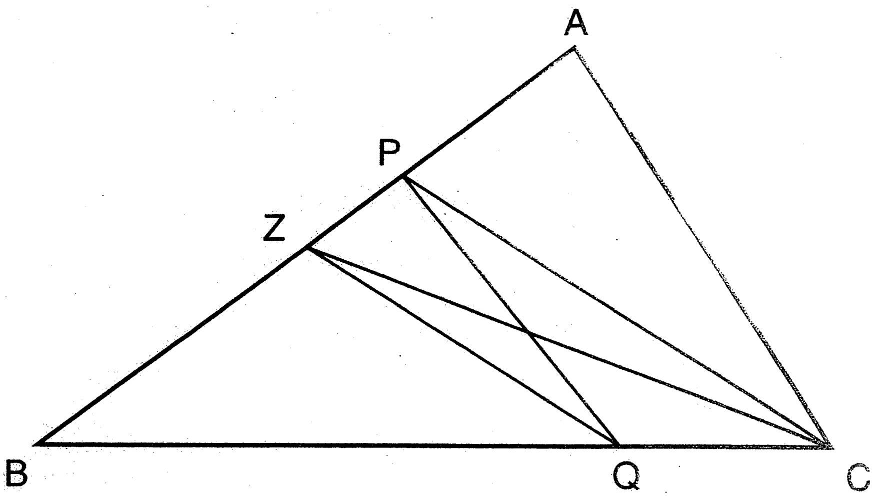
Function

A School Mathematics Magazine

Volume 19 Part 2 **April 1995**



Mathematics Department - Monash University

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

Function is a mathematics magazine produced by the Department of Mathematics at Monash University. The magazine 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, Monash University
900 Dandenong Rd
Caulfield East VIC 3145, Australia
Fax: +61 (03) 903 2227
e-mail: function@maths.monash.edu.au

Alternatively, correspondence may be addressed individually to any of the editors at the mathematics departments of the institutions listed on the inside back cover.

Function is published five times a year, appearing in February, April, June, August, and October. Price for five issues (including postage): $17.00*; single issues $4.00. Payments should be sent to: The Business Manager, *Function*, Mathematics Department, Monash University, Clayton VIC 3168; cheques and money orders should be made payable to Monash University. Enquiries about advertising should be directed to the Business Manager.

*$8.50 for *bona fide* secondary or tertiary students.

EDITORIAL

We welcome new and old readers alike with this issue of *Function* and we hope you find it interesting and enjoyable.

In our main article, John Shanks gives us an account of the classical Archimedes' method for calculating π which uses the perimeters of inscribed and circumscribed regular polygons on a unit circle. The article continues with a remarkable update due to the German mathematician W Romberg. Using Romberg's method with Archimides' data, an accuracy of 16 decimal places can be achieved in just a few steps. If only Archimedes and the many mathematicians who followed him had known this!

The History column is dedicated to Roman numerals which are still used on the faces of clocks, on early pages of some books and in giving the dates of films. Michael Deakin analyses the syntax of this system, listing the rules for constructing a meaningful string. We encourage you to solve the problems included at the end of the article.

This time, you will need a long strip of paper to read the Computers and Computing section, as an algorithm is presented to construct, first with the strip and then on your computer screen, the so-called *dragon curves.*

We hope you find the Problem Corner both challenging and enjoyable. You could also try the problems set for the 1995 Australian Mathematical Olympiad.

We include also a rather intriguing letter which reached the editor: it shows how a particular series leads to the most amazing and unbelievable results. We are always grateful to the active readers who send us letters, solutions to problems, suggestions and – why not – articles.

* * * * *

THE FRONT COVER

Cutting up Triangles

Michael A B Deakin

The front cover for this issue (reproduced here as Figure 1) shows a euclidean construction for dividing a triangle into two pieces of equal area.

ABC is a given triangle and P is a point on its perimeter. It will not significantly limit the problem to suppose that P lies on the side AB. We want to construct a point Q so that the line PQ cuts the triangle into two pieces of equal area.

The construction given here is that from Hall and Stevens's *A School Geometry* (first published in 1903).

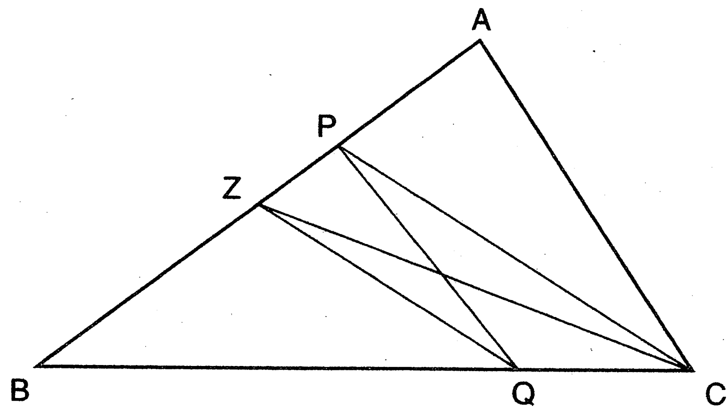


Figure 1

Let Z be the mid-point of AB. It will not significantly limit the problem to suppose that P lies in the interval AZ (rather than in the interval BZ).[1]

Now join PC. Through Z draw a line ZQ parallel to PC and meeting the side BC in Q. The line PQ then bisects the area of the triangle ABC.

To see why this is so, note first that if we join ZC we bisect the area ABC, for the triangles AZC, BZC have equal bases AZ, BZ and the same perpendicular height (the length of the perpendicular from C to AB).

[1] Of course it *could* happen that P *coincided* with Z. In this case the required line is ZC as Q will coincide with C. This comes about as a result of the argument developed in the more general case.

Now consider the triangles PZC and PQC. These two have equal area as they rest on the *same* base PC and their perpendicular heights are equal, being both equal to the distance between the parallels PC, ZQ.

Then

$$\begin{aligned}\tfrac{1}{2}\text{Area }(\triangle ABC) &= \text{Area }(\triangle AZC)\\ &= \text{Area }(\triangle AZC) - \text{Area }(\triangle PZC) + \text{Area }(\triangle PQC)\\ &= \text{Area }(APQC).\end{aligned}$$

Figure 2 gives the diagram for a similar but more difficult problem. This time we are given the triangle ABC and a point X lying on one side, which we may take to be BC. We now seek lines XH and XK that *trisect* the area ABC – i.e. cut it into three equal parts.

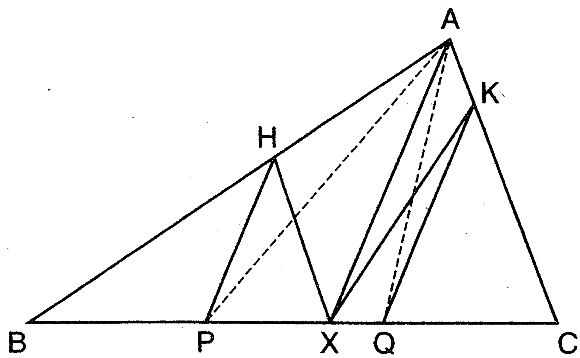


Figure 2

In the situation diagrammed P, Q trisect the side BC, i.e. $BP = PQ = QC$, and X lies in the interval PQ. We join AX and construct PH and QK through P, Q respectively and both parallel to AX. H is to lie on AB and K on AC. Finally join XH and XK. These lines will trisect the area ABC.

This time there are other cases to consider. What if X lay not in the interval PQ but in (e.g.) BP? How would the construction proceed then?

Can you generalise the problem and show how to cut a triangle into four, five, etc. equal areas?

* * * * *

ARCHIMEDES, ROMBERG AND PI

John A Shanks, Dunedin, New Zealand

There are numerous techniques for estimating the number π. Most modern methods use various infinite sums or infinite products discovered over the last few hundred years. This article, however, discusses one of the earliest methods of estimation often referred to as the *classical method*, developed by Archimedes and used by many mathematicians since. We then jump forward over 2000 years and examine a simple but remarkable technique due to the German mathematician W Romberg, which allows us to produce a very accurate estimate of π from Archimedes' data.

Archimedes of Syracuse (287-212 BC) was the greatest mathematician of the ancient world, and perhaps the most famous of his mathematical achievements was the "measurement of the circle". The crux of this problem is the calculation of the number π, the ratio of the circumference to the diameter of a circle.[1] In the ancient Orient π was frequently taken as 3 (II Chronicles 4.2: *He made a molten sea of ten cubits from brim to brim, ... and a line of thirty cubits did compass it round about.*) and the Egyptians were known to use $\pi = (4/3)^4 = 3.1604\ldots$. The first scientific attempt to compute π appears to be that of Archimedes, who calculated the perimeters of regular polygons which were circumscribed and inscribed around a circle of diameter 1. Such a circle has π as its circumference. Figure 1 shows a circle of diameter 1 and two regular hexagons, one circumscribed and one inscribed. (We will assume that all polygons mentioned in this article are regular.) If α_n and β_n are the perimeters of the n-sided $(n \geq 3)$ circumscribed and inscribed polygons respectively, then clearly

$$\beta_n < \pi < \alpha_n \tag{1}$$

(at least the left inequality is clear, the right inequality may take some thought). The values α_n and β_n bracket the value of π and as n is increased the values of α_n decrease, the values of β_n increase, and hence the approximation given by (1) becomes more accurate.

[1] Although we use "π" in this article, this notation only gained wide acceptance after it was adopted by Euler in 1737, following its introduction by English mathematicians in the early 18th century.

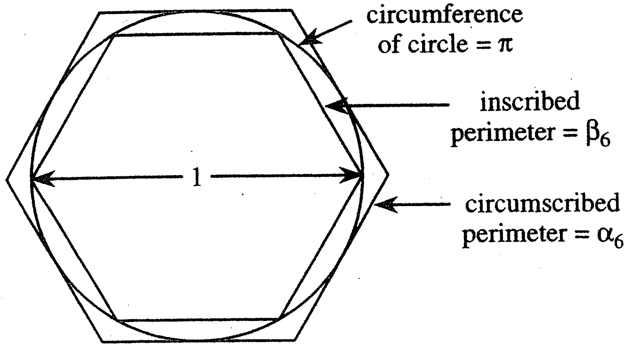


Figure 1

Archimedes' method

Archimedes' greatest achievement here was not the idea of estimating the circumference of the circle by the perimeter of a polygon, but his development of a method for finding the latter. In fact, he found formulae (later called the *Archimedes recurrence formulae*) which allowed the calculation of the perimeters of a whole sequence of inscribed and circumscribed polygons. Each polygon in the sequence has twice the number of sides as the previous polygon. As we will see, Archimedes considered polygons with 6, 12, 24, 48 and 96 sides.

The mathematics relies almost solely on the properties of similar triangles.

Figure 2 shows the arc CMD of a circle together with one side $(\overline{AB})$ of the circumscribed n-sided polygon and one side $(\overline{CD})$ of the inscribed n-sided polygon. Let O be the centre of the circle, $AB = 2v_n$ and $CD = 2u_n$. Let M be the mid-point of $\overline{AB}$ and N the midpoint of $\overline{CD}$. Finally consider the tangent to the circle at C and let this meet the line $\overline{AB}$ at G.

Then $\overline{CM}$ is one side of the inscribed $2n$-sided polygon, whose side lengths we will call $2u_{2n}$. Also, CG and GM are both half of one side length of the circumscribed $2n$-sided polygon, whose sides we will call $2v_{2n}$. Thus we have that $CG = GM = v_{2n}$ and $CM = MD = 2u_{2n}$.

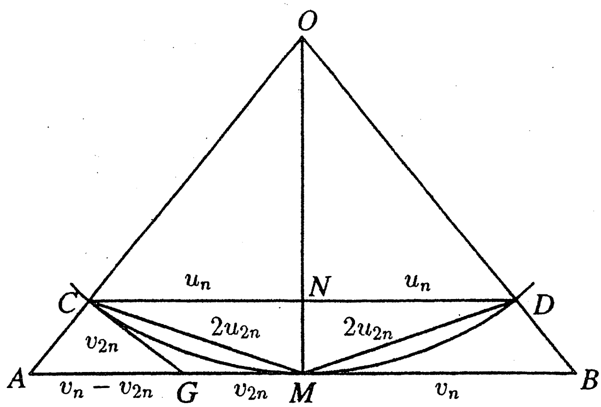


Figure 2

We now consider three pairs of similar triangles. Firstly, triangles AMO and ACG are similar, so

$$\frac{MO}{AO} = \frac{CG}{AG} = \frac{v_{2n}}{v_n - v_{2n}}. \tag{2}$$

Secondly, triangles AMO and CNO are similar, so

$$\frac{CO}{AO} = \frac{CN}{AM} = \frac{u_n}{v_n}. \tag{3}$$

But CO and MO are equal, both being radii of the circle so that the ratios in (2) and (3) are equal, thus

$$\frac{v_{2n}}{v_n - v_{2n}} = \frac{u_n}{v_n}$$

and hence:

$$v_{2n} = \frac{u_n v_n}{u_n + v_n}. \tag{4}$$

Finally, isosceles triangles CDM and CMG are similar (angles $M\hat{C}D$ and $C\hat{M}G$ are equal), so

$$\frac{2u_n}{2u_{2n}} = \frac{2u_{2n}}{v_{2n}}$$

or, equivalently,

$$2u_{2n}^2 = u_n v_{2n} \tag{5}$$

Given that

$$\alpha_n = 2nv_n,\ \beta_n = 2nu_n,\ \alpha_{2n} = 4nv_{2n},\ \beta_{2n} = 4nu_{2n}$$

and introducing these quantities into the relations (4) and (5), we obtain:

$$\alpha_{2n} = \frac{2\alpha_n\beta_n}{\alpha_n + \beta_n},\ \beta_{2n} = \sqrt{\beta_n\alpha_{2n}} \tag{6}$$

The relations in (6) are the *Archimedes recurrence relations*, because once the values of α_n and β_n are calculated, it is possible to calculate α_{2n} and β_{2n}. It's worth noting that α_{2n} is the harmonic mean of α_n and β_n, while β_{2n} is the geometric mean of β_n and α_{2n} (see the article by Ken Evans in *Function, Vol.* 15, *Part* 4, pp. 98-106).

Archimedes started with hexagons, the perimeters of which (α_6 and β_6) are easily calculated from equilateral triangles, namely

$$\alpha_6 = 2\sqrt{3},\quad \beta_6 = 3$$

for a circle of radius $\frac{1}{2}$.

Using (6) with $n = 6$ we find

$$\alpha_{12} = \frac{12\sqrt{3}}{2\sqrt{3}+3} = 12(2-\sqrt{3})$$

and

$$\beta_{12} = 6\sqrt{2-\sqrt{3}}.$$

Applying (6) again, but now with $n = 12$, we have

$$\alpha_{24} = 24(2-\sqrt{3})(2\sqrt{2-\sqrt{3}}-1)(7+4\sqrt{3}).$$

The expression for β_{24} becomes rather complicated.

It is known that Archimedes did not work with such exact expressions. At each calculation of β_n he would have used an approximation to the square root. With this approximation, all Archimedes' calculations produced *rational* numbers, being approximations to the exact values of α_n and β_n. Nevertheless, with some extremely laborious effort (especially so in view of the ciphered numeral system used by the Greeks), Archimedes produced values up to α_{96} and β_{96}.

Table 1 shows the values of the sequence up to β_{96} (correct to 6 decimal places).

n sides	α_n	β_n
6	3.464102	3.000000
12	3.215390	3.105829
24	3.159660	3.132629
48	3.146086	3.139350
96	3.142715	3.141032

Table 1

Archimedes found that $\alpha_{96} \approx 3\frac{10}{70}$ (= 3.142857...) and $\beta_{96} \approx 3\frac{10}{71}$ (= 3.140845...), with his approximations causing inaccuracies in the fourth decimal place when compared to the true values. However, he concluded correctly that

$$3\frac{10}{71} < \pi < 3\frac{10}{70}.$$

Further developments

Many attempts to increase the accuracy in estimating π have been made since. Notable among those using Archimedes' classical method are the following:

- *ca.* 530 AD: The early Hindu mathematician Aryabhata produced the value 62832/20000 = 3.1416, possibly from the classical method with polygons of 384 sides.
- 1579: The French mathematician François Viète found π to 9 decimal places using polygons with 393 216 sides.
- 1593: Adrianus Romanus of the Netherlands found π correct to 15 decimal places using polygons having 2^{30} sides.
- 1610: Ludolph van Ceulen of Germany spent a large part of his life devoted to the calculation of π to 35 decimal places using polygons with 2^{62} sides. This magnificent feat was considered so extraordinary that his approximation was engraved on his tombstone.
- 1621: The Dutch physicist Willebrord Snell (known best for Snell's law of refraction) modified the classical method so that from each pair of bounds on π he produced considerably closer bounds. By his method (later verified by the Dutch mathematician Huygens) van Ceulen's 35 decimal places can be achieved using polygons with only 2^{30} sides.

- 1630: Grienberger used Snell's refinement to compute π to 39 decimals. This was the last major attempt to estimate π using the classical method.

From about 1650, the determination of π has centred on various series and product identities. Gregory's infinite series (1671):

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots \quad (-1 < x \leq 1)$$

yields a very slowly converging series for $\pi/4$ when $x = 1$. In 1699 Abraham Sharp used $x = 1/\sqrt{3}$ to give a series for $\pi/6$, from which he calculated π to 71 decimal places.

Romberg's acceleration method

In 1955 Romberg published a paper concerning the numerical estimation of integrals. The recursive technique he introduced now bears his name and is widely used. As an offshoot, the technique can be used in areas of mathematics other than numerical integration. We will use it to produce vastly improved estimates of π merely by taking a few simple combinations of the polygon perimeters determined by Archimedes.

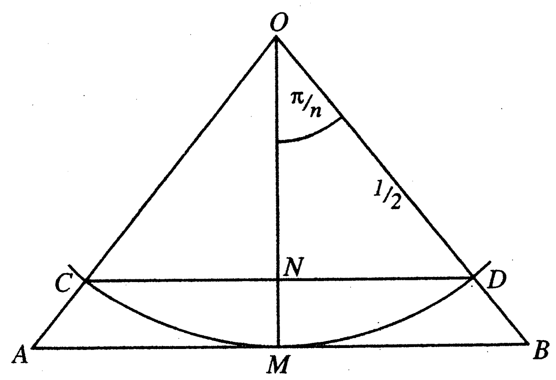


Figure 3

Figure 3 is similar to Figure 2 and shows an arc of a circle of radius $\frac{1}{2}$. If $\overline{AB}$ and $\overline{CD}$ are sides of the circumscribed and inscribed n-sided polygons respectively then the angle $A\hat{O}B$ must be $\frac{2\pi}{n}$. Thus the angle $M\hat{O}B$ is $\frac{\pi}{n}$. In the right-angled triangle NOD, we see that $OD = \frac{1}{2}$ (being a radius) and hence $ND = \frac{1}{2}\sin\frac{\pi}{n}$. Thus $CD = \sin\frac{\pi}{n}$ and the total perimeter of the inscribed n-sided polygon is $\beta_n = n\sin\frac{\pi}{n}$.

Now, $\sin x$ can be expressed as the series[2]

$$\sin x = \frac{x^1}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

Replacing x with $\frac{\pi}{n}$, we have

$$\beta_n = n\left[\frac{\left(\frac{\pi}{n}\right)^1}{1!} - \frac{\left(\frac{\pi}{n}\right)^3}{3!} + \frac{\left(\frac{\pi}{n}\right)^5}{5!} - \frac{\left(\frac{\pi}{n}\right)^7}{7!} + \cdots\right],$$

and defining $s_1 = -\frac{\pi^3}{3!}$, $s_2 = \frac{\pi^5}{5!}$, $s_3 = -\frac{\pi^7}{7!}, \ldots$, we have

$$\beta_n = \pi + \frac{s_1}{n^2} + \frac{s_2}{n^4} + \frac{s_3}{n^6} + \frac{s_4}{n^8} + \cdots \qquad (7)$$

The form of the expansion in (7) involves even powers of n^{-1} and is suitable for Romberg's acceleration technique. (Also from Figure 3, $MB = \frac{1}{2}\tan\frac{\pi}{n}$ and so $\alpha_n = n\tan\frac{\pi}{n}$ giving an expansion for α_n similar to that in (7).) We can regard the terms after the π on the right-hand side of (7) as "error" terms. As n gets larger these all tend to zero and hence β_n tends to π. Romberg's method eliminates these error terms one by one, starting with s_1/n^2, which is the largest. The perimeter of the polygon with double the number of sides is β_{2n}, which must have the expansion (replacing n by $2n$ in (7)):

$$\beta_{2n} = \pi + \frac{s_1}{4n^2} + \frac{s_2}{16n^4} + \frac{s_3}{64n^6} + \frac{s_4}{256n^8} + \ldots \qquad (8)$$

If we define

$$\gamma_n = \frac{4\beta_{2n} - \beta_n}{3},$$

and $r_2 = -\frac{1}{4}s_2$, $r_3 = -\frac{5}{16}s_3$, $r_4 = -\frac{21}{64}s_4, \ldots$, then from (7) and (8) we find

$$\gamma_n = \pi + \frac{r_2}{n^4} + \frac{r_3}{n^6} + \frac{r_4}{n^8} + \ldots \qquad (9)$$

Notice that γ_n has the same form as β_n, but the error terms start at n^{-4} instead of n^{-2}. This leads us to two conclusions: firstly, we can expect that γ_n will be more accurate than β_n (in fact its error will also *decrease* faster since n^{-4} tends to zero faster than n^{-2}). Secondly, it means that we can now "process" γ_n much as we did β_n: we have that γ_{2n} (which is $(4\beta_{4n} - \beta_{2n})/3$), has the expansion:

$$\gamma_{2n} = \pi + \frac{r_2}{16n^4} + \frac{r_3}{64n^6} + \frac{r_4}{256n^8} + \ldots$$

[2]This is called the Maclaurin series expansion for $\sin x$.

and then

$$\delta_n = \frac{16\gamma_{2n} - \gamma_n}{15}$$

succeeds in eliminating the term in n^{-4} (defining the constant q_i much in the same way as we did with s_i and r_i):

$$\delta_n = \pi + \frac{q_3}{n^6} + \frac{q_4}{n^8} + \dots$$

This time δ_n will be more accurate than γ_n, and by continuing in this way, taking simple combinations of previous values we successively eliminate the error terms and hence produce more accurate estimates of π.

Table 2 shows this technique applied to Archimedes' data. The perimeters β_n appear in the first column. The second column contains the values γ_n (positioned midway between the β_n values that they depend on – for example, γ_6 = 3.1411... = (4 × 3.1058... – 3.0000)/3). The third column shows the values δ_n (for example, δ_6 = 3.14159... = (16 × 3.14156... – 3.14110...)/15) and we can extend the table further for two more columns using similar combinations of values. (Here, of course, we have used accurate values for the perimeters β_n, not Archimedes' approximations.)

β_6=3.00000000000				
	3.14110472164			
β_{12}=3.10582854123		3.14159245389		
	3.14156197063		3.14159265355	
β_{24}=3.13262861328		3.14159265044		3.14159265357
	3.14159073296		3.14159265357	
β_{48}=3.13935020304		3.14159265353		
	3.14159253350			
β_{96}=3.14103195089				

Table 2

The improvement across the table is nothing less than incredible. The last value (on the right of the table) is correct to all figures shown; in fact, it would be accurate to 15 decimal places had we started with such accuracy in the left-most column. By taking these special combinations of the very inaccurate data (β_{96} is correct to only 2 decimal places) we produce a tremendous result.

If we can allow ourselves some day-dreaming here, we might suppose that Archimedes had produced his data to 16 decimal places and hit upon the idea of taking certain simple combinations of his values. Then his estimate of π would not have been equalled by using his classical method alone until Romanus's 15 decimal place accuracy in 1593. (This same Romberg scheme applied to Romanus's data, given to greater accuracy, would yield π to about 500 decimal places.)

Further reading

Howard Eves discusses various methods for the computation of π in his book *An Introduction to the History of Mathematics*. The first chapter in this book covers early numeral systems including the Ionic ciphered system used by the Greeks.

The description of Archimedes' derivation is adapted from the book *100 Great Problems of Elementary Mathematics* by Heinrich Dörrie.

The classic *A Short Account of the History of Mathematics* by W W Rouse Ball dates from 1888 but is nevertheless very readable and includes much on early numeral systems and the works of Archimedes.

A discussion of Romberg's acceleration method for numerically estimating integrals can be found in most introductory texts on numerical analysis.

* * * * *

John Shanks was born in England, and received his PhD from Oxford University. He joined the University of Otago, Dunedin, New Zealand in 1974 where he is currently a Senior Lecturer. His research interests are in various branches of numerical analysis; at present, he is working on a problem in motion analysis, studying the centre of rotation in a golf swing.

* * * * *

HISTORY OF MATHEMATICS

Roman Numerals

Michael A B Deakin

The everyday "Arabic" numerals that we all use nowadays came to us relatively recently, and in fact their origin is Indian, although the forms of some of the characters have altered. Prior to their introduction into Europe, the Roman method of representing numbers was employed. This still finds limited use, on the faces of clocks, on early pages of some books, on certain monuments and routinely in giving the dates of films.

Although the Roman system is more cumbersome than the familiar one, it is quite logical and may even be used, without excessive trouble, for moderately difficult arithmetic computations. Its major drawback is that, unless we continue to invent new symbols *ad infinitum*, we reach a limit beyond which numbers cannot be usefully represented. (The Romans *did* in fact have a systematic way to invent such new symbols, but it becomes very cumbersome very quickly.) With our number symbols, this is not such a problem, but the English language (in common with almost all others) runs out of *words* once the numbers become large enough.

There are actually many different conventions that have from time to time been used in the Roman method of number representation. They differ particularly over the representation of larger numbers. For more details see *Number Words and Number Symbols* by Karl Menninger (MIT Press, 1969), which also contains much else of interest. The account given here gives a slightly modernised version of the simplest, most logical and most widespread system.

It will be useful to bear in mind as you read this article that *the Roman system is a decimal system like our own* – that is to say, it gives a special place to the number ten and its powers. This principle is used to set up the *elements* from which other number representations are constructed.

Every Roman numeral is a *string* of letters written one after another and read from left to right. The letters occurring in the string are called the *elements*. There are five *basic elements*, which are:

I	(meaning "one")	X	(meaning "ten")
C	(meaning "a hundred")	M	(meaning "a thousand")
K	(replacing a more cumbersome Roman symbol, and meaning "ten thousand").		

In addition, there are four *auxiliary elements*, namely:

V	(meaning "five")	L	(meaning "fifty")
D	(meaning "five hundred")	W	(meaning "five thousand").

(W, like K, here replaces a more cumbersome Roman symbol.)

Strings are constructed from these nine elements in accordance with certain rules. These rules of themselves tell us nothing about a string's *meaning*, but only *whether* it has a meaning at all. Although I will make use of the meanings in my account of the rules, this is not, strictly speaking, necessary. However, it does simplify matters. If all of the rules are satisfied, then the string has a meaning and may be *interpreted*; the rules may be thought of as a *grammar* or *syntax* of a small and specialised language, whose words are the elements.

First, we need some notation. I will represent elements by lower case letters, a, b, c, etc., and their *meanings* by the notation $n(a), n(b), n(c)$, etc. Thus if $a = \mathrm{V}$, we mean that a stands for the element V. In this case, $n(a) = n(\mathrm{V}) = 5$. Later on, we will also use the symbol $n()$ to apply to strings, and will find that, for example, $n(\mathrm{MDCCCXLIX})$ is the number 1849.

Now for our rules. We begin with:

Rule One: If a and b are elements such that $n(a) < n(b)$, then a may precede b (i.e. be written to its left) in a string only if a *immediately* precedes b.

Thus we may write IV or IX, but not IXV (as the I does not immediately precede the V). Nor may we write XCM (as the X does not immediately precede the M), nor XXL (as the first X does not immediately precede the L). Combinations like VI or CX are not covered by the rule, and we shall see shortly that these two are permissible.

Rule Two: If a is an auxiliary element, and b is any element, then a may precede b only if $n(a) > n(b)$.

This rule implies that each auxiliary element occurs *at most once* in each string and that the auxiliary elements, if present, must be arranged from left to right in the order W, D, L, V. Thus LX is permissible, as is LIX; VX or WK are not allowed.

These two rules are the main ones, but three somewhat less important ones are usually employed as well.

Rule Three: If a and b are elements, and $n(a) < n(b)$, then the combination aba is disallowed.

Thus CXC is permissible; XLX is not.

Rule Four: If a and b are elements, then a may not precede b if $n(b)/n(a) > 10$.

This rule allows XC, XL, XX, XV and XI, but not XK, XW, XM or XD. IC (allowed in some versions) is also excluded. Again note that by Rule Two, we only encounter a further restriction if a is a basic element.

Rule Five: If a is an element, the combination $aaaa$ is disallowed.

Once again, this only further restricts us when a is a basic element. It outlaws strings such as XXXXX and IIII. (Some variants allow the latter; we will not.)

With these rules behind us, we now have access to a *Method* (or *Canon*) *of Interpretation* which allows us to translate Roman into Arabic numerals and *vice versa*. The main rules allowing this are rules One and Two; the others merely make for a neater, tidier system and ensure that for every number there is *precisely one* string that represents it.

An element a may be used in a string in either of two ways. To decide which is being employed, look at a's successor (the element to its immediate right, b let us call it); if $n(b) > n(a)$ then a is used in the *subtractive* mode, otherwise it is used in the *additive* mode.

Rule One ensures that *if an element is used in the subtractive mode, then its successor is used in the additive mode.*

Rule Two tells us that *only basic elements may be used in the subtractive mode.* Auxiliary elements may not.

Each element in a string may now be assigned unambiguously either to the additive or to the subtractive mode. As only the first two of our rules are here involved, we could interpret strings without use of rules Three, Four and Five; in this case, however, the string could be ungrammatical in the same sense as is the, comprehensible to be sure, English sentence "I go'd Saturday footy".

Our Canon of Interpretation is now:

Add the numerical equivalents of all elements used in the additive mode, and from the result subtract the sum of the numerical equivalents of all elements used in the subtractive mode.

An example will help to clarify this. In the string MDCCCXLIX, the first X and also the I are used in the subtractive mode. All the other elements are used in the additive mode. Then

$$\begin{aligned} n(\text{MDCCCXLIX}) &= [n(\text{M}) + n(\text{D}) + 3n(\text{C}) + n(\text{L}) + n(\text{X})] - [n(\text{X}) + n(\text{I})] \\ &= (1000 + 500 + 300 + 50 + 10) - (10 + 1) = 1860 - 11 \\ &= 1849. \end{aligned}$$

It is equally easy to go back the other way. As both systems are essentially decimal, numbers written in the Arabic characters may be converted to Roman form very easily. A single example will suffice to illustrate the technique. We choose, say, the number 24197. Then:

20000	is	2×10000	or	KK
4000	is	$5000 - 1000$	or	MW (M in subtractive mode)
100	is	100	or	C
90	is	$100-10$	or	XC (X in subtractive mode)
7	is	$5 + 2$	or	VII.

Hence our number is expressed by putting together all the expressions in the right-hand column: KKMWCXCVII. Notice how this process, with a little attention to detail, gives a string that automatically obeys the given rules of syntax.

We may summarise these results as a *Theorem*, which you may care to prove formally:

Every sufficiently small whole number corresponds uniquely to a permissible string of elements (i.e. to a Roman numeral) and vice versa.

The question is often raised as to whether the system of Roman numerals could be used in practical computation. In fact it is usually asserted (without foundation) that it cannot. Actually, Roman numerals may be used in basic arithmetic and may produce answers to questions of addition, subtraction, multiplication and division almost as easily as our techniques with Arabic numerals. One may even employ them to find square roots.

There have been several algorithms (i.e. computational procedures) suggested for doing arithmetic with Roman numerals. I have details of two published versions (and I have seen others, although I now no longer have the details of these). However, in 1975, a group of four American philosophers devised methods for addition and multiplication and published these in the learned journal *Archive for History of Exact Sciences, Volume* 15. In 1981, an engineer with interests in computational mathematics for use in the aerospace industry, and who had come up with very elegant methods for multiplication and division, published these in the *American Mathematical Monthly, Volume* 88. A few years ago, I myself devised ways to do all the four basic operations in Roman numerals, although I did not publish them and will not here go into the full details, which are probably not of great general interest. The basic outline is not all that different from the methods taught in primary school for use with our own familiar number representation.

The principal complication is the use of elements in the subtractive mode. Our Canon of Interpretation makes every string a difference of two strings (although the second string could be the *null string* with no elements, and thus representing zero). Thus if A and B are strings, then

$$A = A_1 - A_2 \text{ (say)} \quad \text{and} \quad B = B_1 - B_2 \text{ (say)}.$$

In every one of the strings A_1, A_2, B_1, B_2, each element is used in the additive mode. We now have, in the case of subtraction

$$\begin{aligned} A - B &= (A_1 - A_2) - (B_1 - B_2) \\ &= (A_1 + B_2) - (A_2 + B_1). \end{aligned}$$

Similarly in the case of multiplication

$$\begin{aligned} A \cdot B &= (A_1 - A_2) \cdot (B_1 - B_2) \\ &= A_1 \cdot B_1 + A_2 \cdot B_2 - A_1 \cdot B_2 - A_2 \cdot B_1 \\ &= (A_1 \cdot B_1 + A_2 \cdot B_2) - (A_1 \cdot B_2 + A_2 \cdot B_1). \end{aligned}$$

Because each of the products here involves only elements used in the additive mode, we may thus proceed, using simpler methods than if subtractive elements were still present. (And note, incidentally, how simple the Roman multiplication table is!)

Although the methods mentioned here have been devised, it is far from certain that the Romans themselves used such algorithms. There must have been Romans sufficiently smart to devise such algorithms, but no written record of computation with Roman numerals has come down to us. This is in contrast with the Greeks, whose methods of number representation were different again, and even less familiar to us. In *their* case there are surviving records of techniques for (e.g.) long division involving fractions. Even here it has been speculated that to simplify matters, they may have used the *abacus*: a device that enabled calculations to be done mechanically by moving counting beads about a counting frame. (Such devices are still in widespread use in China, even in tourist hotels.)

Here we are on surer ground. The abacus *was* used by the Romans. Two examples survive to this day; one is in a museum in Rome, the other in Paris. Both are small hand-held devices, about the size of a present-day pocket calculator. Photographs of these are to be seen in Menninger's book (referred to above). The abacus was in fact a most versatile device; it could represent decimal fractions and also numbers up to a million and beyond.

Because we have this physical evidence of mechanical calculation and no written records of symbolic computation, Menninger's verdict is that the Romans, not unlike many students of today, were hooked on their calculators.

In closing, I leave you with a few problems.

1. What is the largest number that can be represented by a permissible string? I.e., how small is "sufficiently small" in the *Theorem*?

2. Write computer programs to decide if strings are permissible and to convert Roman into Arabic numerals and *vice versa*.

3. What is the longest permissible string? What fraction of strings of this length or less are permissible?

4. See if you can devise algorithms for addition, subtraction, multiplication and division using Roman numerals.

* * * * *

COMPUTERS AND COMPUTING

A Folding Program

Cristina Varsavsky

In this article we will design a computer program to obtain the curve known as the *dragon curve*, which was introduced by the NASA physicist John Heighway. There are several methods for generating the dragon curve, but for the purpose of designing an algorithm to draw it on the screen, perhaps the most useful one is by paper folding. Take a long horizontal strip of paper and fold it in half, left over right, marking a sharp crease. Repeat this again, folding in the same direction. Now you have three creases. Fold it two more times, always in the same direction. When the strip is unfolded, an interesting pattern of creases appears. Looking at the folded strip from the side, we see a polygonal curve with 15 vertices as shown in Figure 1.

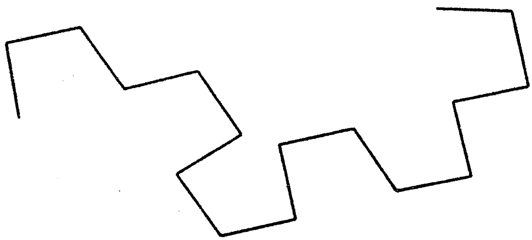


Figure 1

If the strip is arranged so that we have 90° angles at the vertices, we obtain what is called the *dragon curve of order 4* which is depicted in Figure 2. In general, the result of folding a paper strip n times and arranging it using angles of 90° at the vertices is called the *dragon curve of order n*. The strip may become unmanageable after a few steps, so we probably cannot obtain, by folding a strip of paper, a dragon curve of order greater than six or seven. As the generation of the dragon curve is an iterative exercise, what better tool than a computer to do the folding for us to explore dragon curves of higher orders?

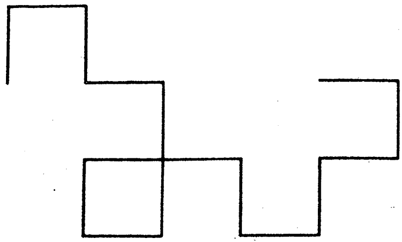


Figure 2

We shall construct the dragon curve starting at one end and then drawing segments of constant length, turning right or left following the pattern of creases we have in Figure 2. Therefore we need to discover a rule that determines in what direction we should draw each segment.

Figure 3

By opening the paper strip as in Figure 3, we see a sequence of upward and downward creases which we will label as U and D respectively. A table is set up below, with n indicating the position of each crease counting from the left, and the corresponding *direction* which is either U(pwards) or D(ownwards):

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
direction	U	U	D	U	U	D	D	U	U	U	D	D	U	D	D

We observe that $direction(1) = direction(5) = direction(9) = direction(13) =$ U. Also, $direction(3) = direction(7) = direction(11) = direction(15) =$ D. Therefore we will have an upward crease if the remainder after n is divided by 4 is 1, and a downward crease if the remainder is 3. We also observe that if n is even then

$$direction(n) = direction(n/2)$$

You can check these rules with any possible folding order.

Now we have all we need to write a program that draws a dragon curve of any order. We will do it in QuickBasic as usual, but you can easily adapt it to any other programming language. The program follows:

```
REM Dragon curve with angle of 90 degrees at the vertices

SCREEN 9
WINDOW (-2, -2) - (3, 2)
pi = 3.141593
angle=pi/2 : m = 0

PRINT "Order?": INPUT order
length = 1/(2^order)

x = length: y = 0:
LINE (0,0) - (x,y)

FOR i = 1 TO 2^order - 1

    p = i

    WHILE (p MOD 2) = 0
        p = p / 2
    WEND

    IF p MOD 4 = 1 THEN bend = 1 ELSE bend = -1

    m = m + bend MOD 4
    x = x + length*COS((m - order/2)*angle)
    y = y + length*SIN((m - order/2)*angle)

    LINE - (x,y)

NEXT i
```

As the program is run with order set to 10, the curve depicted in Figure 4 appears on the screen. You can see that the curve resembles a Chinese dragon, which explains its name. The program draws dragon curves with horizontal and vertical lines only for even orders. Can you explain what happens with odd orders?

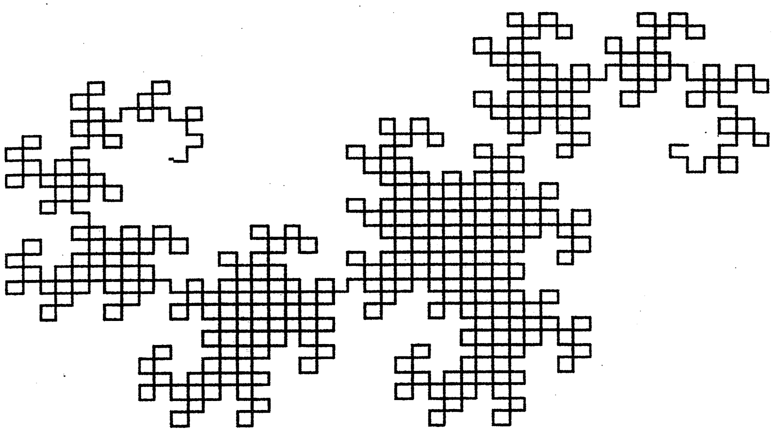


Figure 4

As the direction in which each segment is drawn is specified in terms of the angle we want to have at each bend, you can easily experiment with dragon curves with any angle at the bends. You only need to be careful to express the angles in radians, rather than degrees.

John Heighway's original dragon curve had the corners rounded off, as in Figure 5. You can modify the program to replace each segment with three segments to round off the right angles at the bends.

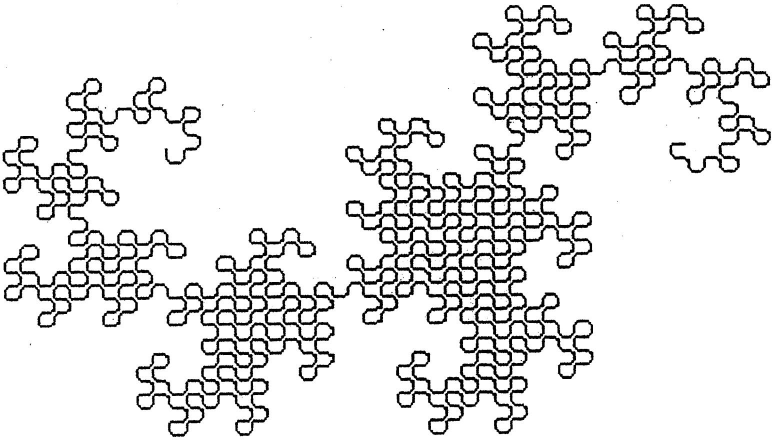


Figure 5

Figure 5 shows one of the many properties of the dragon curve: it never crosses itself. If you would like to find out more about the family of dragon curves, there are two interesting articles by C Davis and D Knuth ([1], [2]) which analyse the properties and number representation of these curves. The articles also explore curves obtained using different folding patterns, for example, folding the strip alternately right and left rather than always in the same direction as we did.

References

[1] Davis, C and Knuth, D E, 1970, Number Representations and Dragon Curves-I, *Journal of Recreational Mathematics, Vol 3, Number 2.*

[2] Davis, C and Knuth, D E, 1970, Number Representations and Dragon Curves-II, *Journal of Recreational Mathematics, Vol 3, Number 3.*

* * * * *

LETTER TO THE EDITOR

Dear Editor,

It is some time since I have had any communication from my erratic correspondent, the eccentric Welsh physicist and mathematician Dai Fwls ap Rhyll. Nor can I claim to have heard from him directly this year either. However, a mathematical derivation circulating on the Internet could not possibly be by anyone else, and so after my usual fashion I would like to share it with your readers.

He begins by noting that

$$\frac{1}{2} = \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \cdots$$

and then using each side of this equation as the exponent to which the number -1 is raised. Thus

$$\begin{aligned}(-1)^{1/2} &= (-1)^{1/3+1/9+1/27+\ldots} \\ &= (-1)^{1/3} \times (-1)^{1/9} \times (-1)^{1/27} \times \cdots \\ &= (-1) \times (-1) \times (-1) \times \cdots.\end{aligned}$$

This simple argument, whose leading equation you may readily check by summing the geometric series, thus shows that i (i.e. $\sqrt{(-1)}$) is the value of $(-1)^\infty$. The result has some quite remarkable consequences. Taking logs (to base -1) of both sides of the equation $(-1)^{1/2} = (-1)^\infty$, we find $\infty = 1/2$. Moreover, if we square both sides of the same equation, we find that $(-1) = (-1)^\infty = (-1)^{1/2}$, so that $i = -1$, and squaring both sides of *this* equation we find $+1 = -1$. And so we may continue finding many new, surprising and truly wonderful theorems! I leave their discovery to your readers.

Kim Dean
Union College
Windsor

* * * * *

PROBLEM CORNER

SOLUTIONS

PROBLEM 18.5.1

In Figure 1, the point P is free to move along $\overline{OL}$.

(a) What is the maximum value of θ (θ_{max})?

(b) With AB remaining at 1 unit, what is the length of $\overline{OA}$ so that $\theta_{\text{max}} = 30°$?

(The problem can be solved without calculus.)

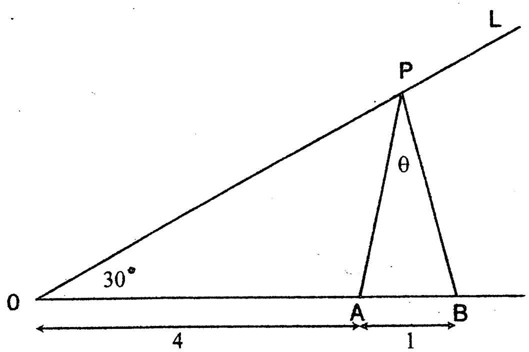


Figure 1

SOLUTION

Construct the circle passing through A and B with $\overline{OL}$ as a tangent. We claim that if P is chosen to be the point of tangency then $\theta = \theta_{\text{max}}$. In order to see this, let P' be another point on $\overline{OL}$, and let Q be the point where the bisector of the angle $AP'B$ cuts the circle. Then $\angle AQB = \theta$, and clearly $\angle AP'B < \theta$.

Let C be the centre of the circle, and let D be the midpoint of $\overline{AB}$. Let r be the radius of the circle, let $h = OC$, and let $\phi = \angle POC$. By a theorem of elementary geometry, $\angle ACB = 2\theta$, so $\angle ACD = \theta$. (See Figure 2.)

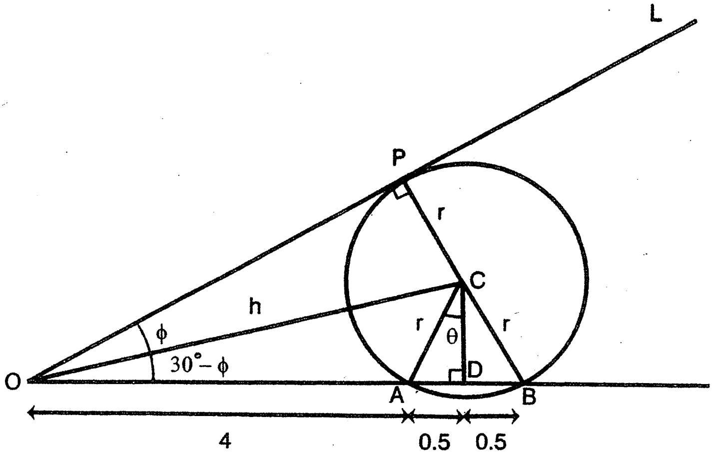


Figure 2

From the right triangles $\triangle ACD$ and $\triangle OCD$ we obtain respectively $CD^2 = r^2 - 0.5^2$ and $CD^2 = h^2 - 4.5^2$. Equating the two expressions for CD^2 and simplifying yields the equation:

$$r^2 = h^2 - 20 \tag{1}$$

From $\triangle OCD$ we obtain:

$$\cos(30° - \phi) = \frac{9}{2h}$$

After expanding $\cos(30° - \phi)$ we have

$$\frac{\sqrt{3}}{2}\cos\phi + \frac{1}{2}\sin\phi = \frac{9}{2h} \tag{2}$$

From $\triangle OPC$ we obtain:

$$\sin\phi = \frac{r}{h}$$

$$\cos\phi = \frac{\sqrt{h^2 - r^2}}{h} = \frac{\sqrt{20}}{h}$$

where the last equality follows from (1). Substituting these results into Equation (2) and solving for r (and noting that h disappears from the equation), we obtain:

$$r = 9 - \sqrt{60}$$

Finally, from $\triangle ACD$ we can calculate the value of θ as follows:

$$\theta = \arcsin\frac{1}{2r} = \arcsin\frac{1}{2(9-\sqrt{60})} = \arcsin\frac{9+2\sqrt{15}}{42} \simeq 23.5^\circ.$$

Part (b) of the problem can be solved by setting $OA = x$ and $\theta = 30^\circ$, repeating the steps shown above, and then solving for x. The answer is $x = 3$.

PROBLEM 18.5.2 (Peter Oliphant, student, Monash University)

The numbers in the equilateral triangle shown in Figure 3 represent the areas of their respective regions. Find the area of the central triangle.

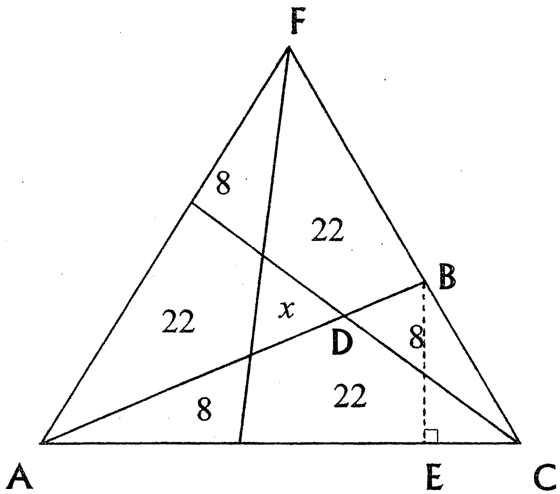


Figure 3

SOLUTION by Peter Oliphant

Figure 3 shows the triangle as it was given in the problem, but with some points labelled, and with the addition of the dashed line $\overline{BE}$ perpendicular to $\overline{AC}$.

The triangles $\triangle ABC$ and $\triangle BCD$ are similar, with $\overline{BC}$ on $\triangle BCD$ corresponding to $\overline{AB}$ on $\triangle ABC$. For similar triangles, the squares of corresponding lengths are proportional to the areas of the triangles. Therefore we obtain:

$$\frac{AB^2}{BC^2} = \frac{\text{Area}(\triangle ABC)}{\text{Area}(\triangle BCD)} = \frac{38}{8} = \frac{19}{4}$$

For the right triangle $\triangle BCE$ for which $\angle BCE = 60°$, the lengths of the other two sides can be established in terms of BC:

$$BE = \frac{\sqrt{3}}{2}BC \text{ and } EC = \frac{1}{2}BC$$

The right triangle $\triangle ABE$ can now be solved using Pythagoras to find AE in terms of BC:

$$AE = \sqrt{AB^2 - BE^2} = \sqrt{\frac{19}{4}BC^2 - \left(\frac{\sqrt{3}}{2}BC\right)^2} = 2BC$$

Thus the length of AC is:

$$AC = AE + EC = 2BC + \frac{1}{2}BC = \frac{5}{2}BC$$

Since $\triangle ACF$ is equilateral, $CF = \frac{5}{2}BC$ also.

The triangles $\triangle ACF$ and $\triangle ABC$ with bases $\overline{CF}$ and $\overline{BC}$ respectively have the same altitude, so their areas are proportional to the lengths of their bases. Therefore:

$$\frac{\text{Area}(\triangle ACF)}{\text{Area}(\triangle ABC)} = \frac{CF}{BC}$$

i.e.

$$\frac{90 + x}{38} = \frac{5}{2}.$$

Solving this equation for x gives $x = 5$.

Solution to an earlier problem

Our current editorial policy is to include in Problem Corner only problems for which we have a solution, and to publish the solution in the following issue but one. It has not always been so, however, and there are problems from earlier issues for which a solution has never appeared. Over the next few issues, as space permits, we will publish some of these solutions.

The problem below appeared in the June 1991 issue of *Function*.

PROBLEM 15.3.3

Inscribe a rectangle of maximal area into the ellipse given by

$$x^2 + 2y^2 = 18.$$

The four vertices of this rectangle lie on the hyperbola with the equation $x^2 - y^2 = a^2$.

(a) Determine these four vertices, the area of the rectangle and the value of a.

(b) Show that the foci of the ellipse and the hyperbola coincide. Show that the two curves intersect in right angles.

SOLUTION

(a) The equations $x^2 + 2y^2 = 18$ and $x^2 - y^2 = a^2$ are satisfied simultaneously at the four vertices. Solving for x and y, we obtain

$$(x, y) = \left(\pm\sqrt{\frac{2}{3}a^2 + 6},\ \pm\sqrt{6 - \frac{1}{3}a^2}\right)$$

as the coordinates of the vertices. The area A of the rectangle can now be expressed as a function of a :

$$A = 4\sqrt{(\frac{2}{3}a^2 + 6)(6 - \frac{1}{3}a^2)}$$

For maximal area, $\frac{dA}{da} = 0$; equivalently, $\frac{d(A^2)}{da} = 0$, which is a little easier to handle. Upon solving this equation (the calculations are routine, so we omit them), we obtain $a = 0$ or $a = \pm 3/\sqrt{2}$. Only the positive solution $a = 3/\sqrt{2}$ is meaningful in this context, so the vertices are $(\pm 3, \pm 3/\sqrt{2})$. The area of the rectangle is $A = 36/\sqrt{2}$.

(b) Using the standard formulae pertaining to conic sections, we obtain $(\pm 3, 0)$ as the foci of both the ellipse and the hyperbola.

To show that the ellipse and the hyperbola intersect in right angles, we find $\frac{dy}{dx}$ for each equation by differentiating implicitly with respect to x:

$$2x + 4y\frac{dy}{dx} = 0 \quad \text{(ellipse)}$$

$$2x - 2y\frac{dy}{dx} = 0 \quad \text{(hyperbola)}$$

(If you are not familiar with implicit differentiation, you could solve each equation for y first and then differentiate, but that requires more work.)

Now solve each equation for $\frac{dy}{dx}$. For the ellipse we have:

$$\frac{dy}{dx} = \frac{x}{2y} \tag{1}$$

and for the hyperbola:

$$\frac{dy}{dx} = \frac{x}{y}. \qquad (2)$$

Upon substituting in turn the coordinates of each of the four points of intersection into these two equations, we find that the product of the two derivatives is -1. Thus the curves intersect in right angles.

PROBLEMS

PROBLEM 19.2.1 (from *School Science and Mathematics*; submitted by M Deakin)

Show that $a^{a^{1/e}} \geq 1/e$ for $a > 0$.

PROBLEM 19.2.2 (from *Parabola, Vol 30, Part 3*, 1994, University of NSW)

Four weary explorers have to cross a bridge over a river one night. Owing to their various degrees of exhaustion, they would individually take 5, 10, 20 and 25 minutes (respectively) to cross the bridge. However, the old and rickety bridge will take only one or two people at a time. Furthermore, it is too dangerous to cross the bridge in the dark, and the expedition has only one torch. How can all four explorers cross the bridge in the least possible total time?

PROBLEM 19.2.3

A 6×6 array of squares is completely covered with 18 dominoes, in such a way that each domino covers two adjacent squares. Prove that there must be at least one line, either horizontal or vertical, that divides the array into two parts without passing through any of the dominoes.

PROBLEM 19.2.4 (from IX Mathematics Olympiad "Thales", in *Epsilon*, 26, 1993, p 107)

The brothers Al Caparroni are trying to open a safe in the Peseta Bank. The combination is composed of an increasing sequence of three non-zero digits. In the pocket of the cashier they have found the following information:

- The sum of the digits is 17.
- The product of any two digits added to the third digit is a perfect square.

What is the correct combination?

THE 1995 AUSTRALIAN MATHEMATICAL OLYMPIAD

Hans Lausch

The contest was held in Australian schools on February 7 and 8. On either day students had to sit a paper consisting of four problems, for which they were given four hours.

Paper 1

1. Show that there are not more than 9 prime numbers between 10 and 10^{29} whose representation in base 10 is a string of ones (i.e. numbers like 11 or 1111).

2. On the circumference of a circle, $4n$ points (n a positive integer) have been chosen and numbered consecutively $1, 2, 3, \ldots, 4n$ clockwise round the circle. All the $2n$ even-numbered points are divided into n pairs, and the points of each pair are joined by a green chord. Similarly, all the $2n$ odd-numbered points are divided into n pairs, and the points of each pair joined by a gold chord. The choice of points and chords turns out to be such that no three of the chords (same colour, or a mixture of both) meet at a common point. Prove that there are at least n points where a green chord intersects a gold chord.

3. A straight line cuts two concentric circles in the points A, B, C and D in that order; AE and BF are parallel chords, one in each circle; GC is perpendicular to BF at G, and DH is perpendicular to AE at H. Prove that $GF = HE$.

4. Determine all pairs (x, y) of positive integers which satisfy

$$y^2(x-1) = x^5 - 1.$$

Paper 2

5. Determine all real numbers r such that there is precisely one pair (x, y) of real numbers satisfying the conditions

   (i) $y - x = r$,

   (ii) $x^2 + y^2 + 2x \leq 1$.

6. For each positive integer n let $f(n,0), f(n,1), \ldots, f(n,2n)$ be integers defined as follows:

   the polynomial $f(n,0) + f(n,1)x + f(n,2)x^2 + \ldots + f(n,2n)x^{2n}$ is identical with the polynomial $(x^2 + x + 1)^n$.

   Prove that there are infinitely many values of n for which precisely three of the integers $f(n,k), k = 0, 1, \ldots, 2n$, are odd.

7. The lines joining the three vertices of triangle ABC to a point in its plane cut the sides opposite vertices A, B, C in the points K, L, M respectively. A line through M parallel to KL cuts BC at V and AK at W. Prove that $VM = MW$.

8. Let $a > 0$ be real. Determine all real-valued functions f defined on the set of all positive reals such that

   (i) $f(xy) = f(x)f(\frac{a}{y}) + f(y)f(\frac{a}{x})$ for all positive reals x, y;

   (ii) $f(1) = \frac{1}{2}$.

There were 102 entrants. Gold Certificates were received by the Year 12 students:

Christopher Barber	WA	Carine SHS
Richard Davis	Tas	The Hutchins School
Gordon Deane	Qld	Kenmore High School
Lisa Gotley	Qld	All Saints Anglican School
Jian He	Vic	University High School
Ben Lin	NSW	North Sydney Boys High School
Andrew Over	Qld	Anglican Church Grammar School
Martin Sheppard	SA	Marbury School
Nigel Tao	SA	Westminster School
Trevor Tao	SA	Brighton Secondary
Zhi Ren Xu	NSW	Canterbury Boys' High School

Congratulations to all! Twenty-seven students, including all Gold Certificate winners, were invited to represent Australia at the Seventh Asian Pacific Mathematics Olympiad. Students from fifteen countries of the Asia-Pacific Region take part in this competition which was started in 1989. The competition paper will also be used by Argentina and by Trinidad and Tobago.

* * * * *

Published by Monash University Mathematics Department