although we might not have used every edge in the diagram.

We will now show that if the path doesn't use every edge, it can be extended. For example, suppose in Figure 1 we had obtained a path by starting at 00 and following the edges 000, 001, 010, 100. We can insert some of the unused edges into the path using the following technique. We trace the path we have found until we reach a vertex with an unused edge leaving it. In this example, 01 is the first such vertex. Starting at 01, we follow unused edges until we can't go any further. When this occurs, we must be back at 01. (The reason is much the same as before, when we explained why the original path had to finish where it began.) For example, from 01 we can follow the edges 011, 110, 101. We now trace the original path to its end at 00. By inserting the new sequence of edges into the original path in this way, we produce a new longer path.

[3]See also M Deakin's article "A Walk across the Bridges" in *Function, Vol* 13 *Part* 1, pp. 20-27.

If there are still unused edges remaining, we can repeat this insertion process as often as necessary. In the present example, the edge 111 has not been used. We trace the existing path to 11, then a new path along the unused edge 111, and finally the existing path to its end. The result is a path that uses all of the edges, as required.

There is one small matter that needs to be resolved before we can claim that this method works for any Good diagram. If there are unused edges remaining at any stage of the process, can we be sure that at least one of them will leave from a vertex in the existing path at that stage? The answer is yes; we will omit the proof, but you may wish to explore the matter yourself.

We can sum up our findings as follows:

In any Good diagram, it is possible to find a path that uses every directed edge exactly once and returns to the starting vertex.

The answer to our original question now follows immediately:

For every natural number n there is a de Bruijn sequence of order n.

Now that we have answered our main question, the way is left open for you to explore further. In particular, let's return to the combination lock problem with which we began this article. In order to tackle this and similar problems, we will need to allow the characters in the sequences to come from an arbitrary finite set rather than just $\{0,1\}$. How would our results need to be modified? Is it still possible to find a de Bruijn sequence of any order in this more general setting? A good starting point for your investigations would be "ternary" de Bruijn sequences composed of the digits 0, 1 and 2. Try drawing some Good diagrams and see what happens.

You could also investigate the more difficult problem of determining how many different de Bruijn sequences there are of a given order. What does it mean for two de Bruijn sequences to be different? Do different paths through a Good diagram always correspond to different de Bruijn sequences? These are just some of the problems you could explore.

* * * * *

Imagination is more important than knowledge.

– Albert Einstein

* * * * *

BOW AND ARROW IN THE GROVES OF IND

Michael A B Deakin, Monash University

and

Otto Steinmayer, Universiti Malaya[1]

The Roman poet Vergil[2] (70 BC – 19 BC) has some lines in his *Georgics* which run, in translation, as follows:

> "Or why should I mention the groves that India close to Ocean bears, a recess of the farthest circle of the world, where not a single arrow in its flight has been able to conquer the loft beyond the summit of a tree?"

This is intriguing as a way to measure the height of a tree, and it leads us to ask how high these trees were, and what kind of bow shot these arrows.

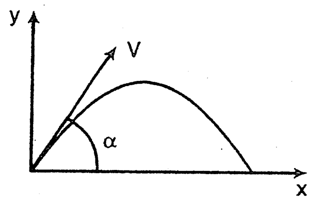


Figure 1. The standard parabolic trajectory; the path followed by an arrow in a vacuum. The case drawn is for $\alpha = 60°$.

Let us begin with some basic considerations. If we neglect air resistance, we may determine the trajectory of an arrow from a knowledge of two parameters. These are V, the velocity with which it is released, and α, the angle of discharge.

Both are illustrated in Figure 1. Use a set of co-ordinates as also illustrated in Figure 1, where the point of discharge is taken as the origin, the x-axis is horizontal in the direction of flight and the y-axis is vertical.

[1]Current address: PO Box 13, 94500 Lundu, Sarawak, Malaysia.

[2]Or Virgil, as it is sometimes spelt.

We now have two equations of motion, derived from Newton's Laws. The equations are:

$$\frac{d^2x}{dt^2} = 0 \quad \text{and} \quad \frac{d^2y}{dt^2} = -g, \tag{1}$$

where g is a constant known as the acceleration due to gravity, and the left-hand sides are the second derivatives with respect to t, the time elapsed since the release of the arrow.

In SI units, g has the value 9.81, in other words about 10.

The solutions of equations (1) are

$$x = (V\cos\alpha)t \quad \text{and} \quad y = (V\sin\alpha)t - \frac{1}{2}gt^2. \tag{2}$$

We also have (by differentiating these)

$$\frac{dx}{dt} = V\cos\alpha \quad \text{and} \quad \frac{dy}{dt} = V\sin\alpha - gt. \tag{3}$$

If we now combine the two equations (2), eliminating t, we reach

$$y = x\tan\alpha - \left(\frac{g}{2V^2}\sec^2\alpha\right)x^2, \quad \alpha < 90°. \tag{4}$$

This is the equation of the arrow's trajectory – a parabola.[3]

Now let us examine some consequences of these equations. First up we want to know how high an arrow can be fired. It should be quite obvious (and moreover can be proved mathematically) that the maximum height will be achieved when the arrow is fired vertically; i.e., when $\alpha = 90°$. It may then be shown (work this out as an exercise using equations (2) and (3)) that the height reached under these circumstances, H let us call it, is given by

$$H = \frac{V^2}{2g}. \tag{5}$$

The other quantity we shall need is the greatest horizontal range that can be covered. This is achieved when $\alpha = 45°$. Call the distance R. Again we leave the proof as an exercise, but quote the result, which is

$$R = \frac{V^2}{g}. \tag{6}$$

---

[3]For more on parabolic trajectories, see *Function, Vol* 16 *Part* 4, p. 100.

Comparing equations (5) and (6), we find the simple result:

$$H = \frac{1}{2}R. \tag{7}$$

Now, while H is a difficult thing to measure, R is not. So here (still as long as air resistance is being neglected), we have the key to the question we are asking. There is quite an amount of data on how *far* a longbow can shoot. The English longbow that was so deadly in medieval warfare was probably as efficient as a wooden "self-bow" can be.[4] These had ranges in the region 150-200m and, for ease of subsequent calculation, we will take 160m as a standard range.

Taken at face value, this would seem to give a figure for H as 80m, and so Vergil seems to be saying that the trees in the groves of India were over 80m high. This seems wrong. Although the giant sequoias of North America can attain such heights, mature rainforest trees typically stand only some 40-45m tall. Was Vergil indulging in poetic licence? Or were we wrong to neglect air resistance?

Well, we'll come back later to the question of air resistance, but the main answer we would suggest to the question posed by the disparity is that Vergil was not referring to the *military* bow (which the Romans did not use), but to some other sort of bow. And if we posit this, we must ask what other kind of bow there was. Now there *were* military bows in use in Indian antiquity; indeed, a passage in the *Mahâbhârata* (a classic Indian epic) has warriors shooting arrows high into the sky. But we think it was not these military bows that Vergil had in mind for otherwise he could hardly have written as he did.

There is indeed another possibility, and once it is advanced it seems obvious, and furthermore, we can advance other evidence in its favour. Throughout India and Southeast Asia, one can still find descendants of the original tribes that lived there long before the invasion of the now-dominant cultural groups. Many of these used the bow and arrow for purposes of hunting (rather than of warfare).

Hunting bows were of a much lighter construction than military bows. They were not intended to be fired over long ranges, nor did they rely on the force of impact to achieve their intended effect. Rather, the arrows were

[4] A "self-bow" is a bow whose body is shaped from a single homogeneous piece of material and is not laminated or reinforced.

tipped with poison, and this (fired from quite close range) was what made them lethal. The ancient Greeks and Romans had *this* tradition (rather than the military one) as their source of knowledge about the bow. This is attested even by the words they used in talking about bows and arrows, and has its echoes still in the English of today.

Strange as it may seem, our word *toxic*, meaning "poisonous", derives from an earlier word for "bow". It reached English via the Latin *toxicum*, meaning "a poison", but the Latin is in its turn derived from the Greek *toxicon pharmakon*, which meant "arrow poison". Between the Greek and the Latin, the word *pharmakon* was dropped[5] and only the single word *toxicon* retained. This word, in *its* turn, derived from an earlier Greek word *toxon*, meaning a bow, and itself came from an Asiatic source, the Scythian **taksha-*, also meaning "bow".[6]

This is in fact only one element in the evidence that the Greeks (and following them the Romans) saw the bow as being of Asian origin. It also strongly suggests that they thought of the hunting weapon, the one with the poison-tipped arrows, when they referred to the bow. So when Vergil spoke of the bow in India, he was referring to the indigenous bow of that land, and his lines now make quite literal sense. Stories of India would have reached him along with the spices that travelled the same route (and sold in Roman markets cheaply, and thus in quantity, as Vergil's contemporary Horace tells us).[7] So this is how Vergil would have learned of the size of the trees in India, and the measurement would have been in terms of the bowshot of the indigenous people of that country.

However, just to be sure of this conclusion, we have to backtrack and consider the question of air resistance, which we postponed. Perhaps surprisingly, it turns out to be quite important. Equations (1-7) refer to an arrow fired in a vacuum, but when air is involved, this exerts forces on the arrow and these may need to be taken into account. After all, we would expect the path of an arrow to be affected by the breeze, and even where (as here) we assume no breeze as such, we may reflect that, as it travels, an arrow creates its own "breeze" because it is in motion relative to the surrounding air.

---

[5]Its modern-day derivatives mean "pertaining to drugs", rather than specifically "poison".

[6]The Scythians were Asian nomads who, between the 7th and the 1st centuries BC, settled in what is now the Ukraine on the northern shores of the Black Sea. The asterisk indicates an inferred word (rather than one still preserved), for Scythian was never a written language.

[7]In his *Epodes*.

The interaction between the moving arrow and the air is in fact very complex, but the best and also the most widely used approximation supposes that the air provides a resistive force directed exactly against the direction of travel and with a magnitude proportional to the square of the speed. This leads to two equations, which replace the simpler equations (1):

$$\frac{d^2x}{dt^2} = -K\frac{dx}{dt}\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$$

and

$$\frac{d^2y}{dt^2} = -g - K\frac{dy}{dt}\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \tag{8}$$

where K is a constant still to be determined.[8]

The value of K depends on many factors[9]; here we shall note that we have already chosen $g \approx 10$ and $R \approx 160$ in SI units. This gives, from equation (6), $V \approx 40$, again in SI units. K depends on V and also on a number of other parameters: the density of air, the density of the wood from which the arrow is made, the length of the arrow, the so-called "kinematic viscosity" of air, and the approximate width of the arrow. We may estimate or look these up to find values in SI units of, respectively, $1.225 \times 10^{-3}, 0.5, 1, 1.5 \times 10^{-5}, 0.01$. The second, third and last of these figures are rough approximations; we are assuming an arrow made of light wood (about the middle of a large range) about 1m long and 1cm across. Fortunately these figures do not affect the calculation very much; the most critical is the density and even this does not matter to any great extent.

An arrow is designed to be a slender, streamlined body. It comprises a long thin shaft, a head and (at the tail end) a set of *fletchings* (often feathers or such) that act to keep the head pointing in the direction of travel. These last localise the drag of the air, so that while the arrow is correctly aligned, air resistance is minimised. There has been much work done on the motion through the air of slender, streamlined bodies, and if we may apply this theory to the arrow, we may make use of the results to find a value for K of 3×10^{-4} in SI units.

This seems quite a small number, but in such cases, we always need to ask "small in relation to what?", and so some further analysis is called for. The

[8]These same equations were discussed (in connection with the sport of long-jumping) by M N Brearley in *Function, Vol* 3 *Part* 3.

[9]For an account of the details and for numerical data, see *Basic Mechanics of Fluids* by H Rouse and J W Howe (New York: Wiley, 1953), p.181.

way to do this is to go over from SI units to a set of other units especially adapted to the problem in hand.[10] These we so choose to ensure that $g = 1$ and $V = 1$. The effect on equations (8) is to simplify them to:

$$\frac{d^2x}{dt^2} = -\varepsilon \frac{dx}{dt}\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$$

and

$$\frac{d^2y}{dt^2} = -1 - \varepsilon \frac{dy}{dt}\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \tag{9}$$

where we have written ε in place of K, in order to emphasise its new significance. In these new units it is the ratio of the force of the air resistance to that of gravity, and its value is about 0.05.

This is a pure number, and it is quite small. If it were to be completely neglected, we would have, from equations (5) and (6), $R = 1$ and $H = 0.5$. We now ask how the non-zero value of ε affects these figures.

Equations (9), although considerably simpler than equations (8), still present great difficulties. Indeed, they cannot be solved exactly; only certain special cases may be fully resolved. One of these is the case $\alpha = 90°$, that of the arrow being fired vertically. When this applies, we have only one equation[11]

$$\frac{d^2y}{dt^2} = -1 - \varepsilon\left(\frac{dy}{dt}\right)^2 \tag{10}$$

and from this it may be deduced that

$$\left(\frac{dy}{dt}\right)^2 = \left(1 + \frac{1}{\varepsilon}\right)e^{-2\varepsilon y} - \frac{1}{\varepsilon}. \tag{11}$$

Now when $y = H$, $\frac{dy}{dt} = 0$. This tells us that

$$H = \frac{\ln(1+\varepsilon)}{2\varepsilon}. \tag{12}$$

---

[10]For a fuller account of the very powerful techniques being employed here, see *Function, Vol* 10 *Part* 1.

[11]This equation applies on the way *up*. On the way *down* it must be modified to read $\frac{d^2y}{dt^2} = -1 + \varepsilon\left(\frac{dy}{dt}\right)^2$ – can you see why?

These formulae will be useful later on.

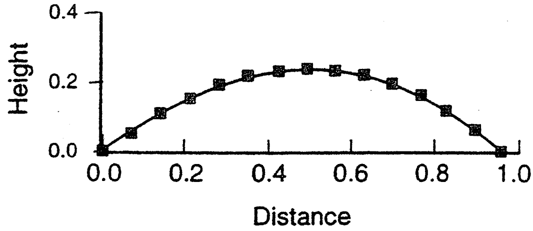


Figure 2. The path of an arrow in an idealised case but with air resistance. The parameter values are $\varepsilon = 0.05$ and $\alpha = 45°$. The points shown are those calculated by the numerical program solving the equations; the curve is interpolated between these.

The other thing we can show about equations (9) is that if ε is small, then its effect on the solutions is small. This is illustrated in Figure 2. The curve shown was calculated for the case $\varepsilon = 0.05$ using numerical techniques on a computer, but it lies very close to an exact parabola and the range is reduced from the value 1 to about 0.96.

When ε is not small, then the only approach is to use a computer and numerical approximation. This is illustrated in Figure 3. In this case we took $\varepsilon = 0.6$ to find the range reduced to about 0.7. This was deliberately chosen because it is the value that has been found in careful experiments.[12]

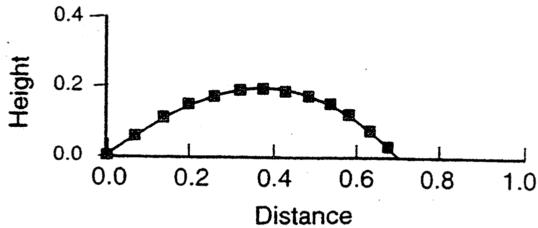


Figure 3. A more realistic path. The parameter values for this case are $\varepsilon = 0.6$ and $\alpha = 46°$. Again, the points are calculated first and then the curve interpolated.

[12]See *New Scientist*, 5 June 1993, pp. 24-25.

But this implies that the calculations that led us to the value of 0.05 for ε were incorrect. What went wrong?

To answer this, think back to that special case in which the arrow was fired vertically. As it goes *up*, its three components (head, shaft and fletchings) all travel along the same path and so it is well and truly streamlined. This gives us our value of H (which works out to be 0.49 in the case $\varepsilon = 0.05$), and we need not for this purpose consider the downward journey. However, it is instructive to consider what happens next. The arrow must now come back to earth and, given a reasonable distance of travel, it will eventually return head downwards. In other words, there will be a period of tumbling motion. While this is going on, the arrow will be anything but streamlined! Thus the value of ε will be greatly increased.

These considerations apply to the general case as well. The direction of travel is changing all along the trajectory and the "tumbling", although less pronounced, is distributed more evenly along the entire flight-path. The *effective* value of ε thus greatly exceeds the theoretical; indeed, its value will vary as the motion proceeds. We found that if we take an average of about 0.6, this produces the observed reduction in the range. However, for the purpose of determining H, the theory is very good and we use $\varepsilon = 0.05$.

So now let us see what this means for our discussion of Vergil. In a vacuum we have, from equations (6) and (7), $\frac{H}{R} = \frac{1}{2}$. When air resistance is taken into account, we have (in the special units) $H = 0.49$ and $R \approx 0.7$. The value of the ratio is thus increased to about 0.7. So now if we go back to SI units, for which $R = 160$, we find that $H \approx 160 \times 0.7 = 112$. Thus if we had had the *military* bow in mind, Vergil would have been saying that the trees were well over 100m high. This is an even greater unlikelihood than the one we dismissed earlier and so taking air resistance into account considerably *strengthens* our case that he spoke of the indigenous hunting bow.

Further Reading

Equations (9) have been widely studied, but the results are not always accessible, nor in convenient form. During World War I, the mathematician J E Littlewood[13] was assigned by the British military to study generalised versions of these equations, which are important in gunnery. A few of his results appeared in his book *A Mathematician's Miscellany* and other bits

[13] Who appeared briefly in *Function, Vol* 19 *Part* 3, pp. 83-85.

and pieces were published elsewhere. In 1971, *Function*'s British counterpart, *Mathematical Spectrum*, printed a two-part account of his work.[14] A more accessible popular article is the summary "Ballistics and Projectiles" in *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences* (see *Vol* 2, p. 1069). This gives a lot of the history and many references, although it is rather terse when it comes to technical detail.[15]

* * * * *

> I thought the following four [rules] would be enough, provided that I made a firm and constant resolution not to fail even once in the observance of them. The first was never to accept anything as true if I had not evident knowledge of its being so; that is, carefully to avoid precipitancy and prejudice, and to embrace in my judgment only what presented itself to my mind so clearly and distinctly that I had no occasion to doubt it. The second, to divide each problem I examined into as many parts as was feasible, and as was requisite for its better solution. The third, to direct my thoughts in an orderly way; beginning with the simplest objects, those most apt to be known, and ascending little by little, in steps as it were, to the knowledge of the most complex; and establishing an order in thought even when the objects had no natural priority one to another. And the last, to make throughout such complete enumerations and such general surveys that I might be sure of leaving nothing out. These long chains of perfectly simple and easy reasonings by means of which geometers are accustomed to carry out their most difficult demonstrations had led me to fancy that everything that can fall under human knowledge forms a similar sequence; and that so long as we avoid accepting as true what is not so, and always preserve the right order of deduction of one thing from another, there can be nothing too remote to be reached in the end, or too well hidden to be discovered.

– René Descartes, *Discours de la Méthode*

* * * * *

[14]It is rather unlikely that very many of *Spectrum*'s target audience actually read or followed Littlewood's account, which is very heavy going. However, the editors must have published it for its historical importance.

[15]There is still much to learn. One problem that seems to be unsolved for the present: the air-resisted trajectories of Figures 2 and 3 (and many other cases besides) can be approximated to wonderfully high orders of accuracy by cubic curves. We are unaware of any theoretical reason why this should be so.

HISTORY OF MATHEMATICS

Michael A B Deakin

St Pavel Florensky and Others

In *Function*, *Vol* 10 *Part* 5, pp. 14-16, I gave brief biographies of a number of mathematicians who had advanced, or might perhaps one day advance, along the path to sainthood in the Roman Catholic church. Later this was updated (see *Vol* 13 *Part* 1, p. 30 and *Vol* 14 *Part* 1, p. 24). The immediate spur to my writing of these articles was the (belated) news that a mathematician, Francesco Faà di Bruno, was officially designated as the Venerable Francesco Faà di Bruno in the minutes of the Vatican in 1971. (This is the first main step towards sainthood; the next is "beatification" – the Blessed ... , and after that comes "canonisation" – St ... .)

In view of the absolute flurry of new saints being created by the present pope, it occurred to me to wonder if Faà di Bruno's status might not by now have been upgraded. I put out an enquiry on the Internet but no-one was able to enlighten me on this point. However, I did learn a lot of other things and these are what I want to share with you in this column. They concern the situation in other branches of the Christian church.

The Anglican church does not have the formal canonisation procedures of its Roman counterpart, so the question, in a formal sense, does not arise in their case. The same may be said for the Scottish and other Protestant traditions. However, there is a long history of British mathematicians taking holy orders.

One of the best-known of these was Charles Dodgson (Lewis Carroll)[1]. Dodgson was a minor mathematician, but is better known as a children's author. However, he never practised as a cleric, nor are there any records of outstanding piety on his part (quite the contrary, in fact – his friendships with young children led to some unsavoury, if quite untrue, gossip).

Another such figure was the Reverend Thomas Kirkman (1806-1895). Kirkman was English, but (probably for financial reasons) studied in Dublin. It seems that he studied divinity rather than mathematics, but he was interested very much in the latter. He is today best remembered for "The Schoolgirls Problem":

[1]See *Function*, *Vol* 7 *Part* 3 and *Vol* 18 *Part* 1.

> Fifteen young ladies in a school walk out three abreast for seven days in succession: it is required to arrange them daily so that no two shall walk twice abreast.

Problems of this type are now seen as very important in the design of statistical trials, and are now studied under the heading "experimental design".[2]

However, his mathematical interests extended beyond this and (in large measure as a result of work on quaternions[3]) he was elected a Fellow of the Royal Society. He also worked in geometry and published a number of works of philosophy and theology. He was active in church affairs and his religious convictions were strong.

But perhaps the name that comes to mind most in such a context as this is that of George Salmon (1819-1904). Certainly this was the one that was passed on to me following my query on the Net. Salmon was an Irishman, and a member of the Church of Ireland, but this actually means (apart from technicalities) a member of the Church of England.[4] His principal association was with Trinity College, Dublin, where he held a number of posts. He was active both in mathematics and also in theology, publishing four well-received technical books in each of these fields. In mathematics, he is perhaps best remembered for his geometrical treatises.

Certainly his best research papers lay in this field. In all, there were 41 of these as well as the four extended works. At least two of his geometric treatises can still be described as classic works in their field; they may still be consulted with profit. He would probably have stayed in mathematics, had his work received the recognition it deserved; however, he failed to gain promotion and found himself not only overworked but also compelled to teach at a level that was well beneath his capabilities. In consequence he changed his career path and in due course achieved his desired promotion, but in the field of divinity rather than of mathematics. This later led him to high administrative office.

However, the main surprise for me in response to my question on the Internet was to hear of a full saint in the Russian Orthodox Church. This is St Pavel Florensky. His name is not widely known in the West, but he

[2]For related problems, see *Function*, *Vol* 4 *Part* 4, pp. 16-23 and *Vol* 8 *Part* 2, pp. 8-9.

[3]See *Function*, *Vol* 5 *Part* 3, pp. 22-25, *Vol* 18 *Part* 3, pp. 81-86 and *Vol* 19 *Part* 5, pp. 152-155.

[4]In other words, he was not a Roman Catholic. It should also be noted that this was before Irish independence from Britain.

is an interesting figure, and a much more modern one than those discussed above. He was most certainly a mathematician, although he did not (as far as I am able to judge) make as significant a contribution to the field as either Kirkman or Salmon.

I base the rest of this article almost entirely on a single source, the only one that I have been able to get my hands on. Florensky is not an easy figure to study. There is a lot of published material, but most of it is in Russian, and even that from very obscure publishing houses. The one article I have managed to obtain is by Graham Flegg (who is, like his subject, both a mathematician and an ordained orthodox priest) and it appeared in the journal *Modern Logic*.[5] Even so, this journal is taken by no Australian library and I had to send to the UK to find a copy.[6]

However, it gives the information that Florensky was born in 1882. Although the town of his birth lies in what is now Azerbaijan and his mother was ethnically Armenian, his father and his cultural heritage were both Russian. His high-school years were spent in Georgia and for his university studies he travelled to Moscow. He graduated in 1904, having specialised in mathematics and mathematical physics. By then he had already published several works. Two at least were technical mathematical studies; others were of a religious nature.

He took a second degree in 1908, this time in philosophy, and by then he had written more studies, mostly on religious or philosophical themes. He became a member of the faculty specialising in the history of philosophy and produced yet more works, mostly religio-philosophical, but one a Russian grammar. He married in 1910 and was ordained a priest in 1911. He seems at this time to have passed through a turbulent period of spiritual questioning, wondering if all his intellectual endeavour was worthwhile and expressing longings for a simple monastic life. However, he continued to write books and papers, mainly of a philosophico-religious nature, and gained a Master's degree in 1913. His thesis for this award became the basis for his book *The Pillar and Foundation of Truth*, published in Russian in 1914.

This work, from Flegg's description, is very wide-ranging, and it has been translated into German, Italian and French. Mathematics was definitely a very poor second in his mind at this time, but there were a couple of papers. It would seem that it was the 1917 revolution that drove him back to more

[5] *Vol* 4 (1994), pp. 266-276.

[6] There is also some background in an article in the journal *The Mathematical Intelligencer*, *Vol* 13 (1991) *No* 2, pp. 24-32.

technical disciplines. After the Bolsheviks came to power, the theological academy was closed and he occupied a number of scientific posts. During the 1920s, he did work on theological matters, but his main thrust was in the scientific field. Toward the end of the decade, he was appointed to the editorship of the Soviet Technical Encyclopedia.

He showed conspicuous courage by continuing to dress quite openly as a priest, despite official displeasure. Flegg regards it as likely that the state (and Lenin in particular) regarded his expertise in the scientific arena as too valuable for them to act against him. For instance, when 100 dissident scholars were banished in 1922, Florensky was not among them. However, Lenin died and was succeeded by Stalin, whose methods became more and more ruthless. In 1933, Florensky was sentenced to ten years in the Gulag. He saw internment in at least three different Siberian concentration camps, and according to official sources, died in one of these in late 1943.

Flegg accepts the story that he was murdered on the eve of his release, murdered on official orders. There is no direct evidence for this, but the indirect case is quite strong, and seems to be accepted by the Russian church. He was canonised in 1981, and is regarded as a martyr for his faith. In Russia and the former Soviet Union, history has long been very much a matter of who is in power and what is the official line. The current edition of their *Philosophical Encyclopedia* has a favourable article on Florensky, including the words "In 1933 he was repressed, and posthumously rehabilitated in 1956". It would seem that there is much current interest in his work, and posthumous editions of his works continue to appear.

It is very clear that he was active in a very broad range of intellectual and spiritual activity, and this has led, according to Flegg, to his being termed "The Russian Leonardo da Vinci". It is also clear that he lived a life of heroic sanctity and thus merited his title of "Saint". What makes me a little uneasy about the account of his life and indeed of similar accounts of other such lives, however, is the tendency of their authors to adopt an uncritical attitude toward their subject.

We might suggest that the argument goes a bit like this:

Florensky was a good man
Florensky was a mathematician
Therefore Florensky was a good mathematician.

It is never this explicit, of course, but it is a worry and indeed, we see it in other contexts. In this case, assessment is difficult. Flegg lists some dozen or so mathematical works, mostly papers, as I would judge. These deal with discontinuous functions, set theory, Cantor's transfinite numbers[7], and various other topics. My hunch is that the later work, produced in the 1920s, is mainly of an expository nature. Flegg quotes no mathematical results associated with his work, and indeed gives no real detail of any of his mathematical writing.

A little more background is supplied by Charles Ford in the article referenced in footnote 6. This deals in more detail than does Flegg with the persecution of the president of the Moscow Mathematical Society, Dmitri Egorov, who was systematically victimised through the late 1920s, and finally expelled from Moscow in 1930. It would not be putting it too strongly to say that he was hounded to his death in 1931.[8]

Egorov's principal disciple was Nicolai Luzin, a considerable mathematician and also a good friend of Florensky. Ford gives quite some detail of the friendship, and it would seem that Luzin was almost persuaded *away* from mathematics by Florensky. It was Egorov who persuaded him to apply his talents to the things he did best. There is available in English a detailed account of Luzin and the Moscow mathematical school.[9] It does not mention Florensky, and so I tend to think that his influence on mathematics itself must really have been quite slight.

It is interesting to compare Florensky and Salmon. The latter, like the former, began life as a mathematician and probably would have remained such, had he not been ambitious for advancement, advancement he could not find while he remained just a mathematician. Florensky, by contrast, abandoned mathematics, only to resume it as a means to earn a living when those who espoused the religious life came under persecution.

* * * * *

[7]See *Function, Vol* 2 *Parts* 1, 2.

[8]One sad aspect of the persecution of Egorov is the part played in it by two very eminent mathematicians, Pontryagin and Gelfond. One would wish things were not so. Just as there is the erroneous "syllogism" quoted earlier, so there is another:

Pontryagin and Gelfond were good mathematicians
Pontryagin and Gelfond were men
Therefore Pontryagin and Gelfond were good men.

This is of course equally invalid. However, it behoves us in judging such cases to ask ourselves how heroically *we* would act under a tyrannical dictatorship.

[9]See the article by Esther Phillips in *Historia Mathematica, Vol* 5 (1978), pp. 275-305.

A Brief Update

Appendix C of my column in *Function*, *Vol* 20 *Part* 3 gave a proof of Fasbender's Theorem that I found in Ivan Niven's *Maxima and Minima without Calculus.* I referred to it as "Niven's Proof" and in footnote 8 said that it "would seem to be Niven's own". However, Andy Liu of the University of Toronto has written to say that it goes back further. Niven's book appeared in 1981, but the proof is to be found on p. 208 of *Convex Figures* by I M Yaglon and V G Boltyanskiĭ. This is a translation of a Russian original dating from 1951. Professor Liu speculates that the proof, which is not sourced (i.e. Yaglon and Boltyanskiĭ do not say where they got it from, any more than does Niven), may indeed be earlier than this.

* * * * *

Paul Erdös

Paul Erdös was a Hungarian mathematician who became a globetrotting citizen of the world. He loved to pose and solve problems and was legendary for his creativity and his energy. He published some 2000+ papers on some counts and had contacts around the world. His links with Australia were strong and long-standing. Among his close associates here was Marta Sved, who served on *Function*'s editorial board from 1989 to 1991. Erdös also featured in the story of Louis Pósa in our very first issue. The solution published in this present issue to the long-outstanding Problem 15.1.8 is due to Erdös.

Sadly, Erdös died last year – fittingly at a mathematics conference. He suffered a heart attack on September 20 at a Warsaw hotel. It is a special loss to Australia for, had he lived, he would have been here among us as this issue of *Function* goes to press.

* * * * *

No human investigation can claim to be scientific if it doesn't pass the test of mathematical proof.

– Leonardo Da Vinci

* * * * *

COMPUTERS AND COMPUTING

Modelling Shapes with Polynomials

Cristina Varsavsky

A drawing program is an essential tool for graphic designers who need to draw objects of many different shapes. Drawing curved shapes is one of the most important features of a drawing computer program. The simplest facility is the freehand drawing tool – you draw the curve by moving the mouse or the cursor keys; but if you have tried this technique you would certainly agree that it rarely produces a nice smooth curve. Freehand drawing is also very inefficient as all points of the curve must be stored to reproduce the curve.

More advanced computer packages have a drawing tool which uses Bézier curves[1]. These have their origin in car design; the French mathematician P E Bézier introduced them while working in the automotive industry. Bézier curves are mathematically speaking very simple; the x and y components of each point on the curve are polynomials depending on a parameter t; only cubics will be considered in this article[2]. These cubics provide an intuitive feel for curve design; their coefficients are directly related to the control points defining the curve. Figure 1 shows the *control polygon* determined by the points P_0, P_1, P_2, and P_3, and the corresponding cubic Bézier curve. The points P_0 and P_3 belong to the curve – they are the starting and ending points – while the points P_1 and P_2 determine the shape of the curve.

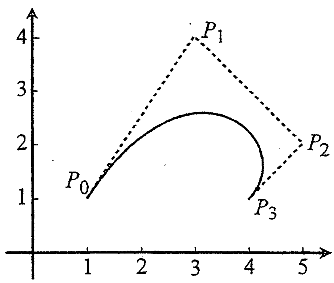


Figure 1

[1]See also *Function, Vol* 15 *Part* 4.
[2]For an introduction to parametric curves see *Function, Vol* 20 *Part* 5.

The cubic parameterisation of the Bézier curve $C(t) = (x(t), y(t))$ is:

$$(x(t), y(t)) = (1-t)^3 P_0 + 3t(1-t)^2 P_1 + 3t^2(1-t)P_2 + t^3 P_3, \quad 0 \le t \le 1. \quad (1)$$

Let us interpret equation (1) with our example in Figure 1. The vertices of the control polygon are

$$P_0 = (1,1), \quad P_1 = (3,4), \quad P_2 = (5,2) \text{ and } P_3 = (4,1),$$

so

$$C(t) = (1-t)^3(1,1) + 3t(1-t)^2(3,4) + 3t^2(1-t)(5,2) + t^3(4,1), \quad 0 \le t \le 1$$

That is,

$$x(t) = (1-t)^3 + 9t(1-t)^2 + 15t^2(1-t) + 4t^3 = -3t^3 + 6t + 1$$
$$y(t) = (1-t)^3 + 12t(1-t)^2 + 6t^2(1-t) + t^3 = 6t^3 - 15t^2 + 9t + 1$$

The definition given in equation (1) is a special combination of the points P_0, P_1, P_2 and P_3, where the coefficients of the combination, which are functions of the parameter t, add up to one. Note that these coefficients are the terms of the binomial expansion of $[(1-t)+t]^3$. When $t = 0$, $(x(0), y(0)) = P_0$, and when $t = 1$, $(x(1), y(1)) = P_3$. As the value of the parameter t moves away from 0 and towards 1, the point P_0 has less weight in the definition of the curve while the influence of the other three points becomes greater. Figure 2 shows the graphs of the four coefficient functions of t. The coefficient of P_1 has greater dominance around $t = 1/3$ where it achieves its maximum, while the coefficient of P_2 achieves its maximum at $t = 2/3$ where it dominates the other points in the definition of the shape of the curve.

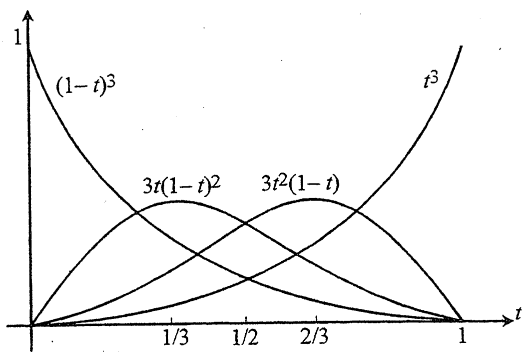


Figure 2

Let us now find the slope of the curve. This is given by the derivative with respect to t of each component of $C(t)$:

$$C'(t) = -3(1-t)^2 P_0 + 3(t-1)(3t-1)P_1 + 3t(2-3t)P_2 + 3t^2 P_3$$

Then $C'(0) = 3(P_1 - P_0)$ and $C'(1) = 3(P_3 - P_2)$, which means that the control polygon is tangent to the curve at the end points P_0 and P_3.

Bézier curves allow the graphic designer to have a good feel for the shape of the curves as there is a direct relationship between this shape and the chosen control polygon. Editing the curve shape is also very straightforward, as only the control points have to be moved around. Cubic Bézier curves can take many different shapes including loops and S-shaped curves; a sample of these with the corresponding control polygons is shown in Figure 3. However, cubics do not provide sufficient flexibility to model any desired shape.

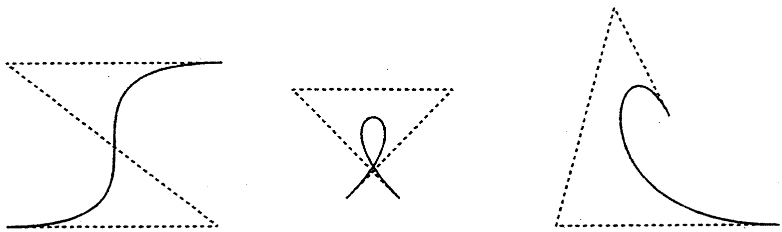


Figure 3

Flexibility can be increased in two ways: by increasing the degree of the polynomial or by joining several cubic Bézier curves. Increasing the degree of the polynomials by one is equivalent to adding one more control point. When joining two Bézier curves $C(t)$ and $D(t)$, as shown in Figure 4, not only must the end point of $C(t)$ coincide with the starting point of $D(t)$, but also the two curves have to be connected with the same slope if a smooth joint is needed. Therefore $C'(1) = D'(0)$, which means that

$$3(P_3 - P_2) = 3(M_1 - M_0).$$

Therefore the common point is half way between the two nearest points P_2 and M_1.

In general, it is also desirable to have the same curvature at the joint. This will mean that the second derivatives have to agree, that is, $C''(1) = D''(0)$. Can you find out the condition on the points to achieve the same curvature at the joint?

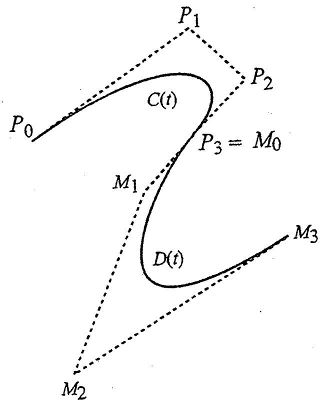


Figure 4

I leave to you the exercise of writing a program which upon the input of control points will draw the corresponding Bézier curve on the screen. Can you use the program to draw the cartoon in Figure 5?

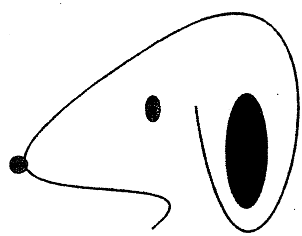


Figure 5

* * * * *

LETTERS TO THE EDITOR

Dear Editor,

I enjoyed reading M Deakin's article 'How to Calculate Cube Roots with Square Roots' (*Function, Vol* 20 *Part* 4, pp. 125-127). It was fun using an iterative process on a simple calculator and actually seeing the convergence with your own eyes. Although the result is no big feat these days, this illustration should provide an excellent introduction to iterative processes in the classroom. After studying Deakin's article I began to wonder whether you could extend the method. Indeed you can. The idea extends quite naturally to finding roots corresponding to any power of $2^n - 1$ for n greater or equal to 2.

For simplicity, let us use $n = 3$ (i.e. to find a seventh root) and pick, say 5. For convenience let us use the following abbreviations:

(1) $\surd$ for the square root key

(2) $\times$ for the multiplication key

(3) $=$ for the result key

(4) M for the memory in key

(5) R for the memory recall key.

We now proceed as follows.

Step 1. Key in: $5, M, \surd$, which enters 5 into the memory and takes $\sqrt{5}$ as x_0, so here $x_0 = 2.236068$ correct to 7 significant figures.

Step 2. Key in: $\times, R, =, \surd, \surd, \surd$ provides the first iteration:

$$x_1 = 1.352250.$$

Now repeat step 2 until convergence is reached.

We find in turn: $x_2 = 1.269853$, $x_3 = 1.259913$, $x_4 = 1.258676$, $x_5 = 1.258521$, $x_6 = 1.258502$, $x_7 = 1.258499$, and that's as far as we need to go for 7 significant figures. Hence the seventh root of 5 is x_7. By the way, the ε rule Deakin found on page 126 for cube roots may also be applied here, but here the error ε will be reduced by a factor of 8. This is best seen by finding the 7th root of 2187 whose result we know to be 3. Starting with $x_0 = \sqrt{2187}$ we eventually find that $x_4 = 3.002012$, so by our present ε rule

$$x_5 = 3 + 0.002012/8 = 3.000252,$$

whilst the fifth proper iterate is actually found to be 3.000251.

Should we wish to find a root corresponding to $n = 4$ (i.e. the 15th root), we need to replace step 2 above by: $\times, R, =, \sqrt{}, \sqrt{}, \sqrt{}, \sqrt{}$, and the ε rule for estimating the next iterate by a factor of 16. Thus the higher our n, the faster will be our convergence.

Julius Guest

Dear Editor,

Recently, consecutive issues of *Function* (*Vol* 20 *Part* 3, p. 76 and *Part* 4, p. 115) contained instances of *arithmetico-geometric* series. In each instance the sum of the series was obtained by differentiating an appropriate geometric series. We now present a non-calculus approach to sum such series. This is akin to the summation of familiar geometric series. Let's begin with the finite arithmetic series,

$$a + (a+d) + (a+2d) + \ldots + [a + (k-1)d]$$

and the finite geometric series

$$1 + r + r^2 + \ldots + r^{k-1}.$$

We multiply each term of the arithmetic series by the corresponding term of the geometric series to form the (finite) arithmetico-geometric series. Let S_k denote this sum. That is,

$$S_k = a + (a+d)r + (a+2d)r^2 + \ldots + [a + (k-1)d]r^{k-1}. \qquad (1)$$

Multiply both sides of equation (1) by r.

$$rS_k = ar + (a+d)r^2 + \ldots + [a + (k-2)d]r^{k-1} + [a + (k-1)d]r^k. \qquad (2)$$

Subtraction of (2) from (1) yields

$$\begin{aligned}(1-r)S_k &= a + dr + dr^2 + \ldots + dr^{k-1} - [a + (k-1)d]r^k \\ &= a + dr(1 + r + r^2 + \ldots + r^{k-2}) - [a + (k-1)d]r^k.\end{aligned}$$

The middle part, $1 + r + r^2 + \ldots + r^{k-2}$, equals $\dfrac{1-r^{k-1}}{1-r}$ by the familiar geometric series summation formula. Hence

$$\begin{aligned}(1-r)S_k &= a + \frac{dr(1-r^{k-1})}{1-r} - [a+(k-1)d]r^k,\\ S_k &= \frac{a}{1-r} + \frac{dr(1-r^{k-1})}{(1-r)^2} - \frac{[a+(k-1)d]r^k}{1-r},\ r \neq 1. \qquad (3)\end{aligned}$$

When $-1 < r < 1$, r^{k-1}, r^k, and $(k-1)r^k$ approach 0 as k approaches infinity and we have

$$S_\infty = a + (a+d)r + (a+2d)r^2 + \ldots = \frac{a}{1-r} + \frac{dr}{(1-r)^2}.$$

In particular, for $1 + 2r + 3r^2 + \ldots$, we have $a = d = 1$ and

$$S_\infty = \frac{1}{1-r} + \frac{r}{(1-r)^2} = \frac{1}{(1-r)^2},\ -1 < r < 1.$$

K R S Sastry

Dear Editor,

Regarding the question posed in the article "The front cover, Subdivision of Triangles" (*Function, Vol* 20 *Part* 3), the following elegant proof, due to Michael Yoder, shows by *reductio ad absurdum* that at most 3 copies can exist. N here represents the number of copies that can exist. This was n in the original article.

Assume $N \geq 4$. Then by induction we will show that for every $n \geq 0$, there are at least $2n+4$ triangles (either further divided, or not) whose area is expressible as a degree n term in a and b (i.e., $a^r b^s$ where $r+s=n$ and r, s are nonnegative integers). This means we have an infinity of triangles, so $N \leq 3$.

For $n = 0$, we have at least 4, so the base of the induction holds.

Otherwise, suppose this is true for $n = m$. Of the $2m+4$ triangles, at most $m+1$ are not further subdivided (one each for $a^n, a^{n-1}b, \ldots, b^n$), so at least $m+3$ are further subdivided; this means there are at least $2(m+3) = 2(m+1)+4$ triangles whose area is expressible as a degree $m+1$ term.

Derek Garson
Lane Cove, NSW

For another proof of this result, see the problems section of this issue. – Ed

PROBLEM CORNER

SOLUTIONS

PROBLEM 20.4.1 (K R S Sastry, Dodballapur, India)

Show that the graph of the polynomial $p(x) = x^4 - 2x^2 + 2x + 2$ has a common tangent line at two distinct points on it.

SOLUTION

The problem can be solved in many ways. One way is to rewrite the polynomial as follows:

$$x^4 - 2x^2 + 2x + 2 = (x^2 - 1)^2 + 2x + 1.$$

It is easy to see that the graph of $y = (x^2 - 1)^2$ has the x-axis as a common tangent line at the two points $(-1, 0)$ and $(1, 0)$. By adding $2x + 1$, we simultaneously transform $y = (x^2 - 1)^2$ into the given polynomial, and the x-axis into the line $y = 2x + 1$. Hence $y = 2x + 1$ is the common tangent line.

This problem turned out to be popular with readers. Solutions, using a variety of methods, were received from Mike Hassall (Daylesford Secondary College, Vic), Peter Bullock (Norwood Secondary College, Vic), Claudio Arconcher (São Paulo, Brazil), John Barton (Carlton North, Vic), and the proposer. K R S Sastry showed that the result can be generalised to a wider class of polynomials, and John Barton found a characterisation of all quartic polynomials with the required property.

PROBLEM 20.4.2 (from *Yidiot Achronot* newspaper, Israel; posted on the Internet by Greg Barron)

Use each of the numbers 1, 5, 6 and 7 exactly once, together with the four basic operations of arithmetic $(+, -, \times, /)$ and parentheses, to obtain an expression equal to 21. The numbers may be used in any order, and there is no restriction on the number of times each operator may be used.

SOLUTION

The answer is $6/(1 - 5/7)$.

PROBLEM 20.4.3 (Claudio Arconcher, São Paulo, Brazil)

Let $ABCD$ be a convex quadrilateral. Let P be a point inside $ABCD$. From P, draw perpendiculars to the sides of $ABCD$, extending them outside the quadrilateral: $p_1 \perp \overline{AB}$, $p_2 \perp \overline{BC}$, $p_3 \perp \overline{CD}$, and $p_4 \perp \overline{DA}$. Let Q_1 be a point on p_1 outside $ABCD$. Construct Q_2 on p_2 with $BQ_1 = BQ_2$, Q_3 on p_3 with $CQ_2 = CQ_3$, and Q_4 on p_4 with $DQ_3 = DQ_4$.

(a) When this construction is possible, is it always true that $AQ_4 = AQ_1$?

(b) Investigate what happens if Q_1 moves along p_1 towards $\overline{AB}$ until one of Q_1, Q_2, Q_3 and Q_4 reaches the quadrilateral. Under what conditions on the quadrilateral do all four points reach the quadrilateral simultaneously?

SOLUTION by Claudio Arconcher

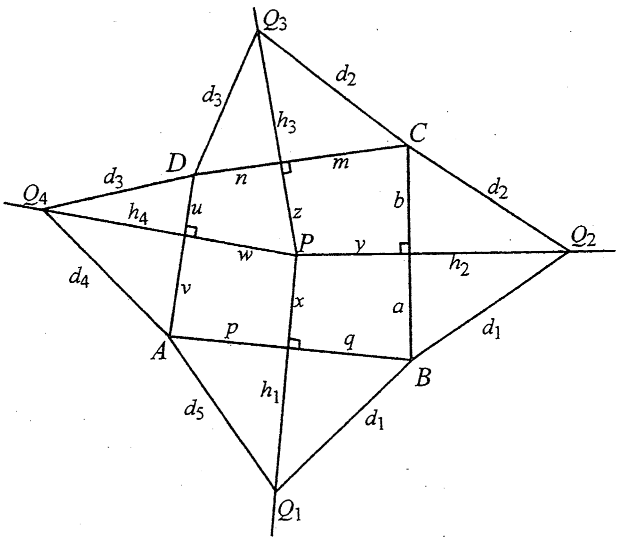


Figure 1

From Figure 1 we obtain:

$$\begin{aligned} x^2 + q^2 &= a^2 + y^2 \\ y^2 + b^2 &= m^2 + z^2 \\ z^2 + n^2 &= u^2 + w^2 \\ w^2 + v^2 &= p^2 + x^2 \end{aligned}$$

Adding these four equations together and simplifying, we obtain:

$$q^2 + b^2 + n^2 + v^2 = a^2 + m^2 + u^2 + p^2. \tag{1}$$

Referring again to Figure 1, we get successively:

$$d_1^2 = h_1^2 + q^2$$

$$d_2^2 = h_2^2 + b^2 = d_1^2 - a^2 + b^2 = h_1^2 + q^2 - a^2 + b^2$$

$$d_3^2 = h_3^2 + n^2 = d_2^2 - m^2 + n^2 = h_1^2 + q^2 - a^2 + b^2 - m^2 + n^2$$

$$d_4^2 = h_4^2 + v^2 = d_3^2 - u^2 + v^2 = h_1^2 + q^2 - a^2 + b^2 - m^2 + n^2 - u^2 + v^2$$

Therefore, by (1), $d_4^2 = h_1^2 + p^2$, so $d_4^2 = d_5^2$. Hence $d_4 = d_5$, as required.

If Q_1 moves along p_1 towards $\overline{AB}$, then Q_2, Q_3 and Q_4 must also move towards their respective sides of the quadrilateral if the conditions are to remain satisfied. The first of these points, Q_i, to reach its corresponding side is on the line p_i for which the part of the side of $ABCD$ from p_i to one of the two adjacent vertices is the longest. If $ABCD$ is a cyclic quadrilateral with P at the centre of its circumcircle, then all four points will reach the quadrilateral simultaneously.

The problem arises in the context of cutting a sheet of paper into a shape that can be folded into a pyramid with an arbitrary quadrilateral as its base.

PROBLEM 20.4.4 (Juan-Bosco Romero Márquez, Universidad de Valladolid, Valladolid, Spain)

Find positive integer solutions to the equation $x + y + xy = x^2 + y^2$.

SOLUTION

This is another problem that can be approached in a number of different ways. One method is as follows. Let $u = x + y$ and $v = x - y$. Then u is a natural number, v is an integer, and $x = \frac{u+v}{2}$ and $y = \frac{u-v}{2}$. Therefore the equation becomes:

$$u + \left(\frac{u+v}{2}\right)\left(\frac{u-v}{2}\right) = \left(\frac{u+v}{2}\right)^2 + \left(\frac{u-v}{2}\right)^2$$

Expanding the terms yields:

$$u + \frac{u^2 - v^2}{4} = \frac{u^2 + 2uv + v^2 + u^2 - 2uv + v^2}{4}$$

On multiplying through by 4 and simplifying, we obtain:

$$4u = u^2 + 3v^2$$

By completing the square with respect to u, we can write this equation in the form:

$$4 = (u-2)^2 + 3v^2$$

It follows that $3v^2 \leq 4$, and so $v^2 \leq 4/3$. Since v is an integer, v must equal $-1, 0$ or 1. If $v = -1$ then either $u = 1$, which gives a value of 0 for x, or $u = 3$, which yields the solution $x = 1, y = 2$. If $v = 0$ then either $u = 0$ or $u = 4$; again the first option can be disregarded since it gives $x = 0$, while the second option gives the solution $x = 2, y = 2$. If $v = 1$ then either $u = 1$ or $u = 3$; in this case too, the first option does not yield a valid solution, while the second gives $x = 2, y = 1$. Therefore there are three solutions: (1,2), (2,2) and (2,1).

This was a popular problem. Solutions, using various approaches, were received from Mike Hassall (Daylesford Secondary College, Vic), Peter Bullock (Norwood Secondary College, Vic), Claudio Arconcher (São Paulo, Brazil) and the proposer.

More on some earlier problems

Andy Liu, of the University of Alberta, Canada, sent comments on a number of the problems. He supplied a simpler and more elegant solution to Problem 16.4.4 (the triangle subdivision problem) than the one described in the front cover article in *Function, Vol* 20 *Part* 3. Suppose four copies of the triangle could be produced. In the terminology of the article, the subdivision could contain at most one undivided triangle at level zero (i.e., one of the four copies left undivided), and at most two undivided triangles at level one (either both in the same copy or in two different copies). Since there are four copies of the original triangle altogether, this would leave at least four level one triangles which must be subdivided further. But this is now the start of an "infinite descent", since these four triangles must themselves be subdivided using the same rules as applied to the original four copies. Therefore there cannot be four copies of the triangle, so at most three copies are possible.

The problem appeared in the Spring 1995 Junior A-Level paper of the International Mathematics Tournament of the Towns.

Andy Liu pointed out that the arithmetico-geometric series in the article (page 76) can be evaluated without using differentiation. This is also done by K R S Sastry in his letter, which appears elsewhere in this issue. Finally, Andy Liu noted that Problem 20.1.7 is essentially Theorem 14 in Stan Ogilvy's *Excursions in Geometry*, pp 57-59.

One problem from the previous issue, Problem 20.5.6, will benefit from some clarification. (A solution will be given in the next issue.) Part (c) of the problem asked readers to "resolve the apparent paradox". If you didn't see anything paradoxical about the answers to parts (a) and (b) (and there was no reason why you should have!), you would be at a loss to know what was being asked. A better question would have simply asked readers to generalise the problem, leaving it to them to resolve any paradox if they encountered one.

Here is the solution to a problem from the February 1991 issue of *Function*, for which we have not previously provided a solution.

PROBLEM 15.1.8

An infinite set of points in the plane is such that the distance between any two is an integer. Prove that the points are all collinear.

SOLUTION originally due to the eminent Hungarian mathematician Paul Erdös; posted on the Internet by Robert Israel, University of British Columbia, Canada

Let S be a set of points satisfying the conditions of the problem. Suppose S contains three points, A, B and C, not all on the same line. Let Q denote an arbitrary member of S. By the triangle inequality, the distances AB, AQ and BQ satisfy $|AQ - BQ| \leq AB$, so $AQ - BQ$ is one of the integers from $-AB$ to AB. For any given integer k in the range from $-AB$ to AB, the points Q for which $AQ - BQ = k$ lie on a branch of a hyperbola (or one of its degenerate cases, a line parallel or perpendicular to $\overline{AB}$). Thus we have a finite family of curves associated with A and B, and every point in S is on one of these curves. Similar finite families of curves are associated with A and C, and with B and C, and each point Q in S is on one of the curves in each family. Hence, all of the points in S fall where members of the three families of curves intersect. But there can be only a finite number of such points, because two hyperbolas intersect (in general) in at most two points. (In cartesian coordinates, a hyperbola is given by a second degree relation, so the intersection of two hyperbolas is found by solving a quadratic equation; hence there can be at most two solutions. The possibility that two curves from different families might coincide can be eliminated by using the fact that A, B and C are not collinear.) Hence S contains a finite number of points.

Although we cannot find an *infinite* number of non-collinear points with each pair of points an integer distance apart, it is possible to find an *arbitrarily large finite* set of non-collinear points satisfying this condition (see Problem 21.1.1 below).

PROBLEMS

Readers are invited to send in solutions (complete or partial) to any or all of these problems. All solutions received by 14 April 1997 will be acknowledged in the June issue, and the best solutions will be published.

PROBLEM 21.1.1

Show that it is possible to find an arbitrarily large finite set of points in the plane such that the points are not all collinear and the distance between any two is an integer. (Rather than trying to do this in one step, you may find it easier if you first look for sets of points for which the distance between any two points is a rational number, and then rescale.)

PROBLEM 21.1.2 (from the German mathematics magazine *Alpha*, May/June 1996)

Solve the equation $x^3 - 3y = 2$ in natural numbers.

PROBLEM 21.1.3 (Claudio Arconcher, São Paulo, Brazil)

Let a and b denote real numbers. Find necessary and sufficient conditions over a and b such that:

$$ax + b\lfloor x \rfloor = ay + b\lfloor y \rfloor \text{ if and only if } x = y$$

where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x.

PROBLEM 21.1.4

Let $\triangle ABC$ be a triangle with $AB \neq AC$. Let D be the point of intersection of the angle bisector of A and the perpendicular bisector of $\overline{BC}$. Prove that D is on the circumcircle of $\triangle ABC$.

PROBLEM 21.1.5

Let $f(x)$ be a cubic polynomial with three distinct real roots: a, b and c. Let $v = \frac{a+b}{2}$. Prove that the tangent to $f(x)$ at v passes through $(c, 0)$.

* * * * *

OLYMPIAD NEWS

The 1996 AMOC Senior Mathematics Contest

This Australia-wide competition, designed for mathematically highly gifted year-eleven students, took place on 13 August 1996. The contest paper consisted of five questions, and the 80 participants were given four hours to tackle them. The Australian Mathematical Olympiad Committee (AMOC) invited the most successful students to participate at the problem-oriented School of Excellence, which was held at Melbourne last December. Here are the problems:

1. Let k be a semicircle with diameter AB. Let D be a point such that $AB = AD$ and AD intersects k (at the point E, say). Let F be the point on chord AE such that $DE = EF$. Let BF (extended) meet k at the point C.
   Show that $\angle BAE = 2\angle EAC$.

2. Find all functions f, defined for all real numbers and taking real numbers as values, which satisfy, for all real numbers u and v:
$$f(u+v)f(u-v) = 2u + f(u^2 - v^2).$$

3. Let x be a non-zero real number such that $x + \frac{1}{x}$ is an integer. Prove that $x^n + \frac{1}{x^n}$ is an integer for all positive integers n.

4. The sequence $a_0, a_1, a_2, \ldots, a_{1997}$ has the properties:
   (i) $0 \leq a_n \leq 1$ for all $n \geq 0$,
   (ii) $a_n \geq \dfrac{a_{n-1} + a_{n+1}}{2}$ for all $n \geq 1$.

   (a) Prove that $a_{1997} - a_{1996} \leq \frac{1}{1997}$.
   (b) Find a sequence satisfying (i) and (ii) such that $a_{1997} - a_{1996} = \frac{1}{1997}$.

5. Let ABC be an acute triangle with $\angle ACB = 60°$. Let h_a be the altitude through A and h_b the altitude through B. Prove that the circumcentre of $\triangle ABC$ lies on the bisector of one of the two angles formed by h_a and h_b.

* * * * *

Published by Department of Mathematics and Statistics, Monash University