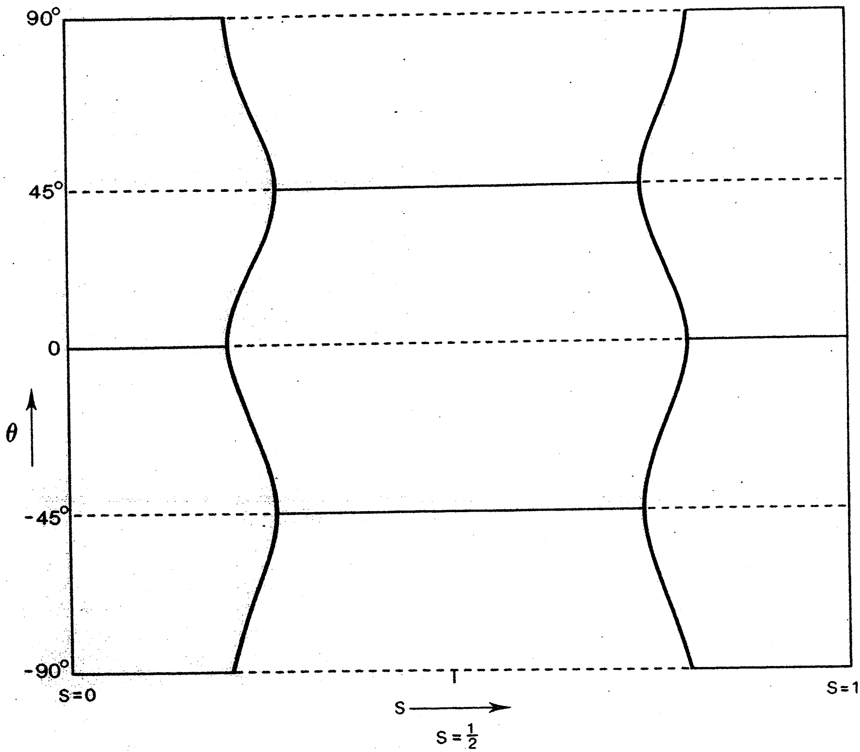
FUNCTION

Volume 11 Part 4 August 1987



A SCHOOL MATHEMATICS MAGAZINE
Published by Monash University

Reg. by Aust Post Publ. No. VBH0171

Function is a mathematics magazine addressed principally to students in the upper forms of schools, and published by Monash University.

It is a 'special interest' journal for those who are interested in mathematics. Windsurfers, chess-players and gardeners all have magazines that cater to their interests. *Function* is a counterpart of these.

Coverage is wide - pure mathematics, statistics, computer science and applications of mathematics are all included. There are articles on recent advances in mathematics, news items on mathematics and its applications, special interest matters, such as computer chess, problems and solutions, discussions, cover diagrams, even cartoons.

* * * * *

BUSINESS MANAGER: Mary Beal (03) 565-4445

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,
Clayton, Victoria, 3168.

Alternatively correspondence may be addressed individually to any of the editors at the institutions shown above.

The magazine is published five times a year, appearing in February, April, June, August, October. Price for five issues (including postage): $10.00*; single issues $2.00. Payments should be sent to the business manager at the above address: cheques and money orders should be made payable to Monash University. Enquiries about advertising should be directed to the business manager.

* $5.00 for *bona fide* secondary or tertiary students.

* * * * *

Registered for posting as a periodical - "Category B"

ISSN 0313 - 6825

CONTENTS

* * * * *

THE FRONT COVER

Michael A.B. Deakin, Monash University

Imagine a log with a square cross-section and set it afloat in a pool of water. What way will it line up? Will it adopt a horizontal position as in Figure A or Figure H opposite? Or perhaps it might lie with its sides at 45° to the water surface as in Figure D or Figure E? Or perhaps at some other angle as in Figures B, C, F or G?

In fact all these behaviours are possible. The angle (θ, let us call it) between the side and the water surface is determined by the specific gravity of the log. That is to say by the ratio of the densities : density of wood/density of water. Let us call this quantity s .

The basic law regarding floating bodies is called Archimedes' law, and legend has it that he discovered it while in his bath, being so carried away that he leapt out, crying "Eureka!" (Greek for "I have found it"). Archimedes' law states that the upward force on a submerged body is equal to the weight of the fluid it has displaced. If the body is floating, the upward (or buoyancy) force so generated must exactly equal the weight of the body.

In our example, the weight of the water displaced will be proportional to the wetted area W, say, while the weight of the body will be proportional to the total area T, say . In fact we must have

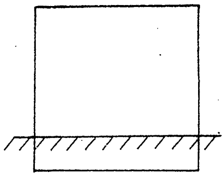
$$W = s\,T \tag{1}$$

and so we can achieve flotation if $0 < s < 1$.

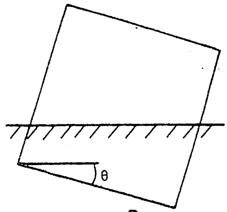
The problem of determining how the log will float, which of the diagrams A – H will be achieved, boils down to determining how θ depends on s .

To do this, we use a form of Archimedes' law, but in a restatement. The geometric centre (equivalent to the centre of mass) of the wetted area is the point through which the upward, buoyancy, force acts. It is called the *centre of buoyancy* . The downward, weight, force acts, of course, through the geometric centre of the square.

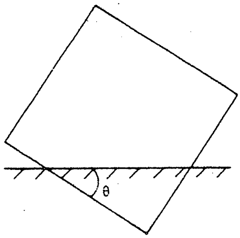
The log will float without rotating if the second of these two points is directly above the first. It will be stable (i.e. tend to right itself if slightly disturbed) if the distance between the two points is minimised.



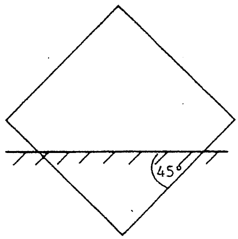
A



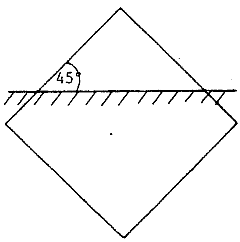
B



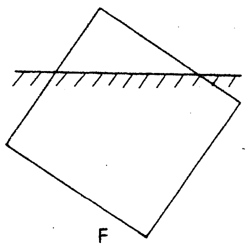
C



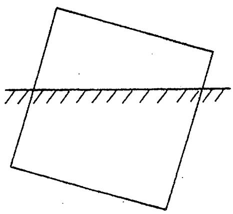
D



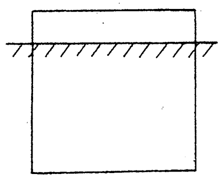
E



F



G



H

This much can be proved from Archimedes' law and our more modern understanding of mechanics. The specific case of the floating log was first worked out by the Dutch physicist and mathematician, Christiaan Huygens (1629–1695), a somewhat senior contemporary of the better-known Isaac Newton (1642–1727).

The cover diagram (reproduced with minor differences of labelling on p.101) shows the results of the analysis. I spare readers the details of the calculations, which can be consulted in Sir Horace Lamb's book *Statics* (Cambridge University Press, 1912), pp.220–234. Here I merely summarise the results.

If

$$0 < s < \frac{1}{2} - \frac{\sqrt{3}}{6} \quad (= 0.2113\ .\ .\ .) \qquad (2)$$

then the log will float in the attitude shown in Figure *A*. Balsawood and corkwood logs adopt this configuration.

But if

$$\frac{1}{2} - \frac{\sqrt{3}}{6} < s < \frac{1}{4} \quad , \qquad (3)$$

then the log will float in the attitude shown in Figure *B*. The precise connexion between θ, s is given by the equation

$$\tan^2 \theta = 12s(1 - s) - 2 \qquad (4)$$

for this range of values of s.

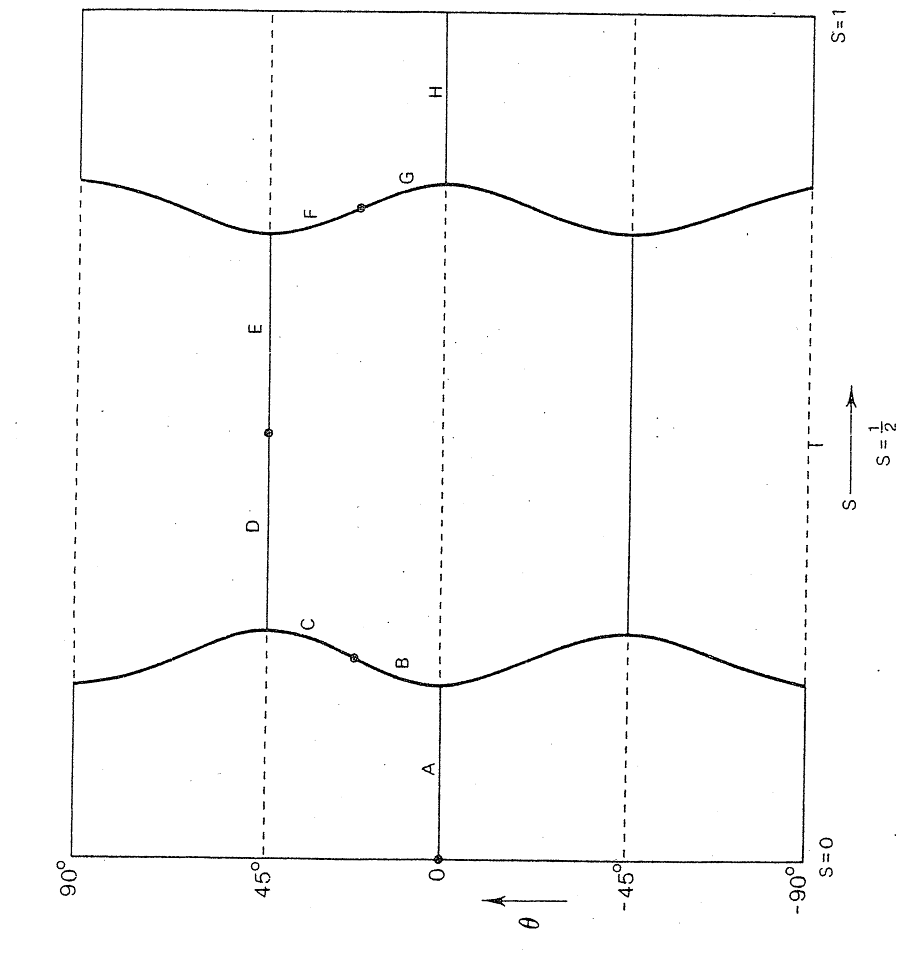
I wasn't able to find a wood whose density lies in the range given by Inequality (3), but presumably some plastic or composite material satisfying these constraints must exist.

If $s > \frac{1}{4}$, and

$$\frac{1}{4} < s < \frac{9}{32} \quad , \qquad (5)$$

then the log floats as shown in Figure *C*. The formula connecting θ, s is in this case:

$$\frac{8s}{9} = \frac{\tan \theta}{(1 + \tan\theta)^2} \quad . \qquad (6)$$


90°
45°
0
-45°
-90°
θ
S=0
S=1/2
S=1
S
A
B
C
D
E
F
G
H
I

Again, I couldn't learn of a wood with the appropriate density, but there must be a plastic or composite . . . .

Once we achieve a higher specific gravity

$$\frac{9}{32} < s < \frac{1}{2}, \tag{7}$$

the log floats with its diagonal upright, as in Figure *D*, and the diagonal remains upright if

$$\frac{1}{2} < s < \frac{23}{32}, \tag{8}$$

the log merely settling lower in the water. Figure *E* shows this behaviour.

Many softwoods satisfy the constraints of Inequality (7). Western Red Cedar is one example. Black Ash and many other timbers have densities that obey Inequality (8).

Now consider higher densities. If

$$\frac{23}{32} < s < \frac{3}{4} \tag{9}$$

the configuration adopted is that of Figure *F*. Oak floats like this.

Next if

$$\frac{3}{4} < s < \frac{1}{2} + \frac{\sqrt{3}}{6} \ (= 0.7886\ .\ .\ .)\ , \tag{10}$$

then Figure *G* gives the situation. Applewood is a reasonable accessible example.

Finally if

$$\frac{1}{2} + \frac{\sqrt{3}}{6} < s < 1\ , \tag{11}$$

the log will float in the position shown in Figure *H*. Heavy timbers like Dogwood or Jarrah wallow low in the water in this fashion. One of the heaviest timbers that still (just) stays afloat is Andaman Marblewood (a form of ebony), for which $s = 0.978$. Of course, some timbers sink. Australian Mahogany Gum is one of these.

The cover diagram (and its variant on p.101) summarise the data given above. Solid lines graph the angle θ as a function of the specific gravity s . The letters attached to the graph on p.101 show the different behaviours and relate the graph to Figures *A–H*.

Note several points.

1st : as we go into Figure *B*, we could tilt either to the right, as shown, or to the left. θ , in other words, could be positive or negative. For θ negative, we could trace a path through the lower half of the diagram on p.101.

2nd : we could begin, not with $\theta = 0$, but with $\theta = \pm 90^\circ$, and these possibilities are also shown in the diagram and we see how they connect up with the cases previously discussed.

3rd : we could even have $\theta = 180^\circ$ or $\theta = -180^\circ$, which would represent the same configuration. I haven't drawn this because the identification of 180° and -180° would require a cylindrical page, which would make *Function* hard to print and to read. You'll just have to use your imagination.

Notice too the dashed lines. These represent equilibria, but these equilibria are unstable. (The distance between the centre and the centre of buoyancy is *maximised* in such cases.) Notice how stable equilibria become unstable at just those points where different curves intersect. This is a phenomenon called "the exchange of stabilities" –it first drew explicit attention from the French mathematician Henri Poincaré (1854–1912) and today it is a very important topic indeed.

Huygens in fact analysed not only the system we have described, but a more general one, in which the cross-section of the log is rectangular. θ in this case depends on two quantities, s (as before) and the shape of the rectangle. You may like to consider how this additional complication affects the diagrams so far drawn.

However, Huygens didn't complete the analysis of the more general case. That was done by another Dutch physicist and mathematician, Diederik Korteweg, who is today best remembered for his work in Hydrodynamics, but who was also a notable historian and editor of many of the volumes of Huygens' collected works.

Lamb tells us that Huygens gave up his calculation, saying, in effect, that it had at most an extremely minute practical usefulness. Huygens' calculations were right, his historical judgement wrong. These calculations have been modified and extended to analyse the stability of offshore oilrigs under tow. And the "practical usefulness" of *those* calculations is measured today in millions of dollars.

CALCULATING THE GREATEST COMMON DIVISOR

R. T. Worley, Monash University

The greatest common divisor (or highest common factor, as it is sometimes called) of two integers a, b is the largest positive integer that divides both a and b. It can be found from the prime factorisations of a and b, but calculating prime factorisations of large integers is not easy. A better method is Euclid's algorithm for calculating the greatest common divisor, gcd(a_0, b_0), of two positive integers a_0, b_0 which is described by J.A. Deakin in an earlier article in *Function* (Vol 8, part 5). The algorithm is basically just repeatedly taking remainders on division, and is quite fast for hand calculation. For example, to find gcd(57, 152) we divide 152 by 57 to get 152 = 2 × 57 + 38. Now dividing 57 by 38 we get 57 = 1 × 38 + 19, then dividing 38 by 19 we get 38 = 2 × 19 + 0. At this point we stop, for gcd(0, 19) is 19.

In essence, the algorithm is just to replace the pair (a_0, b_0), that is (57, 152), by the pair (a_1, b_1), that is, (38, 57), then to replace this by (a_2, b_2) = (19, 38), and finally to replace this by (0, 19). Each new pair has the same gcd as the previous pair. Thus we say that the algorithm works by replacing a pair by a smaller pair which has the same gcd as the original pair. The procedure stops when the gcd is obvious. In Euclid's algorithm the rule for generating the next a, b pair is that b_{i+1} is a_i and a_{i+1} is the remainder on dividing b_i by a_i.

This algorithm for the gcd is quite fast. The number of division steps required is around 2.3 times the number of decimal digits in b, assuming that b is smaller than a. (More precisely, the number of steps is approximately $\log_e b$, the natural logarithm of b.) However the calculation of the remainder is not necessarily easy to do. For example in calculating gcd(1234567890123456789, 82345678901234578) we find the quotient is approximately $1.2345 \times 10^{18} / 8.2345 \times 10^{16} = 14.9918$ (so it must be 14), and the remainder is 1234567890123456789 − 14 × 82345678901234578. This can be done, with a little difficulty, using a calculator, pencil and paper.

There is another alternative algorithm for calculation of the gcd which is slightly more suited to hand calculation, and may also outperform Euclid's algorithm on some small computers. This alternative method is known as the binary method because it uses division and multiplication by 2. It performs well on small computers because they usually divide by 2 much more rapidly than they do the normal divide step required by Euclid's algorithm. The algorithm uses the same basic technique of replacing a_0, b_0 by smaller pairs with the same gcd, stopping when the gcd is obvious.

First, a simple example. Consider gcd(57, 133). Any common divisor of 57 and 133 divides $133 - 57 = 76 = 2 \times 38 = 2^2 \times 19$, and is odd, so it must divide 19. Replacing the larger of 57 and 133 by 19, we look for gcd(57, 19). Repeating this procedure, we write $57 - 19 = 38 = 2 \times 19$, so we replace the larger of 57 and 19 by 19, and look for gcd(19, 19). This is obvously 19, so gcd(57, 133) = 19. The process of dividing an even number by 2 until it becomes odd will be called extracting the "odd part" of a number. The power of two extracted will be called the "even part" of the number. Thus the procedure above is to repeatedly replace the larger of a, b by the odd part of $|b-a|$ until $a = b$.

The algorithm needs a little modification if we start with even numbers; its operation depends on the numbers both being odd. In general, the binary algorithm for finding gcd(a, b) is

(i) break a into its even and odd parts, call them e and x respectively.

(ii) break b into its even and odd parts, call them f and y respectively.

(iii) while $x \neq y$, replace the larger of x, y by the odd part of $|x - y|$.

(iv) gcd(a, b) is the final value of x, multiplied by the smaller of e and f.

For example, to find gcd(1802, 3332) we proceed

(i) $1802 = 2 \times 901$, so $e = 2$, $x = 901$.

(ii) $3332 = 2 \times 1666 = 2^2 \times 833$, so $f = 4$, $y = 833$.

(iii) $|901 - 833| = 68 = 2 \times 34 = 2^2 \times 17$, so $x = 17$, (y remains at 833).

$|17 - 833| = 816 = 2 \times 408 = 2^2 \times 204 = 2^3 \times 102 = 2^4 \times 51$, so $y = 51$.
(x remains at 17)

$|17 - 51| = 34 = 2 \times 17$, so $y = 17$ (x remains at 17)

(iv) gcd is 17 multiplied by the smaller of 2 and 4, that is, gcd(1802, 3332) = 34.

Whether or not the binary algorithm is faster than Euclid's algorithm on a computer depends on a lot of factors. The speed depends on the programming language being used, and even on the actual compiler or interpreter being used. In modern microcomputers the microprocessors being used tend to have special instructions both for obtaining remainders and for multiplication/division by 2. The existence of these special instructions means that both operations are done quite quickly. Previously there were not so many special instructions on small microprocessors, which meant that to obtain the remainders in Euclid's algorithm a lot of subtraction instructions had to be carried out. In this case Euclid's algorithm will probably be much slower. On my computer which does have special instructions for obtaining remainders Euclid's algorithm was a little faster for ordinary integers and the binary algorithm was faster for integers of 9 digits.

There is a lot of interest in factoring numbers quickly, in connection with cryptography, and a key process in factoring is finding gcds, so a fast algorithm for obtaining gcds is of great interest. Euclid's algorithm has another property, which makes it superior to the binary algorithm presented here. In handling very large numbers (say 200 decimal digits), it turns out that the full remaindering operation (which is very time consuming) does not have to be done quite as often. By the use of a modified algorithm many steps can be combined into one, making the algorithm run much faster. With the binary algorithm only a few steps can be combined. However, more recently, a variant of the binary algorithm has been proposed for a piece of computer hardware to gcd calculation - a so-called gcd-chip. Just as when the microprocessor has the special instructions to find remainders, Euclid's algorithm can be faster, so the binary algorithm can be made faster by adding a special chip to the computer.

UNRELIABLE WITNESSES AND BAYES' RULE

G. A. Watterson, Monash University

Recently, I read an article by Dennis Lindlay which interested me, and I would like to pass on his thoughts to you. Lindlay discussed a very old problem, dating back to 1685 if not earlier. The Bishop of Bath and Wells (in England) had raised the problem then, and had tried to answer it himself.

Suppose that two witnesses separately reported that a particular event happened. For instance, they may both have said that "a red-headed man stole a sheep from the farm." Both witnesses are known to be not completely reliable. Witness number 1 has probability p_1 of telling the truth and witness number 2 has probability p_2. What is the probability that the event actually took place?

The Bishop argued as follows. The probability that witness 1 doesn't tell the truth is $1-p_1$, and it is $1-p_2$ for witness 2. If the event *didn't* happen, then both witnesses must have lied, which, if they were acting independently, has probability $(1-p_1)(1-p_2)$. Therefore, the probability that they didn't both lie is $1-(1-p_1)(1-p_2)$, which must be the probability that the event actually took place. He was wrong!

In 1763, another clergyman, Thomas Bayes, introduced what we now call Bayes' rule. He was the rector of Tunbridge Wells (not "Bath and Wells", so Lindlay calls his article "A Tale of Two Wells"). How does Bayes' rule answer the problem? Let us introduce some further notation. Let B denote the event which may, or may not, have happened. We write B' to indicate the complement of B, namely that B didn't happen. Also, let A_1 denote the event that witness 1 said that B happened, and let A_2 denote the event that witness 2 said that B happened. Finally, let $A = A_1 \cap A_2$ indicate that *both* witnesses said that B happened. What the Bishop wanted to know was the probability of B, conditional on A, which we write as $Pr(B|A)$.

Bayes' rule helps us calculate such conditional probabilities. It says that

$$Pr(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B')P(B')} .$$

To see how to introduce the probabilities p_1 and p_2 that the witnesses tell the truth, we might assume that

$$p_1 = P(A_1|B) = P(A'_1|B')$$

and

$$p_2 = P(A_2|B) = P(A'_2|B') .$$

We are assuming that p_1 and p_2 apply whether B did occur or whether it did not; each witness will tell the truth, under either circumstance, with the same probability. Further, if we assume the witnesses acted independently, then

$$Pr(A|B) = Pr(A_1 \cap A_2|B) = Pr(A_1|B)Pr(A_2|B) = p_1 p_2 \tag{1}$$

and

$$\begin{aligned} Pr(A|B') &= Pr(A|B') \\ &= Pr(A_1 \cap A_2|B') \\ &= Pr(A_1|B')P(A_2|B') \\ &= [1-Pr(A'_1|B')][1-Pr(A'_2|B')] \\ &= (1-p_1)(1-p_2) . \end{aligned} \tag{2}$$

Substituting these results into Bayes' rule, we find the answer to the problem:

$$Pr(B|A) = \frac{p_1 p_2 \, Pr(B)}{p_1 p_2 Pr(B) + (1-p_1)(1-p_2)[1-pr(B)]} . \tag{3}$$

Remember that the Bishop said that the answer was

$$1 - (1-p_1)(1-p_2)$$

which is much simpler, but which is really $Pr(A'|B')$, the answer to a different problem!

How would we apply the correct result, (3), in practice? Notice that we need to have some idea of the probability $Pr(B)$ that the reported event would take place, *not* given any evidence from witnesses. The Bishop tried to get an answer to his problem without using $Pr(B)$, but unfortunately it does need to be known. Often, this is a great stumbling block to the use of Bayes' rule; in order to find the probability of B given that A happens, we need to know the probability of B in general, *without* knowing whether A or A' happens. For instance, in India, it might be very unlikely that a red-headed man would steal a sheep (there being very few red-headed men in India), so that even if our witnesses were generally fairly reliable (p_1 and p_2 close to 1), the low value of $Pr(B)$ will cause (3) to yield a low value of $Pr(B|A)$. For instance, if $Pr(B) = 1/1000$, $p_1 = p_2 = 0.9$, then (3) says that $Pr(B|A) = 0.075$, a very low probability in spite of two reliable witnesses saying that B happened.

The other complications about (3) are that you need to know the witnesses "reliabilities" p_1 and p_2 , and also to assume their independence as used in (1) and (2). Witnesses might collude to give false evidence if a defendant in court was actually guilty, but there would be no need for collusion if the defendant were innocent. So (1) might be false but (2) could be true.

* * * * *

Mathematics is the tool specially suited for dealing with abstract concepts of any kind and there is no limit to its power in this field. For this reason a book on the new physics, if not purely descriptive of experimental work, must be essentially mathematical. - P.A.M. Dirac (*Quantum Mechanics*, 1930).

As I proceeded with the study of Faraday, I perceived that his method of conceiving the phenomena [of electromagnetism] was also a mathematical one, although not exhibited in the conventional form of mathematical symbols. I also found that these methods were capable of being expressed in the ordinary mathematical forms, and thus compared with those of the professed mathematicians. - James Clerk Maxwell (*A Treatise on Electricity and Magnetism*, 1873).

SOLVING POLYNOMIAL EQUATIONS I

Neil S. Barnett
Department of Mathematics
and Operations Research
Footscray Institute of Technoloty

INTRODUCTION

Many practical applications of mathematics involve the solution of one or more equations. Quite often not all solutions to these equations are of concern but interest focuses on those solutions of a special type or that lie between particular values, because only in this range do the solutions relate to the physical situation from which the equations arose.

Many types of equations occur. There are polynomial equations, differential equations, integral equations, difference equations and numerous others. Quite often equations involving many variables have to be solved simultaneously and the equations can be linear, non-linear or homo-geneous each requiring their special methods of solution. Expressing systems of equations in matrix form is often an invaluable aid.

The first types of equations for solving, met at school, are simple linear equations followed by simultaneous linear (maybe even a linear and non linear equation in two variables), then quadratic equations and some simple equations involving trigonometrical ratios. Solutions to such equations (if they exist) are numeric.

An equation involving derivatives is called a differential equation and in general the solutions to such are mathematical functions rather than numbers. Consider, for example, solving the equation

$$3\frac{dy}{dx} = x^2 - 1 .$$

By integrating (anti-differentiating) both sides of the equation, y is seen to be

$$\frac{(x^3 - 3x)}{9} + C$$

where C can be any constant. (Differentiating the result will show this to be so.) Very often, some particular values of x and y satisfying this functional solution will be known and substitution of them into the solution will determine the constant C .

Differential equations are among the most common that the applied mathematician meets. They arise quite naturally in problems of mechanics, electronics, and in problems of oscillation, dynamics and hydrodynamics. For certain types of differential equations the method of solution requires determining the solutions of an auxiliary polynomial equation.

You may or may not have wondered, having derived the formula for solving quadratic equations, whether or not such formulae exist for the solution of cubic equations and similar equations of higher degree. It is the solution of such equations that is the substance of this article.

NEGATIVE SQUARE ROOTS AND GRAPHS

A polynomial equation of degree n , where n is a positive integer, is written as

$$a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_{n-1} x + a_n = 0 \,.$$

This yields a simple linear equation when $n = 1$, a quadratic when $n = 2$, etc.

For solving the quadratic

$$a_0 x^2 + a_1 x + a_2 = 0$$

we have the standard formula

$$x = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_o a_2}}{2a_o}$$

and when $a_1^2 < 4a_o a_2$ this involves the square root of a negative number, namely $\sqrt{a_1^2 - 4a_o a_2}$. At an elementary level, it is usually stated, that when this happens the equation simply has no solutions. An example is the equation

$$2x^2 - x + 12 = 0$$

and the quadratic formula gives

$$x = \frac{1 \pm \sqrt{1 - 96}}{4}$$

$$\text{i.e.} \quad x = \frac{1 + \sqrt{-95}}{4} \quad \text{or} \quad x = \frac{1 - \sqrt{-95}}{4} \quad .$$

Rather than discarding such solutions they can be written symbolically as,

$$x = \frac{1 \pm i\sqrt{95}}{4} \quad ,$$

where

$$i = \sqrt{-1} \; .$$

It is thus possible to expand the concept of a solution to include solutions that consist of a number together with another number multiplied by $\sqrt{-1}$. Such solutions are said to be complex. If there is no $\sqrt{-1}$ involved in the solution, then the solutions are said to be real. There is the need at times to find both real and complex solutions, at other times, very often for practical reasons, only one of the types of solutions is of interest. Whilst acknowledging that both complex and real solutions are possible this article focuses on finding real solutions.

As will be familiar from considering both linear and quadratic equations graphically, solutions can be visualised as the intersection of curves with the x axis. For example, solutions to the equation

$$x^2 - 4x + 3 = 0$$

can be obtained by observing where the curve $y = x^2 - 4x + 3$ cuts the x axis (i.e. at $x = 1$ and $x = 3$).

This is equally true for the general polynomial

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0$$

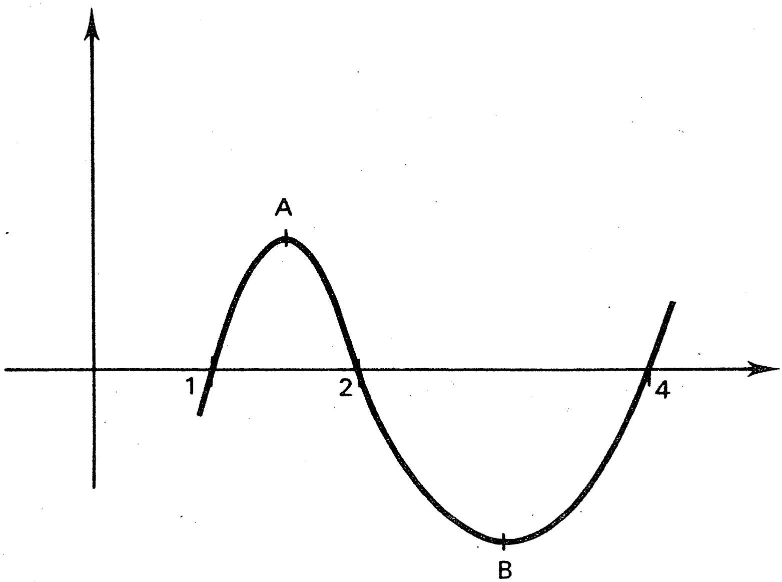
although of course the drawing of the graph of

$$y = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

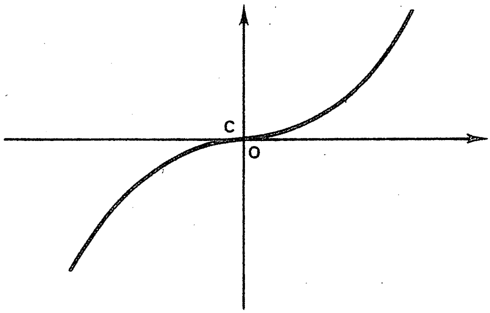
for large n, is a lot more involved than drawing the graph of a quadratic.

When trying to solve the quadratic equation $x^2 - 2x + 3 = 0$ using the formula the solution involves $\sqrt{-1}$: there are no real solutions. This is verified by drawing the graph of $y = x^2 - 2x + 3$ when we find that the curve does not intersect the x axis. The graphical approach is thus useful for approximately locating real solutions only (and these are often the ones of most importance to us). This is true also for cubics, quartics and higher degree polynomials.

Recall that in solving quadratic equations there is never any occasion when one solution is real and one complex : they are either both real or both complex. Consider now a cubic polynomial, for example $y = x^3 - 7x^2 + 14x - 8$. A plot of this looks like:



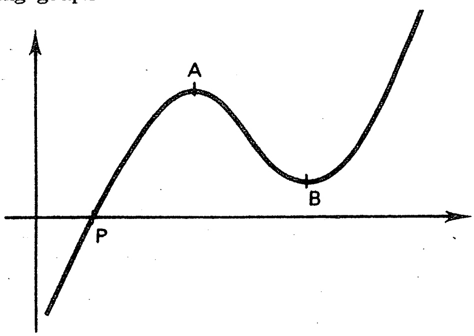
The solutions to the equation $x^3 - 7x^2 + 14x - 8 = 0$ are clearly 1, 2, and 4 (all real). There are also two turning points to the graph - labelled A and B. Another cubic, $y = x^3$ has a different shape entirely. It appears as



The two turning points A and B, of the previous example have merged into one at C.

Clearly, there is just one solution to $x^3 = 0$, namely $x = 0$. However, because $x^3 = 0$ is the equation and not just $x^1 = 0$, the solution $x = 0$ is said to occur three times!

Consider now a cubic $y = x^3 - 6x^2 + 11x - 1$ that gives rise to the following graph

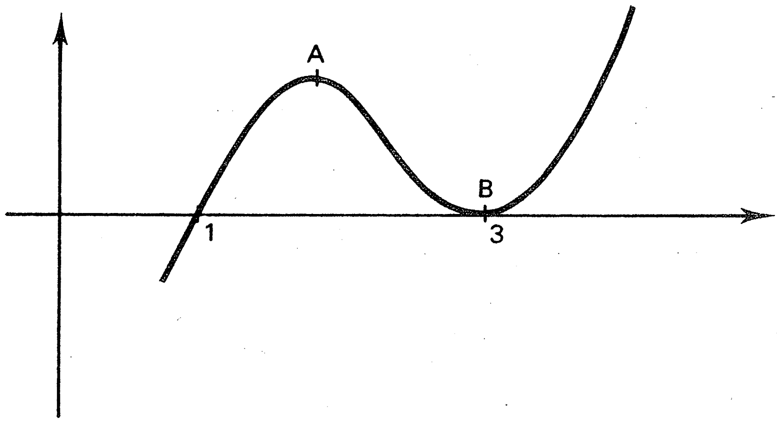


There are two distinct turning points, A and B yet only one real solution located at P.

The existence of just one real solution is a consequence of both turning points lying on the same side of the x - axis.

Solving the equation by non-graphical means would provide the solution P and two other solutions involving negative square roots (i.e. complex solutions). Therefore, there are one real and two complex solutions.

Consider further the cubic $y = x^3 - 7x^2 + 15x - 9$. The graph of this appears as



There are, clearly, two distinct real solutions to $x^3 - 7x^2 + 15x - 9 = 0$ and these are 1 and 3.

Because B lies on the axis the two solutions that would have occurred, had it been below the axis, have merged into one to give a repeated root (or solution) at $x = 3$.

Factorization confirms this, since,

$$x^3 - 7x^2 + 15x - 9 = 0 = (x - 1)(x - 3)^2 = 0 .$$

The power of two, means that $x = 3$ occurs twice (in the same sense as we considered $x^3 = 0$ to have three solutions equal to 0).

To summarize what has been observed for cubic equations:-

(i) there can be *three* distinct (different) real solutions,

(ii) there can be *three* co-incident (equal) real solutions,

(iii) there can be one real and two complex solutions (a total of *three*)

or

(iv) there can be *three* real solutions, two of them coinciding and one distinct.

There is no way that, for unrestricted x values, a cubic can be drawn that doesn't cut the x axis at least once. Thus, allowing for complex and repeated solutions a cubic equation will always have *three* solutions.

Much of what has been observed in the foregoing regarding cubic equations is relevant to polynomial equations of higher degree. It should be noted that the maximum number of turning points of a cubic polynomial is 2 (1 less than the degree)and the number of solutions to the resulting equaiton is always 3 (the degree of the polynomial) provided that repeated and complex solutions are included. It was also observed that for both the quadratic and cubic situations, complex solutions, when they occurred, did so in pairs. A polynomial of degree n has at most $n-1$ turning points and, permitting complex and repeated solutions, the resulting equation has exactly n solutions. Complex solutions occur in pairs and arise when successive turning points lie on the same side of the x axis. Repeated solutions occur when turning points fall on the x axis and/or merge with one another. There are, for unrestricted x values, no breaks in a polynomial curve, i.e. polynomials are said to be continuous.

The question was posed earlier, 'Is there a formula for writing down the solutions of a cubic equation as there is for solving quadratics'? The answer is yes but there are various conditions that make the whole thing very messy. What is possible, however, is that with relative ease the nature of the solutions can be stipulated (i.e. how many real, complex, repeated etc.).

FACTORS AND BOUNDS

Practice examples for solving equations are usually contrived with coefficients chosen so as to simplify the working. For example, there is no major difficulty in solving the equation

$$x^3 - 6x^2 + 11x - 6 = 0$$

provided one is able to see that it factorizes into

$$(x - 1)(x - 2)(x - 3) = 0 \ ,$$

giving the solutions $x = 1, 2$ and 3 .

It is important to realize that factorization is deliberately contrived to simplify the problem. It may well be, in a practical problem, that the coefficients are obtained from calculations on certain measurements and an equation such as $3.26\ x^3 + 1.41\ x^2 - 3.29\ x + 8.25 = 0$ (i.e. with non-integer coefficients) would not be unusual. The chances of being able to factorize such an equation are small indeed. Factorization then cannot be considered a practically viable method of obtaining solutions. For low degree polynomial equations graphing the polynomial will provide an approximation to real solutions but very approximate, especially if solutions are a long way apart. The plotting of higher degree polynomials is impractical as a method of obtaining accurate solutions. When seeking solutions to equations that have coefficients generated from some physical context it will be unusual to obtain solutions exactly. One has often to be content with approximate solutions but this is generally no real handicap provided that we can obtain approximations to a stipulated degree of accuracy. In obtaining even approximate solutions we have first to get a starting approximation which may well mean locating a solution between two integers (say 1 and 2). Although drawing the graph is often not attempted, the features observed regarding turning points and their relation to solutions can be useful.

Whilst factorization, as mentioned previously, is not generally a viable practical approach to solving polynomial equations, once the solutions are known of course, so too are the factors. For example, if a quadratic equation has solutions 1.2 and 3.4 then the factors are $(x - 1.2)$ and $(x - 3.4)$ and the original equation can be re-written as $(x - 1.2)(x - 3.4) = 0$. Similarly, if a cubic has solutions 1.2, 2.6 and 3.4 the original equation can be written (factorized as)
$(x - 1.2)(x - 2.6)(x - 3.4) = 0.$

Generalizing, suppose that a cubic has solutions $x = \alpha_1$, $x = \alpha_2$, $x = \alpha_3$; then the equation can be written as

$$(x - \alpha_1)(x - \alpha_2)(x - \alpha_3) = 0$$

which expanded gives

$$x^3 - x^2(\alpha_1 + \alpha_2 + \alpha_3) + x(\alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_3\alpha_1) - \alpha_1\alpha_2\alpha_3 = 0$$

and in general for an n^{th} degree polynomial with solutions $x = \alpha_1, \alpha_2, \ldots, \alpha_n$ the original equation can be written as

$$x^n - x^{n-1}(\alpha_1 + \alpha_2 + \ldots + \alpha_n) + x^{n-2}(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \ldots \alpha_{n-1}\alpha_n) + \ldots + (-1)^n\alpha_1\alpha_2 \ldots \alpha_n = 0 .$$

This relationship between the solutions and the coefficients of the equation enables bounds to be placed on the largest solution (if they are all real).

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the solutions of $a_0x^n + a_1x^{n-1} + \ldots + a_n = 0$ then the above equation must be the same as

$$x^n + \frac{a_1}{a_0}x^{n-1} + \frac{a_2}{a_0}x^{n-2} + \ldots + \frac{a_n}{a_0} = 0$$

and observing that

$$(\alpha_1 + \alpha_2 + \ldots + \alpha_n)^2 = \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2 + 2(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \ldots + \alpha_{n-1}\alpha_n) ,$$

we conclude that

$$\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2 = \left(\frac{-a_1}{a_0}\right)^2 - 2\left(\frac{a_2}{a_0}\right) .$$

Clearly, the numerically largest solution (which is one of $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$) , α_{max} , is such that

$$\alpha_{max}^2 \leq \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2$$

i.e. the numerical value of $\alpha_{max} \leq \sqrt{\left(\frac{-a_1}{a_0}\right)^2 - 2\left(\frac{a_2}{a_0}\right)}$

Similarly,

$$n\,\alpha_{max}^2 \geq \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2$$

i.e. $\sqrt{\frac{1}{n}\left\{\left(\frac{-a_1}{a_0}\right)^2 - 2\left(\frac{a_2}{a_0}\right)\right\}} \leq |\alpha_{max}| \leq \sqrt{\left(\frac{-a_1}{a_0}\right)^2 - 2\left(\frac{a_2}{a_0}\right)}$

Using this on the equation $2x^3 - x^2 - 26x + 40 = 0$ (we can show that all the solutions are real)

$$2.958 \leq |\alpha_{max}| \leq 5.2$$

i.e. the numerically largest solution lies between these values and thus all solutions lie between -5.2 and 5.2 which is a great help in finding a first approximation to the solutions. (In this particular instance the solutions can be fairly easily found to be -4, 2 and 2.5). The usefulness of these bounds is in situations - higher degree polynomials - where solutions are harder to obtain and rough location is necessary. The usefulness of the bounds does depend largely on their closeness which in turn depends on the magnitude of the coefficients a_0 , a_1 and a_2 .

THE SPRAGUE SEQUENCE

Shyen Wong, Year 10, Knoxfield College

In Problem 10.4.2 the Sprague sequence appeared. This is the sequence where

$$s_{n+1} = s_n + 1/s_n$$

and

$$s_1 = 1 .$$

In *Function*, Volume 10, Part 5 (pp.28-29) the approximation

$$s_n \sim \sqrt{(2n)} = R_n$$

was given.

To test this formula, and also test another formula

$$s_n \sim \sqrt{(2n + 1/2 \ \ell n \ n + 1)} = T_n \ ,$$

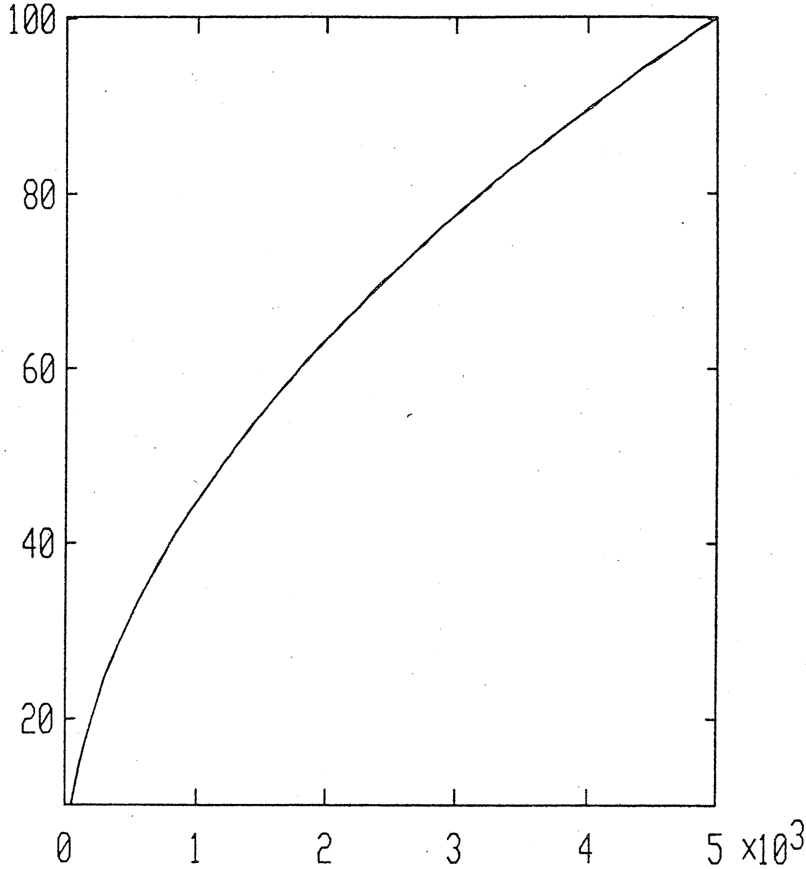
where $\ell n \ n$ is the natural logarithm of n, I computed and graphed s_n, R_n and T_n.

The results I got (see graphs) show that these numbers differ only slightly. The 3 lines joined together and could hardly be told apart.

Graph 2 shows the difference between s_n and R_n. The accurate figure and the approximate figure first move apart, but later come closer together.

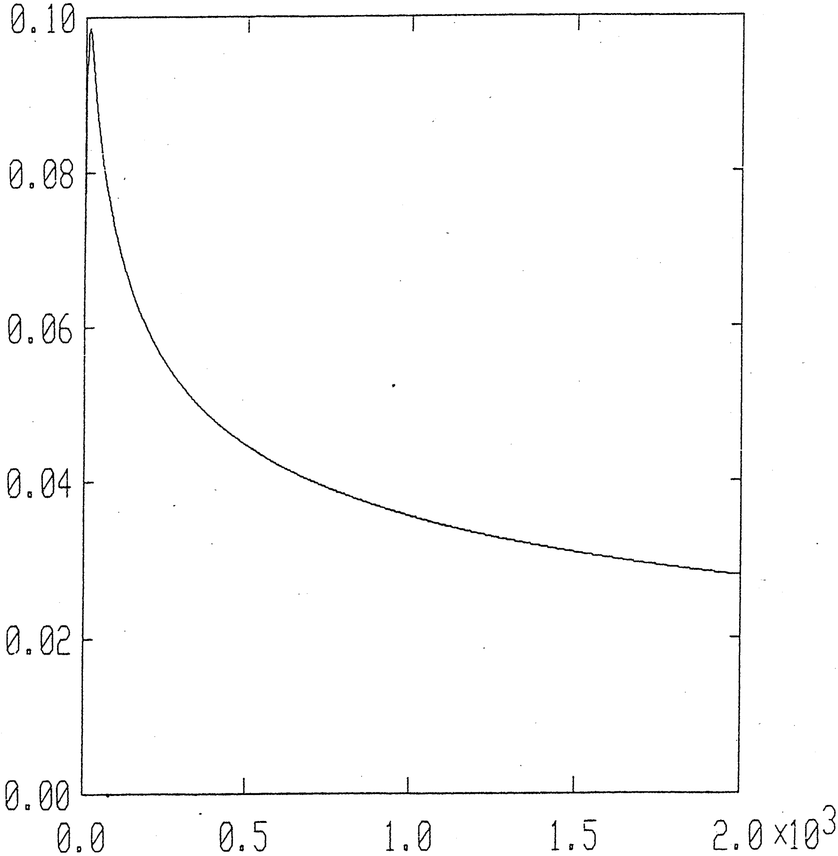
Graph 3 shows the difference between s_n and R_n and the difference between T_n and s_n. The approximate figure of T_n seems more accurate.

[The approximation $s_n \simeq \sqrt{2n}$ was given by David Shaw and John Barton using two separate arguments. Extension of either gives the better approximation $s_n \simeq T_n$. Shyen asks us to thank Geoff Bryan of Monash University for help with the programming.
Eds.]



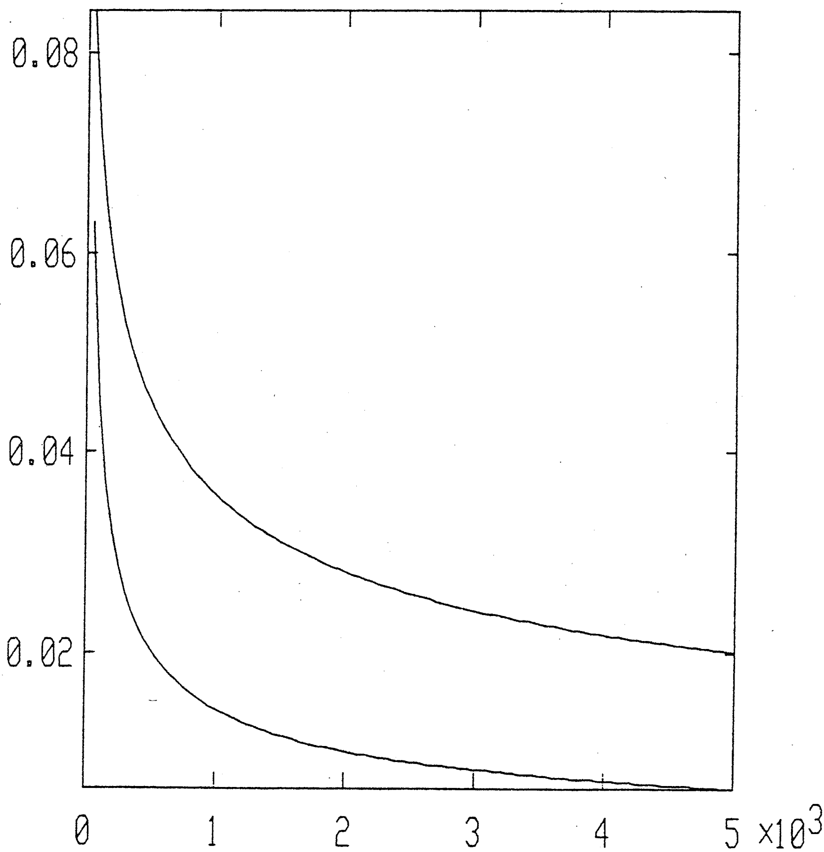
GRAPH 1

The three sequences: R_n, s_n and T_n as functions of n .



GRAPH 2

The difference between s_n and $\sqrt{2n}$.



GRAPH 3

Upper curve: the difference between s_n and R_n : $s_n - R_n$.

Lower curve: the difference between T_n and s_n : $T_n - s_n$.

LETTER TO THE EDITOR

The Front Cover article of Volume 11, No.3 depicts the "Quintenz Balance" and gives its theory as supplied in Barnard's *Statics*. When I first saw this I tried to find who Quintenz was, but I had no luck tracing him at all.

Recently, however, I came upon an article quite by chance. One of the best and oldest of the world's mathematical journals is *Journal für die reine und angewandte Mathematik* (Journal of pure and applied mathematics), known as *Crelle's Journal* after its founding editor A.L.Crelle (1830-1855). Crelle was an engineer who took a great interest in mathematics. While his contributions to mathematics itself were of very slight importance, he was active in both mathematical education and the publication of mathematics. *Crelle's Journal* is his most lasting legacy.

I had occasion to look up the very first volume (1826) of *Crelle's Journal* and, as I turned the pages, the name Quintenz caught my eye. A short article gives the theory of the Quintenz balance, essentially in the same form as Barnard gives it. It also gives a brief background.

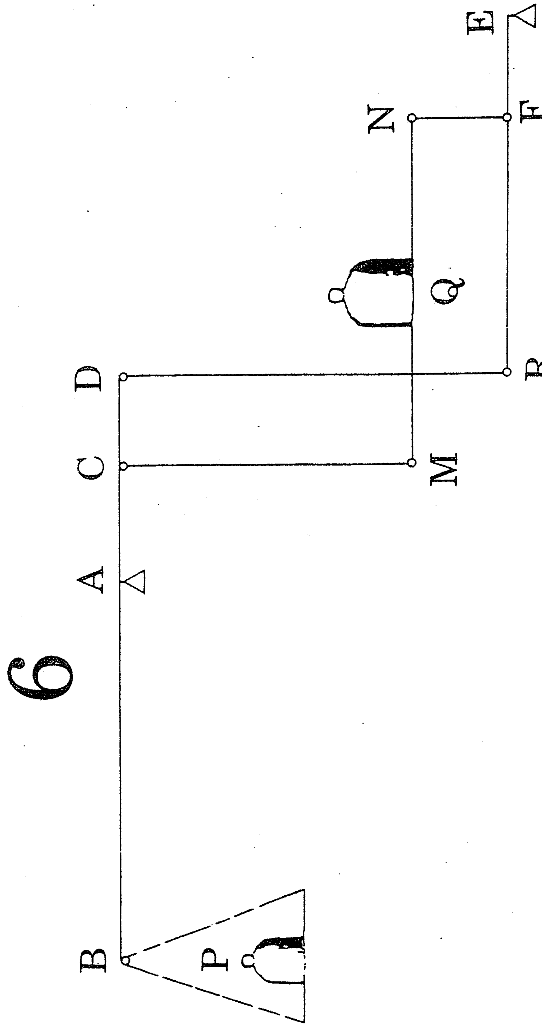
Quintenz seems to have been an inventor. In 1821, he was in Strasburg and it was here that he displayed his newly invented balance. It took the eye of a M.Francoeur† who described it in Issue No. 234 of a journal called *Bulletin de le societé d'encouragement* (presumably this society encouraged inventors and the like) and the article in *Crelle's Journal* is based on Francoeur's description. The author of this article is called E. I don't know who he was. Crelle published short anonymous or cryptonymous articles on topics of lesser mathematical importance and this is one. (For two other such cases, see the article on 0^0 in *Function Vol.5 Part 4*.)

E's picture of the Quintenz balance is reproduced opposite.

Michael A.B. Deakin

---

† Possibly Paul Francoeur (1803 $\geq$ 1865), a Parisian mathematics teacher.


6
B
A
C
D
P
Q
M
R
N
F
E

ARE WE CONSISTENT?

G.A.Watterson, Monash University

When we gamble, are we consistent? We often have to assess whether some games are more favourable to us than others. If so, it would be sensible to play the more favourable ones and not play the others.

Recently, our consistency has been put to the test and we have been found wanting! I here describe an experiment conducted by Amos Tversky and Daniel Kahneman, psychologists working in North America. They asked people to choose between two lotteries, A and B, as to which one would be better to play. They also asked people to choose between two other lotteries, C and D. Before telling you what happened, I invite you to choose whether *you* would prefer to play lottery A or B, and whether *you* would prefer lottery C or D. In all four lotteries, there are some coloured marbles in a barrel, and you win (or lose) some money depending on which colour is drawn out. The percentages of the colours, and the prizes involved, are given in Tables 1 and 2.

Table 1

A choice between lotteries

Lottery	White	Red	Green	Blue	Yellow
A	90%	6%	1%	1%	2%
	$0	Win $45	Win $30	Lose $15	Lose $15
B	90%	6%	1%	1%	2%
	$0	Win $45	Win $45	Lose $10	Lose $15

Table 2

Another choice between lotteries

Lottery	White	Red	Green	Yellow
C	90%	6%	1%	3%
	$0	Win $45	Win $30	Lose $15
D	90%	7%	1%	2%
	$0	Win $45	Lose $10	Lose $15

If you preferred lottery B to lottery A, then you are in good company. All the people in the experiment chose B, probably because they noticed that the proportions of the colours were the same in both lotteries, but in lottery B, the prizes were always at least as good as those in A and were actually better if green or blue were drawn.

Now, which of C or D did you prefer? Most people in the experiment chose C. And they were wrong to do so! Although it looks as if C has the better prizes (in particular, you win if green is drawn in C, but you lose in D), yet notice that the percentages of the colours are not the same in C and D, which makes the choise somewhat confusing : lottery D is certainly the better choice, because lottery D is really equivalent to B, lottery C is really equivalent to A, and B is definitely preferable to A. The equivalences mentioned here are because, in lottery A, there are only four *distinct* prize values ($0, $45, $30, - $15) which have respective percentages (90%, 6%, 1%, 3%). Note that the combined contribution of blue and yellow is 3%; they both yield -$15 in lottery A. And these prizes, and their percentages, are the same as in lottery C. Similarly, you can check that the prizes and the percentages in lottery B are the same as for lottery D.

So if you chose lotteries B and C as being better than A and D respectively, you were being inconsistent. This inconsistency is a stumbling block to theories of economics and probability which assume that peoples' choices between various alternatives are consistent.

* * * * *

Strange as it may sound, the power of mathematics rests on its evasion of all unnecessary thought and on its wonderful saving of mental operations. - Ernst Mach.

A single curve, drawn in the manner of the curve of prices of cotton, describes all that the ear can possibly hear as the result of the most complicated musical performance. ... That to my mind is a wonderful proof of the potency of mathematics.
- Lord Kelvin.

PROBLEM SECTION

SOLUTION TO PROBLEM 11.3.1

Let $f(x) = (x + \frac{1}{x})^2$.

If $0 \leq x < \frac{1}{2}$, if $a = \frac{1}{2} - x$, $b = \frac{1}{2} + x$, then the desired inequality may be written in the form:

$$f(\tfrac{1}{2} - x) + f(\tfrac{1}{2} + x) \geq 2f(\tfrac{1}{2})$$

or

$$f(\tfrac{1}{2}) - f(\tfrac{1}{2} - x) \leq f(\tfrac{1}{2} + x) - f(\tfrac{1}{2}) .$$

In this form it follows at once from the Mean Value Theorem, since $f'(x)$ is increasing.

SOLUTION TO PROBLEM 11.3.2

Let the coins be numbered 1, 2, ..., 12, and be weighed as follows:

1.	1, 2, 3, 4	v.	5, 6, 7, 8
2.	1, 2, 5, 9	v.	3, 6, 10, 11
3.	1, 5, 7, 10	v.	2, 8, 11, 12 .

For each weighing we record either *B* (balance), *R* (right side down), or *L* (left side down). It may be checked that each of the 24 possibilities for the 12 coins may be identified by a unique *B-R-L* sequence.

SOLUTION TO PROBLEM 11.3.3

Since 107 is prime, we must have $b = 1$ or $b = 107$. If $b = 107$, then $a = 61.78$, which is not an integer. So $b = 1$, whence $a = 6$ and $c = 198$.

J. G. Kupka
Monash University