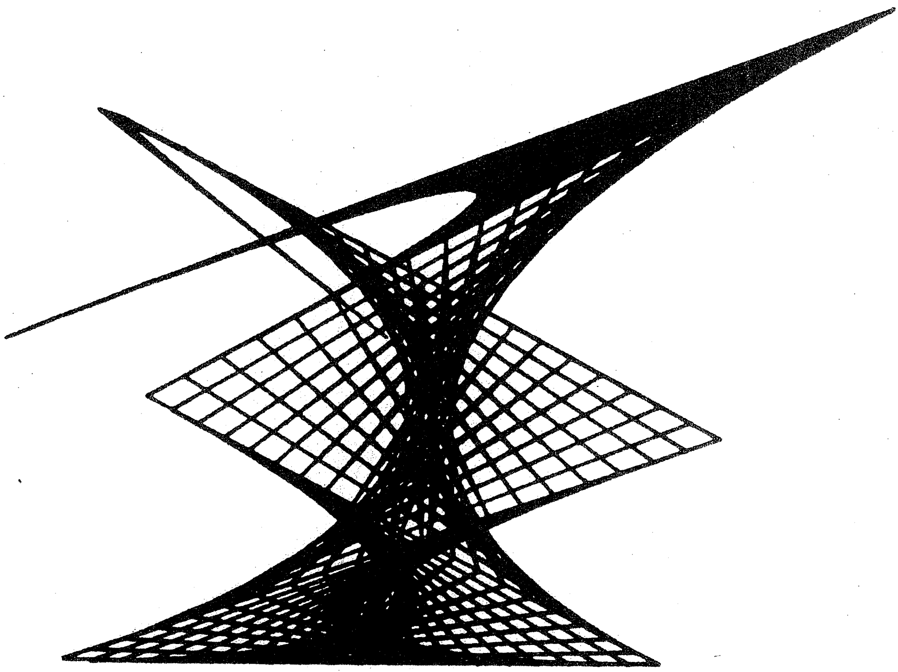
ISSN 0313-6825

FUNCTION

Volume 2 Part 2

April 1978



A SCHOOL MATHEMATICS MAGAZINE

Published by Monash University

Function is a mathematics magazine addressed principally to students in the upper forms of schools. Today mathematics is used in most of the sciences, physical, biological and social, in business management, in engineering. There are few human endeavours, from weather prediction to siting of traffic lights, that do not involve mathematics. *Function* contains articles describing some of these uses of mathematics. It also has articles, for entertainment and instruction, about mathematics and its history. Each issue contains problems and solutions are invited.

It is hoped that the student readers of *Function* will contribute material for publication. Articles, ideas, cartoons, comments, criticisms, advice are earnestly sought. Please send to the editors your views about what can be done to make *Function* more interesting for you.

∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

EDITORS: G.B. Preston (chairman), N.S. Barnett, N. Cameron, M.A.B. Deakin, B.J. Milne, J.O. Murphy, G.A. Watterson, (all at Monash University); N.H. Williams (University of Queensland); D.A. Holton (University of Melbourne); E.A. Sonenberg (R.A.A.F. Academy); K.McR. Evans (Scotch College, Melbourne)

BUSINESS MANAGER: Dianne Ellis (Tel. No. (03) 541 0811, Ext. 2591)

ART WORK: Jean Hoyle

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 addresses shown above.

The magazine will be published five times a year in February, April, June, August, October. Price for five issues (including postage): \$3.50; single issues 90 cents. 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.

∞ ∞ ∞

Registered for posting as a periodical - "Category B"

Throughout its history mathematics has usually been concerned about the rigour of its arguments only to the extent that was necessary. The standards of argument appropriate to your time are those necessary to convince your contemporaries. The article on Archimedes' Method in this issue shows the great clarity of Archimedes' understanding of the difference between a plausible conjecture and a convincing proof. This clarity was not to appear again in the general run of mathematics until the nineteenth century.

The article on Earth, Air, Fire and Water discusses the mistaken reasoning used in early attempts to understand the growth of plants. None of the insight of Archimedes here!

By the end of 1977, the first year of *Function*, we had built up a large correspondence from readers, mainly in sixth form. We have largely a new set of readers this year and ask again for your comments, letters, articles, problems and solutions. Begin by letting us have your comments on David Dowe's views about calculators (see page 31).

CONTENTS

THE FRONT COVER

J.O. Murphy, Monash University

Most computer systems and many programmable hand calculators have a random number generator for a variety of applications in mathematics.

The random number generator program for the Hewlett Packard 25 calculator is simply based on the following formula:

$$u_i = \text{fractional part of } ((\pi + u_{i-1})^5).$$

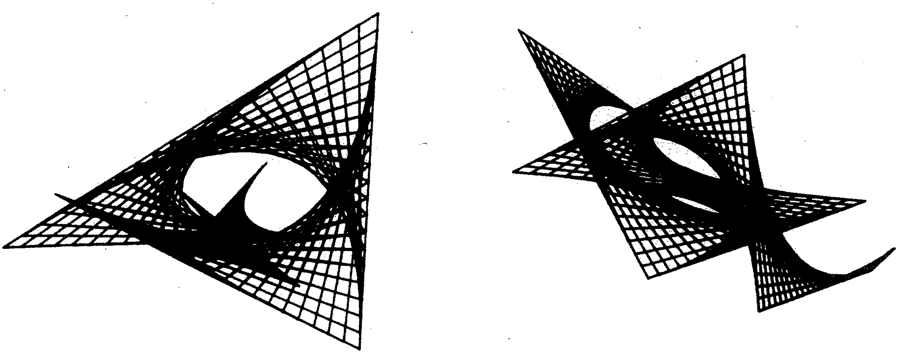
The user has to specify the initial value, u_0, $0 \leqslant u_0 \leqslant 1$, and the program (with $i = 1$) calculates u_1. The program then, for $i = 2$, calculates u_2 from u_1; and from u_2 the program calculates u_3; and so on. The sequence $u_0, u_1, u_2, \ldots$ is our sequence of "random" numbers. By employing the linear transformation

$$U_i = u_i d,$$

we get a sequence of random numbers $U_0, U_1, U_2, \ldots$ now in the range $0 \leqslant U_i \leqslant d$.

Answers to problems, encountered in such diverse areas as simulated traffic flow and nuclear physics, can be found from methods using large sequences of random numbers.

For the front cover a random number generator has been used to construct a sequence of ordered pairs (x_i, y_i), $i = 0, 1, 2, \ldots$, by taking successive pairs from a random number sequence. The diagrams on the front cover and on this page have been constructed, with the aid of a computer graphics terminal, by regarding these pairs as the co-ordinates of points and connecting successive points in the sequence above by straight line segments and then, in the manner described in *Function*, Volume 1, Part 5, drawing enveloped parabolas between pairs of these lines. The lines joining (x_0, y_0), (x_1, y_1) and (x_1, y_1), (x_2, y_2) have been used as the base lines for the first parabola and then the lines joining (x_1, y_1), (x_2, y_2) and (x_2, y_2), (x_3, y_3) for the second and continued interactively until the desired pattern was obtained.



DYNAMIC PROGRAMMING WORKING BACKWARDS

Neil Cameron, Monash University

We wish to travel from start S, to finish, F, on the street plan shown in Figure 1(a), moving always in the direction of the arrows. Each link from node (or junction) to adjacent node has a length, or as we can say, is weighted by distance and the problem is to find a route from S to F whose total length is minimal.

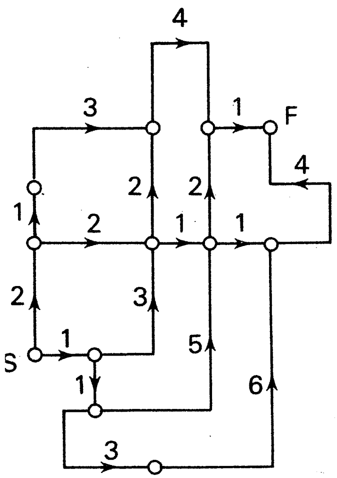


Figure 1(a)

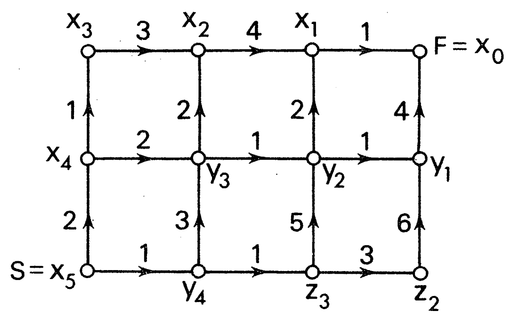


Figure 1(b)

Let us first distort the geometry in Figure 1(a) to produce equivalent information in the grid form of Figure 1(b).

It may now be calculated that there are $\frac{(2 + 3)!}{2!3!} = 10$

possible routes from S to F. Since in this example there are at most two alternatives at each node, we can use a binary code, R for move right, U for move upwards and write down each route along with its length. For example, two possible routes are $RRRUU$ with length 15 and $URRRU$ of length 10. If you consider all ten routes you will find two different routes of minimum length 8. Such routes are described as *optimal* in this context.

Suppose that instead of a 2 × 3 grid as in Figure 1(b) we had a 20 × 20 grid. It would then take a big electronic computer about a month of computation to work out the lengths of all possible routes (more than 10^{11}). For a 100 × 100 grid, no computer (available at the time of publication) could perform the calculations in a lifetime!

Is there a general method for solving such problems which does not involve us in considering all possible routes? One approach uses *dynamic programming*, a procedure based on a simple principle due to the American Richard Bellman and the Russian L.S. Pontryagin, announced in the 1950's. Let us again distort the geometry in our example to produce the equivalent Figure 2. The reason for labelling of nodes in the way shown in Figure 1(b) should now become evident. If we identify our goal F as level 0, then level 1 consists of those nodes x_1, y_1 (on routes) which are one link away from F, level 2 of nodes x_2, y_2, z_2 two links away from F, and so on to level 5 consisting of our starting point S.

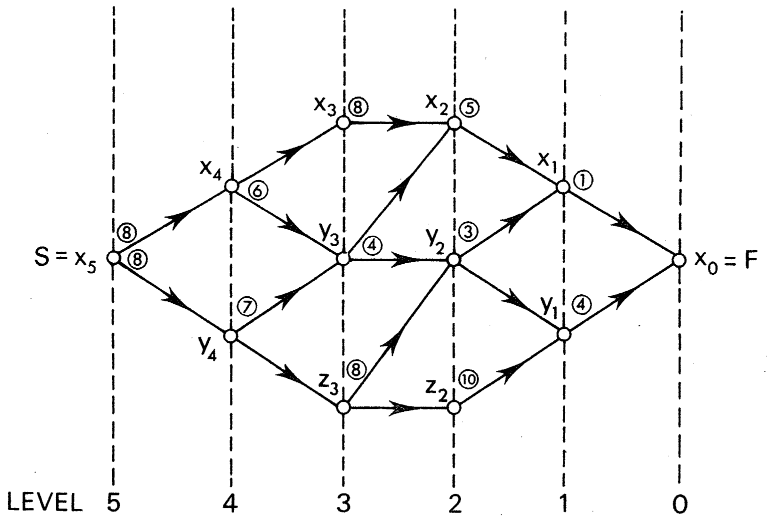


Figure 2

More generally, suppose we have a situation like that of Figure 2, but where the levels are n, n-1, ..., 1, 0. Then routes from $S = x_n$ (the sole node at level n) to $F = x_0$ (the sole node at level 0) have a form

$$x_n u_{n-1} \cdots u_1 x_0$$

where u_i is at level i, $i = 1$ to $n - 1$. The Bellman-Pontryagin principle can be stated as follows:

> An optimal route $R = x_n u_{n-1} \cdots u_1 x_0$ from S to F has the property that each subroute $u_j u_{j-1} \cdots u_1 x_0$ of R from u_j to F is optimal among all routes from j to F, $j = 1$ to $n - 1$.

This is easily proved, for suppose $u_j v_{j-1} \cdots v_1 x_0$, where v_i is at level i, $i = 1$ to $j - 1$, is a route from u_j to F which is better than $u_j u_{j-1} \cdots u_1 x_0$. Then

$$x_n u_{n-1} \cdots u_{j+1} v_j v_{j-1} \cdots v_1 x_0$$

is a route from S to F which is better than R, contradicting the optimal nature of R.

Let us apply this principle to solve our problem (Figure 2). From each node of level 1 there is only one route to F. Enter the length of the route, circled, beside the appropriate node, as in Figure 2. Move back to level 2; here there are two routes from y_2 to F (and two links from y_2 to level 1), whose lengths are $2 + 1 = 3$ and $1 + 4 = 5$. According to the optimal principle we know that if an optimal route from S to F passes through y_2 at level 2 then the subroute from y_2 to F, here $y_2 x_1 x_0$ of length 3, must be optimal among all routes from y_2 to F. Thus the subroute $y_2 y_1 x_0$ from y_2 to F can be discarded from consideration and we enter the length 3 of the optimal route from y_2 to F beside y_2 and above the link to be used in moving from y_2 to level 1. For each of the other two nodes at level 2 there is only one route to F, so enter the length of the route, circled, by the appropriate node. Moving back to level 3 we find two links from each of the nodes y_3, z_3 to level 2. However we need not investigate all three routes from y_3 to F and all three from z_3 to F. For example, if we go from y_3 to y_2 we know the rest of the route to F; the route must be $y_3 y_2 x_1 x_0$, not $y_3 y_2 y_1 x_0$. Why? The possible routes from y_3 to F have lengths $2 + 5 = 7$ and $1 + 3 = 4$, the latter being the better. As at level 1, we select this route ($y_3 y_2 x_1 x_0$) as the best from y_3 to F and enter the length 4 by y_3 and above the link to be used in moving from y_3 to level 2. Similarly we move back to level 4 and finally to S at level 5. We have found two optimal routes from S to F, each of length 8, namely

$$x_5 x_4 y_3 y_2 x_1 x_0 \text{ and } x_5 y_4 y_3 y_2 x_1 x_0.$$

This technique of working backwards from the goal certainly eliminates from consideration several of the ten possible routes from *S* to *F*, and it also solves our problem. For larger problems the effect is much more dramatic. Indeed apparently intractable problems can be solved.

We have solved a problem of minimal length, but the theory is applicable in a much more general context, where the weighting of links measures some decision effect (not limited to binary choice as in our example) and optimal routes become optimal decision policies.

The theory is clearly just as applicable to maximal problems as to minimal problems. Indeed consider our problem as a maximal problem where the link weights are not lengths but rewards of some kind, for example, profits to be maximised. Figure 2 then becomes Figure 3 (where, for clarity, some symbols have been removed and the sole optimal policy is shown arrowed). Notice too that if all links in a given problem are reversed there is an obvious solution to the new problem.

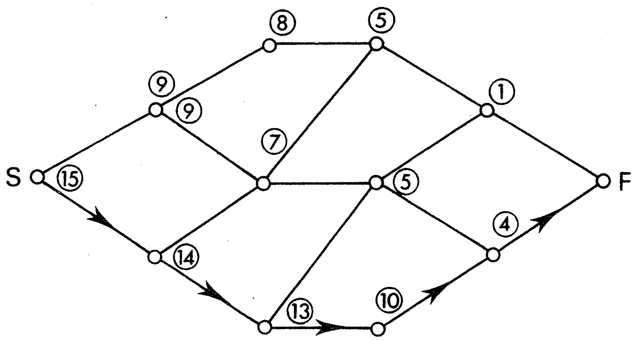


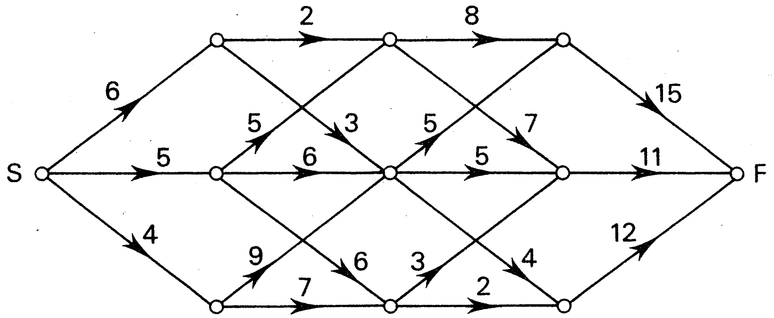
Figure 3

In spite of the very simple procedures applied in dynamic programming, it is an extremely useful technique, widely used in industry and government. In the words of *M.F. Rubinstein

*M.F. Rubinstein, *Patterns of problem solving*, Prentice-Hall, 1975.

"Dynamic programming is more of a fundamental element in the philosophy of problem solving than an algorithm for the solution of a class of problems ..., formulation of a problem as a dynamic programming model is a truly creative task".

Exercise: Find the optimal policies (and their values) (i) as a minimal problem, (ii) as a maximal problem, for the sequential policy decision network shown.



∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

SOLUTION TO PROBLEM 1.5.2 (i.e. Problem 5.2 of Volume 1)

In what follows let Σ denote the sum over j from 1 to ∞. Except for $i = 1$, $j = 1$, we have $X_{i,j} = (X_{i,j-1} + X_{i-1,j})/2$. Hence, if $i > 1$, we have $\Sigma X_{i,j} = (\Sigma X_{i,j-1} + \Sigma X_{i-1,j})/2$. But $X_{i,0} = 0$ and so $\Sigma X_{i,j} = \Sigma X_{i,j-1}$. Hence $\Sigma X_{i,j} = \Sigma X_{i-1,j}$.

(This solution by Geoffrey J. Chappell, Kepnoch High School, Bundaberg; solutions of the special case in Campbell's article also received from David Dowe, Geelong Grammar School, Geelong, and Glen Merlo, Taylor's College, Melbourne.)

∞ ∞ ∞ ∞ ∞ ∞ ∞

PROBLEM 2.1

A man walks in a straight line from A to B, starting at A, at a constant speed of 5km/hr. A fly starts at B at the same time as the man sets off from A and flies to the man's nose, then back to B, then to the man's nose; and so on. The fly flies at twice the speed the man walks. How far has the fly flown when the man reaches B?

PROBLEM 2.2

You can clearly cut a $3 \times 3 \times 3$ cube up into 27 cubes, each $1 \times 1 \times 1$, by 6 cuts. What is the smallest number of cuts that you can use to achieve the same result, perhaps by rearranging the parts after each cut?

ARCHIMEDES' DISCOVERY METHOD

Liz Sonenberg, RAAF Academy

The subject we now call calculus has developed over a long period. Perhaps the most dramatic progress was made during the seventeenth century but much of this development stemmed from ideas discussed by Greek mathematicians in the third and fourth centuries B.C.

One class of problems which had been studied intensively concerned the calculation of areas of figures bounded by curves. Areas of figures which are bounded by straight lines can often be calculated using elementary geometry. One method involves cutting up the area you are interested in and rearranging the parts to form a figure whose area you know. For example with this method triangles, parallelograms and other simple figures 'become' rectangles. (Figure 1)

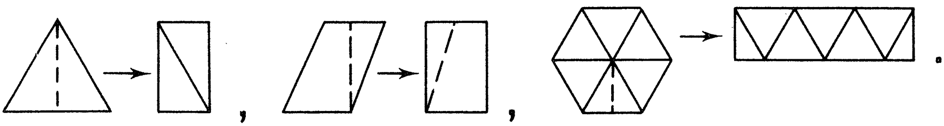


Figure 1

(A detailed discussion of some aspects of this rearranging method appears in the article *Hilbert's Third Problem* which was published in *Function*, Volume 2, Part 1, 1978.)

Area problems involving figures with curved boundaries cannot usually be solved by such elementary means. However geometrical methods can often be used to *suggest* answers to problems where rigorous proofs are rather more difficult to find.

Figure 2 suggests how by 'unrolling' the circle and stretching its circumference along a straight line you might guess the formula

$$A = \tfrac{1}{2}(\text{circumference} \times \text{radius})$$

for the area of a circle (provided you knew the formula $a = \frac{1}{2}(\text{base} \times \text{height})$ for the area of a triangle).

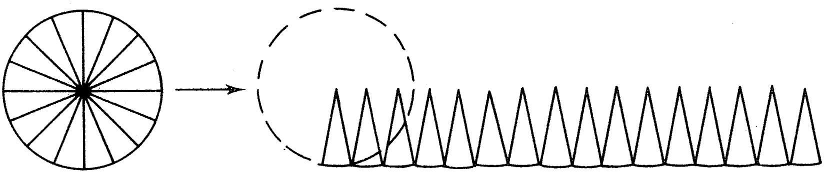


Figure 2

This relation for the area of a circle was known from the fourth century B.C. but was first established on a rigorous basis by the Greek mathematician Archimedes (287 - 212 B.C.).

Archimedes, who was one of the great Greek mathematicians, solved many difficult area and volume problems. He was the first to determine the area and the length of circumference of the circle, that is, to give suitable approximate values of π, and moreover to determine the volume and the surface area of the sphere and of cylinders and of cones. But he went far beyond this; he found the areas of ellipses, of segments of parabolas, and also of sectors of a spiral, the volumes of segments of various solids of revolution, the centroids of segments of a parabola, of a cone, of a segment of the sphere, of right segments of a paraboloid of revolution and of a spheroid. These were amazing achievements, indeed.

Archimedes' proofs used a method, called the method of exhaustion, which had first been applied by the Greek Eudoxus at the beginning of the fourth century B.C. This method is an essentially geometric process in which figures with known areas are circumscribed about and inscribed into the figure whose area one wants to compute. Using this method perfectly rigorous proofs can be given but when one studies these proofs it is difficult to see how the writer might have discovered the result in the first place. Most of the ancient Greek work which is known of today gives no hint as to methods of discovery. There is apparently only one treatise, due to Archimedes, from which we can gain some insight into the methods used to investigate new problems.

This treatise contains a description of Archimedes' discovery method applied to a number of problems and also includes for each problem a rigorous geometric proof of the result. In the preface to the treatise he distinguishes clearly the investigatory methods employed 'to supply some knowledge of the questions' and the formal methods subsequently required to 'furnish an actual demonstration'. His purpose in publishing such an account is made abundantly clear when he says†:

> 'I now wish to describe the method in writing, partly, because I have already

† This translation is from Dijksterhuis, E.J.: *Archimedes*, Copenhagen, 1956, p. 315.

> spoken about it before, that I may not impress some people as having uttered idle talk, partly because I am convinced that it will prove very useful for mathematics; in fact, I presume there will be some among the present as well as future generations who by means of the method here explained will be enabled to find other theorems which have not yet fallen to our share.'

The expressed wish of Archimedes to further the progress of mathematics by the communication of this method appears to have remained unfulfilled. The manuscript which was discovered in Constantinople only in 1906 contains work of Archimedes as copied by a writer in the tenth century. Beyond this, the manuscript seems to have remained unread and unnoticed. It is interesting to note that concepts similar to those discussed in the manuscript were developed during the mediaeval period and were widely publicised in the works of Galileo and Cavalieri in the seventeenth century.

The manuscript includes a letter written by Archimedes to Eratosthenes. The writing is in the hand of a tenth century copyist but later, in the thirteenth century, since nobody there was interested in it any longer, an attempt was made to wash out the old writing and a religious text of the orthodox church was written on the parchment. Fortunately the earlier writing still appears fairly clearly (at least with the aid of a magnifying glass) on most of the 177 leaves of the manuscript.

The part of the manuscript in which we are interested is titled ἔοδοσ, meaning Method. The Method of Archimedes may be called a mechanical infinitesimal method. The fundamental basis of the discovery method is simple and depends upon the imaginary balancing of magnitudes against one another on the arm of a lever. The first proposition proved in the 'Method' concerns the area of a segment of a parabola and may be stated as follows:

PROPOSITION. Consider the segment of a parabola which is bounded by the parabola and the chord AC (Figure 3a). Let D be the midpoint of AC. Draw the line through D which is parallel to the axis of the parabola and let B be the point of intersection of this line with the parabola (Figure 3b). Then the area of the segment of the parabola equals $4/3 \times$ area of triangle ABC.

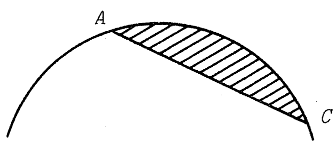


Figure 3a

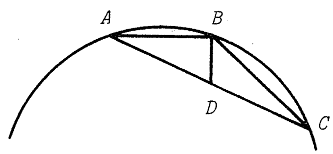


Figure 3b

In his proof of this proposition Archimedes uses a number of properties of the parabola, some of which he has proved in earlier works and others which come from the works of Euclid on conics. We reproduce at the end a translation of Archimedes' proof which was first published in 1912[†] and in this proof we indicate the steps which use these properties of the parabola by a *. You might like to try your hand at finding proofs of these properties.

To understand the proof you also need to know the following three basic properties of a lever.

1. Suppose the lever *XY* has its fulcrum at *W* and the lengths of *XW* and *WY* are *d* units and *d'* units respectively (Figure 4a).

A mass of *m* units suspended at *X* will balance a mass of *m'* units suspended at *Y* if and only if $dm = d'm'$ (Figure 4b).

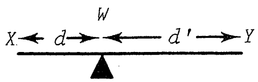


Figure 4a

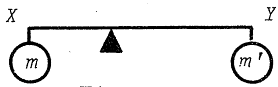


Figure 4b

2. Suppose that we have the masses *m* and *m'* of Figure 4b and also a mass of m_1' units which, when suspended d_1' units from the fulcrum, also balances *m* (Figure 5a).

Then the lever system shown in Figure 5b will be in equilibrium.

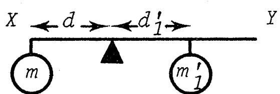


Figure 5a

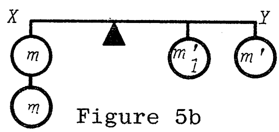


Figure 5b

3. Suppose the lever system syown in Figure 6a is in equilibrium:

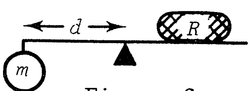


Figure 6a

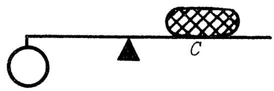


Figure 6b

i.e. a mass of *m* units suspended *d* units from the fulcrum balances the body *R* placed where it is.

We suppose that the body *R* has mass *n* units and that the centre of gravity of the body is at *C* (Figure 6b).

Then a mass of *n* units suspended at *C* will balance a mass of *m* units suspended at *X* (Figure 6c).

[†] T.L. Heath, *The Works of Archimedes*, reprinted by Dover Publications.

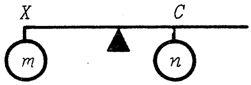


Figure 6c

Proof (of the PROPOSITION stated at the foot of page 10)

'From A draw AKF parallel to DE, and let the tangent to the parabola at C meet DBE in E and AKF in F. Produce CB to meet AF in K, and again produce CK to H, making KH equal to CK (Figure 7).

Consider CH as the bar of a balance, K being its middle point.

Let MO be any straight line parallel to ED, and let it meet CF, CK, AC in M, N, O and the curve in P.

Now, since CE is a tangent to the parabola and CD the semi-ordinate,

$$EB = BD$$

"for this is proved in the Elements [of Conics]†." - - - - *

Since FA, MO are parallel to ED, it follows that

$$FK = KA, \quad MN = NO.$$

Now, by the property of the parabola, "proved in a lemma",

$$MO : OP = CA : AO \qquad - - - - *$$

$$= CK : KN \qquad - - - - *$$

$$= HK : KN.$$

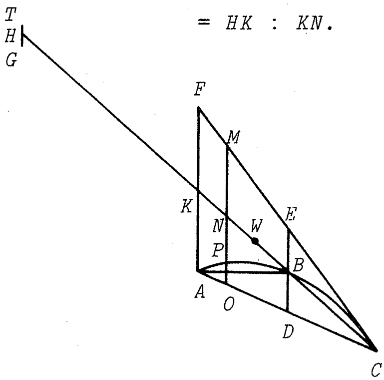


Figure 7

† i.e. the works on conics by Aristaeus and Euclid.

Take a straight line TG equal to OP, and place it with its centre of gravity at H, so that $TH = HG$; then, since N is the centre of gravity of the straight line MO, and

$$MO : TG = HK : KN,$$

it follows that TG at H and MO at N will be in equilibrium about K.

Similarly, for all other straight lines parallel to DE and meeting the arc of the parabola, (1) the portion intercepted between FC, AC with its middle point on KC and (2) a length equal to the intercept between the curve and AC placed with its centre of gravity at H will be in equilibrium about K.

Therefore K is the centre of gravity of the whole system consisting (1) of all the straight lines as MO intercepted between FC, AC and placed as they actually are in the figure and (2) of all the straight lines placed at H equal to the straight lines as PO intercepted between the curve and AC.

And, since the triangle CFA is made up of all the parallel lines like MO, and the segment CBA is made up of all the straight lines like PO within the curve, it follows that the triangle, placed where it is in the figure, is in equilibrium about K with the segment CBA placed with its centre of gravity at H.

Divide KC at W so that $CK = 3KW$; then W is the centre of gravity of the triangle ACF; "for this is proved in the books on equilibrium".

Therefore ΔACF : (segment ABC) $= HK : KW$

$= 3 : 1.$

Therefore segment $ABC = (1/3)\Delta ACF$.

But $\Delta ACF = 4\Delta ABC$.

Therefore segment $ABC = (4/3)\Delta ABC$.

"Now the fact here stated is not actually demonstrated by the argument used; but that argument has given a sort of indication that the conclusion is true. Seeing then that the theorem is not demonstrated, but at the same time suspecting that the conclusion is true, we shall have recourse to the geometrical demonstration which I myself discovered and have already published."'

∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

In dealing with mathematical problems, specialization plays, I believe, a still more important part than generalization.

David Hilbert: *Mathematical Problems*, 1900

THE ORIGIN OF THE SOLAR SYSTEM

A.J.R. Prentice, Monash University

Very shortly I hope to publish the results of several years research into the origin of the solar system. This problem is not a new one but one which has perplexed the minds of astronomers and cosmogonists since the time Copernicus discovered that the planets all revolve about the sun along nearly perfect circular orbits. Perhaps the most plausible explanation of the origin which has so far been advanced was the famous nebula hypothesis of P.S. Laplace, published in 1796. Laplace was impressed by the remarkable order which exists in the planetary system, especially the circularity of the planetary orbits and the fact that all the orbits lie almost in the same plane. He felt that these orderly features were the ones most deserving of attention.

Laplace proposed that at one time the sun was much hotter than it is today and, like a hot balloon, occupied a region of much larger dimension, so large in fact that it encompassed the orbits of all the planets, as we see in Figure 1. As the early sun cooled off it contracted inwards and because of its initial rotation began to spin faster and faster. Gaseous rings were supposed to form near the outskirts of the cloud, like certain bands on Jupiter's surface, and these were successively abandoned at the equator of the sun during its collapse whenever the centrifugal force overcame the gravitational force. Later on, by some unspecified process, the planets were supposed to have condensed from the system of concentric orbiting rings.

This attractive hypothesis held sway for almost one hundred years. In 1884, however, Fouché pointed out that there were certain observational features, notably the present rate of rotation of the sun and the speeds in their orbits of the planets, which in no way, it appeared, could be reconciled with what Laplace had proposed. Other objections were also raised by other people, including Clerk Maxwell. The Laplacian hypothesis fell into disrepute and soon became largely abandoned. Numerous other theories have appeared in its place, none with the same simplicity or appeal, but so far no satisfactory explanation for even a single feature of the solar system has been found.

Whilst studying at Oxford, with the cosmogonist D. ter Haar, some eight years ago, I came to the conclusion that the original objections to the Laplacian hypothesis were, when reviewed in the light of modern data, probably incorrect. Both Fouché and Maxwell, for example, had overlooked the vast amounts of hydrogen and helium gas which must have originally been present in the solar system when the planets were being formed. Encouraged by this turn of events I therefore attempted to develop a modern Laplacian theory. Fresh difficulties soon appeared, however, which seemed to be more serious than the ones originally advanced

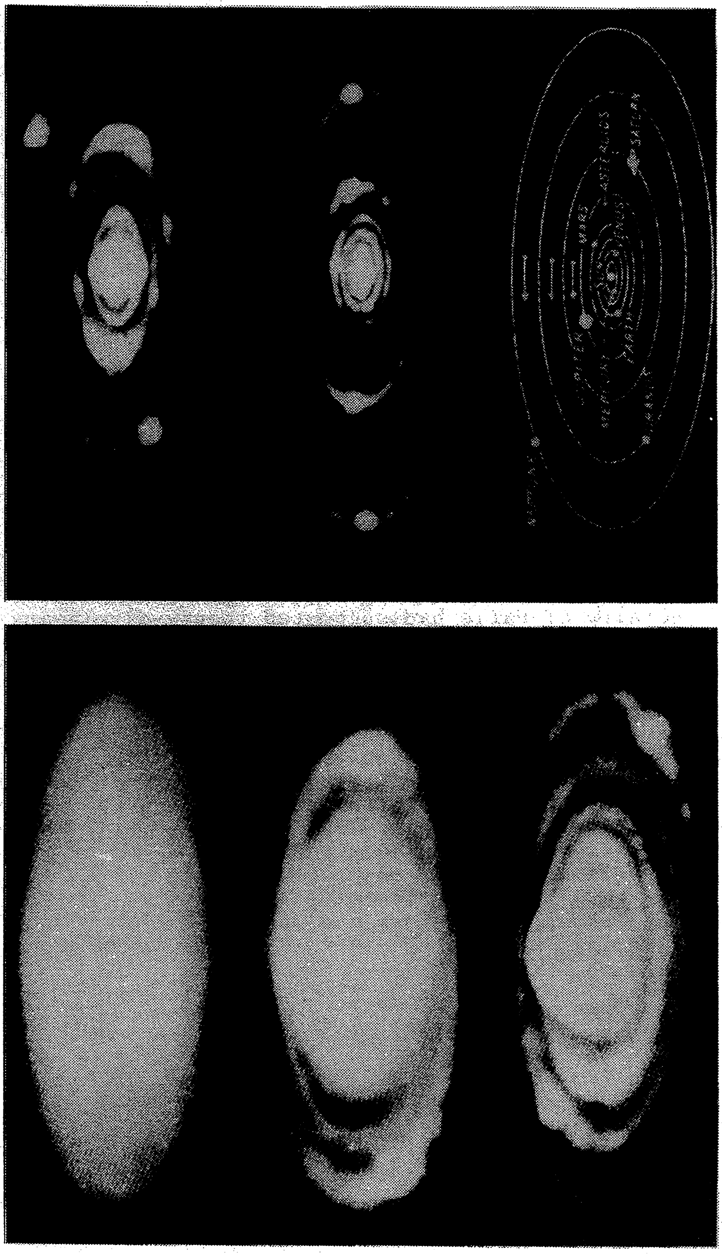


Figure 1: Laplace's contracting nebula hypothesis. A large rotating cloud of gas and dust sheds a system of gaseous rings at its equator as it collapses under its own weight. The planets were later supposed to condense from the rings. (Drawings by Scriven Bolton, Fellow of the Royal Astronomical Society.)

by Fouché. For example, Sir James Jeans had earlier noted in 1928 that if the young sun had given up its excess rotation in the manner proposed by Laplace, then most of its mass must have once been very concentrated towards its centre. This observation follows from the fact that the sun is so massive compared to the planetary system. In addition, no efficient exchange of rotation can take place unless the interior of the sun rotates almost uniformly, like a rigid body. Unfortunately, however, and this was the difficulty, it is not possible to construct from present day physics, stars with the required properties.

A second and perhaps more serious difficulty was to explain why the early sun should seek to dispose of its spin through the shedding at discrete intervals of a system of concentric rings. Surely we should expect the material at the equator to be shed continuously, thus forming a vast disc-like nebula from which a great sheet of rocks and ices covering the plane of the solar system, like Saturn's rings, might emerge.

It therefore became apparent that if the formation of the planets could not be understood in terms of presently known physics, then some hitherto undiscovered physical phenomenon must have been responsible for the origin. My colleague D. ter Haar, of Oxford University, as well as E. Schatzman of the University of Paris had suggested twenty-five years ago that some form of supersonic turbulence may have played a vital role in the formation of the solar system, but no-one knew quite how. With the assistance of a large computer, I therefore attempted to develop the new concept of supersonic turbulent convection which forms the basis of the theory. From a study of very young objects called *T* Tauri stars, which are thought to be young suns producing new planetary systems, I proposed that in the interior of the early sun there may have existed rising and falling columns of hot and cold gas called convective elements or eddies, as shown in Figure 2. So much energy is released during the gravitational collapse of the star that these eddies, often thousands of miles long, can travel many times faster than the speed of sound. When they travel that fast they become long and needle-like and frequently strike into one another.

The upward and downward motion of the eddies creates an additional source of pressure in the star known as supersonic turbulent stress which can be many times larger than the normal gas pressure produced by the motions of the individual molecules. A detailed analysis shows that the net effect of this additional stress is to cause the star to become very centrally condensed and to rotate almost rigidly. Yet these two features are precisely the very ones, mentioned earlier, which we require for the development of a modern Laplacian theory.

Even more exciting was the discovery that in the case of a rotating star, turbulent stress causes the development of a very dense ring of non-turbulent gases at the equator of the star. As the star collapses and begins to rotate faster it tends to bulge out at the equator, as we can see in Figure 3. At the same time the mass of the equatorial belt increases until a

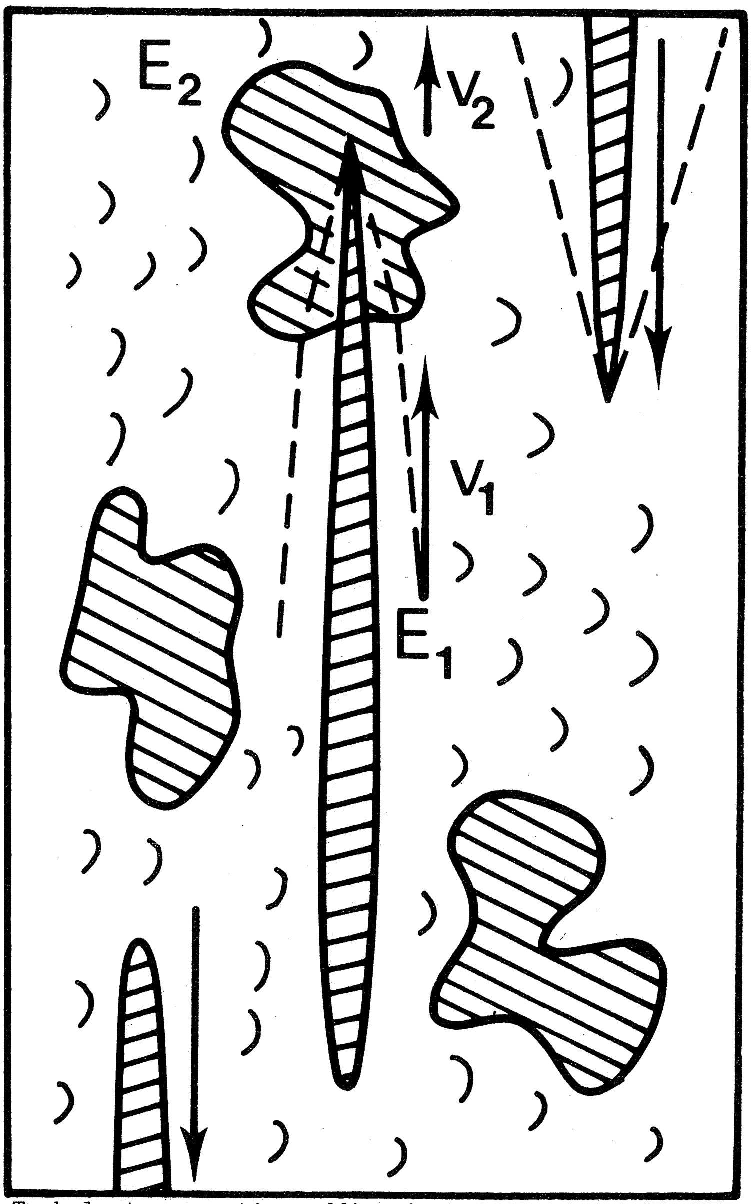


Figure 2: Turbulent convective eddies in a convectively unstable region of a young star. These eddies are rising and falling clumps of hot and cold gas, driven by buoyancy, which become long and needle-like at supersonic speeds. In this diagram a collision is taking place between the long supersonic eddy E_1 and the irregular subsonic eddy E_2; which are travelling with speeds v_1 and v_2 respectively.

critical stage is reached where the centrifugal force just balances the gravitational force. Applied to a star of mass and size of the present solar system, at this point the ring of gases is perfectly suspended in a circular orbit, of radius R_0, say, near the present orbit of Neptune. As the star continues its inward contraction, the non-turbulent ring of gases is unable to follow along with it any further because it has no viscous coupling. The ring is therefore left behind at radius R_0 whilst the collapsing star begins to develop a new equatorial ring in order to maintain pressure equilibrium at its surface. Soon the newly forming ring reaches critical size and is in turn left behind by the sun at a somewhat smaller radius R_1. The whole process repeats itself until the sun reaches its present size.

Supersonic turbulent convection therefore causes the sun to dispose of its excess mass and spin through the successive detachment of a system of gaseous concentric rings. The orbital spacings of these rings R_n $(n = 0, 1, 2, \ldots)$ form a simple geometric sequence which is very similar to the orbital spacings of the planets, whose distances from the sun obey what is known as the Titius-Bode law.

The origin and meaning of the Titius-Bode law has been one of the great unsolved mysteries of cosmogony. According to our theory of angular momentum disposal, the Titius-Bode constant β, which is the mean ratio of one planetary orbit to the next, is given simply by the formula

$$\beta = R_n/R_{n+1} = (1 + m/Mf)^2$$

where m is the mass of the disposed ring, M that of the sun, and $f \simeq 0{\cdot}01$ is a certain constant of the system - called the *moment of inertia coefficient*. Setting $R_n/R_{n+1} \simeq 1{\cdot}7$ the observed value, we predict from this formula that, if our theory is correct, the masses of the primeval rings were each the same and of order 1000 $M_\oplus$ (one $M_\oplus$ is the mass of the earth) of solar material. Such a mass of material, which consists mostly of the light gases hydrogen and helium, contains approximately one earth mass of rocks and about 15 $M_\oplus$ of ice-like materials. It is interesting therefore to note that the earth for example weighs exactly one earth mass whilst the ice-like planets, Uranus and Neptune, weigh 15 $M_\oplus$ and 17 $M_\oplus$ respectively!

Much work remains to be done to test this theory. Nevertheless we feel that our work, at the very least, restores the validity of the original Laplacian hypothesis. The key to the formation lies in the physical process of supersonic turbulent convection. One reason for the slow publication to date of this research has been that the theory that supersonic turbulence is possible has encountered much opposition. Direct observation seems at present very difficult. Nevertheless we are almost certain that a planetary (or satellite) system cannot be formed without something with the properties of supersonic turbulence being

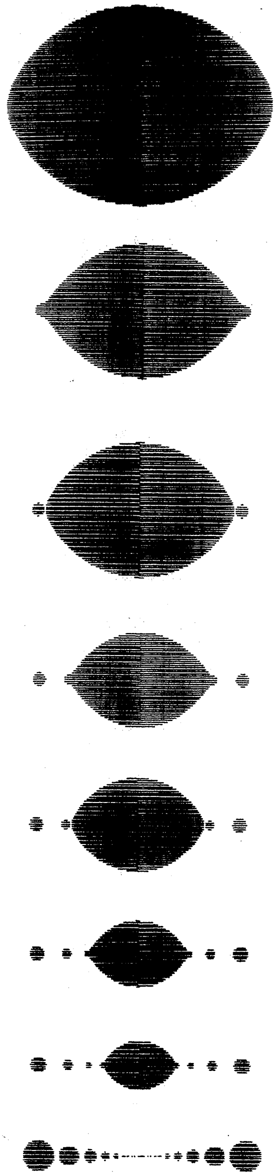


Figure 3: Typical shots taken from a computer-simulated film of the collapse of the large primaeval sun. Each sequence shows a polar cross-section of the rotating protosun at various stages of its gravitational collapse, commencing near the present orbit of Neptune. The sun disposes its excess mass and spin through the successive detachment at its equator of a discrete system of gaseous Laplacian rings from which the planets later condense.

involved. Thus in a field of research where theories often do not survive for more than a few years, it will be interesting to see whether this theory, especially the explanation of the Titius-Bode law, is capable of standing for a longer stretch of time.

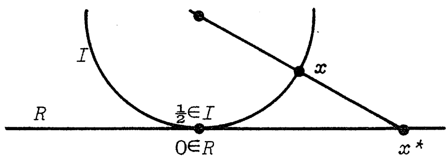
∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

INFINITE NUMBERS II

Neil Williams, University of Queensland

In the last issue of *Function* we introduced a method of deciding whether two sets, finite or infinite, were of the same size. We showed that, with our definition, the set Z^+ of positive integers and the set Q^+ of positive rationals, were, surprisingly, of the same size. We announced that in this article we would show that in fact bigger sets than Z^+ do exist.

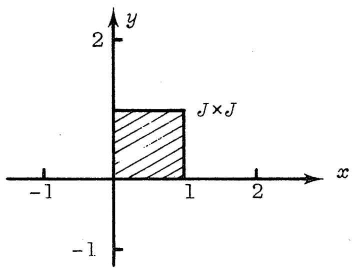
Let us look now at an example involving the set R of all real numbers. Let I be the set of all those real numbers lying strictly between 0 and 1, so $I = \{x \in R|\ 0 < x < 1\}$. Since I is such a small piece of R, surely I will be smaller in size than R? But no, again the two sets are the same size! One way to see this is the following. Think of taking the set I and bending it into a semi-circle, and then balancing the semi-circle on the number line R so the number $\frac{1}{2}$ in I rests on 0 in R. Now for any number x in I, pair x with $x^* \in R$ where x^* is where the line joining the centre of the semi-circle to x cuts the line R. Since the "end points" 0 and 1 are not members of I, this line is never parallel to R so we can always find x^*. You might like to use a little geometry to check that x^* can be found in fact from the formula



$$x^* = \frac{1}{\pi}\tan(\pi(x - \frac{1}{2})).$$

This pairing of x with x^* is enough to show that I and R have the same size.

Now for a really deceptive example. Let us put the "end points" into I, say $J = I \cup \{0, 1\} = \{x \in R|\ 0 \leqslant x \leqslant 1\}$, so I and J are the same size since you have added just two elements to the infinite set I. Let us compare J with $J \times J$, where $J \times J$ is the cartesian product of J with itself, that is,



$J \times J$ is the set of ordered pairs (x, y) where both $x, y \in J$. If you think of the ordinary (x, y) co-ordinates for the plane, then $J \times J$ corresponds to a unit square with one corner at the origin as shown. We can identify the set J with the bottom edge of the square. Thus when we compare J with $J \times J$, we are comparing one edge with the whole square. The square *must* be bigger. But no! In fact J and $J \times J$ are the same size. It is easiest to take a couple of steps to show this. Certainly J is no larger than $J \times J$. This is clear as soon as I say exactly what I meant by "identify the set J with the bottom edge of the square": this means match each $x \in J$ with $(x, 0) \in J \times J$. This gives a pairing of J with some of $J \times J$, and so shows that J is no larger than $J \times J$.

The surprising thing is the second step: also $J \times J$ is no larger than J. Once we know this, so J is not larger than $J \times J$ and also $J \times J$ is not larger than J, by a remark I made before, it follows that $J \times J$ is the same size as J, as I claimed. Before we start, note that each x in J can be written out as a non-ending decimal with 0 before the decimal point, like $\frac{1}{3} = 0{\cdot}3333 \ldots$, $\frac{\pi}{10} = 0{\cdot}3142 \ldots$, $0 = 0{\cdot}0000 \ldots$, $\frac{1}{2} = 0{\cdot}5000 \ldots$, $1 = 0{\cdot}9999 \ldots$. (Usually you don't write an ending of all 0's, just $\frac{1}{2} = 0{\cdot}5$, but for my purposes I want to have the 0's.) Moreover, if we say that apart from $1 = 0{\cdot}9999 \ldots$ representations ending with all 9's are forbidden, then this can be done in exactly one way for each x in J. (I don't want to have, for example, both $\frac{1}{2} = 0{\cdot}5000 \ldots$ and $\frac{1}{2} = 0{\cdot}4999 \ldots$.) So, for each x in J, I can write x uniquely as $x = 0{\cdot}x_1x_2x_3x_4 \ldots$ where each x_i is an integer between 0 and 9. Now to see that $J \times J$ is no larger than J, consider the correspondence which sends the ordered pair $(x, y) = (0{\cdot}x_1x_2x_3x_4 \ldots, 0{\cdot}y_1y_2y_3y_4 \ldots)$ to the number $0{\cdot}x_1y_1x_2y_2x_3y_3x_4y_4 \ldots$ in J. (So, for example, $(\frac{1}{3}, \frac{1}{2}) = (0{\cdot}3333 \ldots, 0{\cdot}5000 \ldots)$ corresponds with $0{\cdot}35303030 \ldots$; $(1, 1) = (0{\cdot}999 \ldots, 0{\cdot}999 \ldots)$ corresponds with $0{\cdot}999999 \ldots = 1$.) You can easily check that this correspondence has the one-to-one property and so gives a pairing of all the elements of $J \times J$ with (some of) the elements of J, thus showing that $J \times J$ is no larger than J. With all this then, we have shown that J and $J \times J$ are the same size.

Perhaps by now you are beginning to wonder if all infinite sets are of the same size. (Of course, this concept of size would not be much use if this was true.) So let me give you an example involving infinite sets of different sizes. I shall show that Z^+ is smaller than $J = \{x \mid 0 \leqslant x \leqslant 1\}$. Certainly Z^+ is not larger than J: the pairing of n in Z^+ with $\frac{1}{n}$ in J shows this. To show that Z^+ really is smaller, we have to show that there is no possible pairing of Z^+ with *all* of J. Realize that this is likely to be a much harder task than anything we have done before. Previously, we have wanted to find pairings between various sets, and as soon as I have shown you *one*, the hunt is over. Now we have to show that *none* is possible, that is that *every* conceivable attempt is doomed to failure.

We shall use a method that is common in mathematics - an argument "by contradiction". To give a "proof by contradiction", you start by assuming that what you want to prove is in fact false, and show that nonsense follows. So if what you want were false, nonsense would hold. So you conclude that what you want must be true.

So here, let us suppose that there *is* a pairing of the elements of Z^+ with all those of J. Think of displaying this pairing with a table, which would look something like the example below, where each element of J is written in decimal form as described above.

element of Z^+	corresponding element of J
1	0·66666 ...
2	0·31415 ...
3	0·25000 ...
4	0·89898 ...
5	0·57721 ...
⋮	⋮

Since the pairing uses up all the members of J, every element of J occurs somewhere in the table. For each positive integer n, define a number x_n, either 1 or 2, as follows: if the number in the table corresponding to $n \in Z^+$ has 1 at its n-th place after the decimal point, then $x_n = 2$; if this number has anything other than 1 at its n-th place, then $x_n = 1$. Consider the number $x = 0 \cdot x_1 x_2 x_3 x_4 \ldots$. (So with the example above, x would start 0·12112 ... .) Certainly $x \in J$. Yet x is different from every number in the table, for x differs from the first in its first decimal place, from the second in its second decimal place, ..., from the n-th in its n-th decimal place, ... So $x \in J$, yet x is not in the table that lists *all* the members of J. And this is indeed nonsense. Thus: if there was a pairing of Z^+ with *all* the members of J, we would get this nonsense. Hence no such pairing is possible. This shows that indeed Z^+ is smaller than J, as I claimed.

We can now answer the question concerning Z^+, Z, Q and R posed at the beginning of this article. We know that Z and Q are the same size as Z^+, and R is the same size as J, so all of Z^+, Z, Q are the same size and all are smaller than R.

∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

I am very well acquainted too with matters mathematical
I understand equations, both the simple and quadratical.

W.S. Gilbert: *Pirates of Penzance*

ALTERNATING SERIES†

In practice, on present-day computers, the values of functions, such as *log* or *sin*, the values of definite integrals, the solutions of differential equations, etc. are nearly always found by approximating the value required by a rapidly converging infinite series. A particularly simple kind of convergent series, although not necessarily rapidly converging, is a special kind of what are called *alternating series*, i.e. series whose terms are all non-zero and are alternately positive and negative.

Perhaps the best known such alternating series is the standard series expression for $\log_e 2$:

$$\log_e 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^{n-1}\frac{1}{n} + \cdots .$$

Note that in this series (a) the successive terms (here $1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \ldots$) are non-zero and are steadily decreasing in numerical (= absolute) value (b) the n-th term (here $(-1)^{n-1}\frac{1}{n}$) tends to 0 as n tends to infinity and (c) successive terms are of alternating sign, i.e. the series is alternating.

Any infinite series which satisfies these three properties (a), (b) and (c) is necessarily convergent. Indeed its sum lies somewhere between the value of its first term and the value of the sum of its first two terms. For example, in the series exhibited, we conclude that $\log_e 2$ lies between 1 and $1 - \frac{1}{2} = \frac{1}{2}$. In fact $\log_e 2$ is approximately $0{\cdot}693$. It is similarly true that for all n, the sum of an alternating series satisfying (a) and (b) lies between the values of the two partial sums, the sum of the first n terms and the sum of the first $n + 1$ terms; and moreover that the error made in taking the sum of the first n terms as an approximation to the infinite sum is less than the numerical value of the n-th term.

For example, applying this statement to the series for $\log_e 2$, we conclude, in turn, that $\log_e 2$ is approximated by

$$\begin{aligned}
1 &\text{ with an error of at most } 1\\
1 - \frac{1}{2} = 0{\cdot}5 &\text{ with an error of at most } \frac{1}{2}\\
1 - \frac{1}{2} + \frac{1}{3} = 0{\cdot}8\dot{3} &\text{ with an error of at most } \frac{1}{3}\\
1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = 0{\cdot}58\dot{3} &\text{ with an error of at most } \frac{1}{4}\\
1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} = 0{\cdot}78\dot{3} &\text{ with an error of at most } \frac{1}{5}\\
1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} = 0{\cdot}61\dot{6} &\text{ with an error of at most } \frac{1}{6},
\end{aligned}$$

†An article on this topic was promised in *Function*, Volume 1, Part, 1, page 26.

and so on; and that $\log_e 2$ lies between

$$\begin{array}{ccc} 1 & \text{and} & 0{\cdot}5, \\ 0{\cdot}5 & \text{and} & 0{\cdot}8\dot{3}, \\ 0{\cdot}8\dot{3} & \text{and} & 0{\cdot}58\dot{3}, \\ 0{\cdot}58\dot{3} & \text{and} & 0{\cdot}78\dot{3}, \\ 0{\cdot}78\dot{3} & \text{and} & 0{\cdot}61\dot{6}, \end{array}$$

and so on.

Since the sum of an alternating series satisfying (a) and (b) is found by alternately adding and subtracting positive quantities which get steadily smaller and which tend to zero, it follows that the behaviour we have exhibited for the series for $\log_e 2$ occurs for any alternating series: the sum of an odd number of terms of the series is always greater than the sum of the infinite series and the sequence of partial sums of an odd number of terms is steadily decreasing with limit the infinite sum. Thus the sequence 1, $0{\cdot}8\dot{3}$, $0{\cdot}78\dot{3}$, ... is steadily decreasing and has limit $\log_e 2$. Similarly, the sequence of partial sums of an even number of terms is steadily increasing with limit again the infinite sum. For example, $0{\cdot}5$, $0{\cdot}58\dot{3}$, $0{\cdot}61\dot{6}$, ..., is steadily increasing with limit $\log_e 2$.

We have said enough to justify all the assertions made. For those who like to use more symbols the situation may be described as follows. Consider the series

$$a_1 - a_2 + a_3 - \cdots + (-1)^{n-1}a_n + \cdots . \quad (*)$$

The n-th term of this series is $(-1)^{n-1}a_n$. Denote the sum of the first n terms by s_n: thus

$$s_n = a_1 - a_2 + a_3 - \cdots + (-1)^{n-1}a_n.$$

Suppose that the series (*) satisfies conditions (a), (b) and (c). This means that each $a_n > 0$, that

$$a_1 \geqslant a_2 \geqslant a_3 \geqslant \cdots \geqslant a_n \geqslant \cdots,$$

and that $a_n \to 0$ as $n \to \infty$.

It follows that $a_1 - a_2$, $a_2 - a_3$, ..., $a_{n-1} - a_n$, are all non-negative. Hence

$$\begin{aligned} s_1 = a_1 &\geqslant a_1 - (a_2 - a_3) = s_3 \\ &\geqslant s_3 - (a_4 - a_5) = s_5 \\ &\geqslant \cdots \end{aligned},$$

i.e. $s_1 \geqslant s_3 \geqslant s_5 \geqslant \cdots \geqslant s_{2n-1} \geqslant \cdots,$

and similarly

$$\begin{aligned} s_2 = a_1 - a_2 &\leqslant (a_1 - a_2) + (a_3 - a_4) = s_4 \\ &\leqslant s_4 + (a_5 - a_6) = s_6 \\ &\leqslant \cdots \end{aligned}$$

i.e. $s_2 \leqslant s_4 \leqslant \cdots \leqslant s_{2n} \leqslant \cdots$.

Moreover, any odd partial sum is greater than each even partial sum, i.e.

$$s_{2n-1} > s_{2k}$$

for any n, k. To see that this is so consider the two cases (i) $2n - 1 > 2k$ and (ii) $2n - 1 < 2k$.

(i) $2n - 1 > 2k$.

We then have

$$\begin{aligned} s_{2n-1} &= s_{2n-2} + a_{2n-1} \\ &> s_{2n-2}, \text{ since } a_{2n-1} > 0, \\ &\geqslant s_{2k}, \text{ as already shown.} \end{aligned}$$

(ii) $2n - 1 < 2k$.

We have

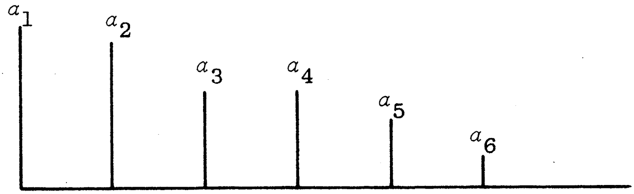
$$\begin{aligned} s_{2k} &= s_{2k-1} + a_{2k} \\ &< s_{2k-1}, \text{ since } a_{2k} > 0, \\ &\leqslant s_{2n-1}, \text{ as already shown.} \end{aligned}$$

In both cases (i) and (ii) we have shown that $s_{2n-1} > s_{2k}$.

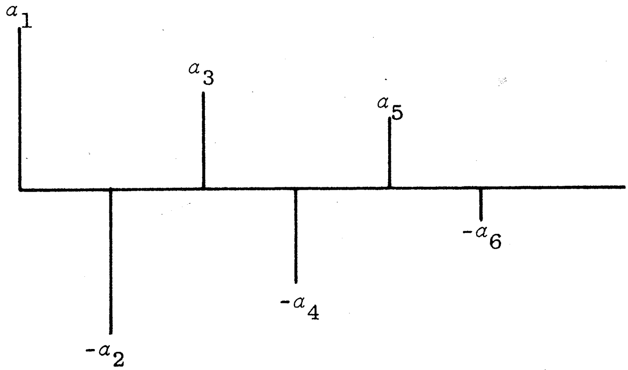
To summarize, we have shown that $s_2 \leqslant s_4 \leqslant \cdots \leqslant s_{2k} \leqslant \cdots < \cdots \leqslant s_{2n-1} \leqslant \cdots \leqslant s_3 \leqslant s_1$. Now observe that $s_{2n-1} - s_{2n} = a_{2n} \to 0$ as $n \to \infty$; in words, the steadily decreasing sequence of odd partial sums gets steadily closer and closer to the increasing sequence of even partial sums, the difference between the two tending to zero.

The common limit to which s_{2n} increases and to which s_{2n-1} decreases as $n \to \infty$ is the sum s of the infinite series (*).

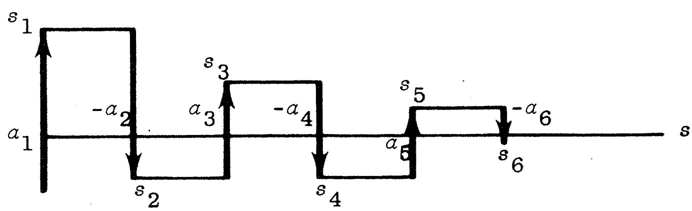
The situation is illustrated in the following pictures.



Steadily diminishing sequence $a_1 \geqslant a_2 \geqslant a_3 \geqslant \cdots$



The same sequence with signs attached



The partial sums s_1, s_3, s_5, ... diminish steadily to s, and s_2, s_4, s_6, ... increase steadily to s.

Finally we mention some well-known alternating series that satisfy conditions (a) and (b).

For all x

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots$$

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \cdots + (-1)^n \frac{x^n}{n!} + \cdots .$$

If $1 \geqslant x > 0$, then these series are alternating series satisfying conditions (a) and (b).

When $x = 1$, in the series for e^{-x}, we get

$$\frac{1}{e} = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \cdots .$$

When $-1 \leqslant x \leqslant 1$,

$$\operatorname{arc}\tan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots$$

which is an alternating series if $0 < x \leqslant 1$. When $x = 1$ we have

$$\frac{1}{4}\pi = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots ,$$

a series that has frequently been used to calculate π to various degrees of approximation. The series was first discovered by the Scottish mathematician James Gregory (1638 - 1675).

∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

PROBLEM 2.3

Fifty knights of King Arthur sit at a round table. Each has a goblet of red or white wine in front of him. At midnight, each passes his goblet to his right hand neighbour if he has red wine, to his left hand neighbour if he has white wine. Assuming that both red and white wine were at the table, prove that someone at the table will be left without wine after midnight.

Is the conclusion still true if the King was also at the table?

PROBLEM 2.4

Let n be an integer greater than 2. Prove that the n-th power of the length of the hypotenuse of a right angled triangle is greater than the sum of the n-th powers of the lengths of the other two sides.

(Hungarian Problems Book II, 1908/2)

TOPICS IN THE HISTORY OF STATISTICAL THOUGHT AND PRACTICE
I. EARTH, AIR, FIRE AND WATER

Peter D. Finch, Monash University

In popular Science Fiction extra-terrestial beings encountering Man are often depicted as astounded and awed by the rapidity of his scientific and technological development over the past few hundred years. It is not mentioned, however, that they could not but also notice our repeatedly demonstrated reluctance to abandon old ideas, even when they have been proved wrong, and be puzzled by the fact that for so long a time it was scarcely recognized that one cannot discover how the world 'works' simply by thinking about it. Even today it is not well understood outside the hard sciences that speculation without the design and analysis of experiments is a fruitless exercise. This important thesis became more widely recognized during the 16th and 17th centuries and was eloquently argued as far back as 1620 in the *Novum Organum* of Sir Francis Bacon (1561 - 1621), the most versatile and highly cultured man of his age; by means of it the road to scientific enquiry was at last discovered. In subsequent centuries, statistical thought developed only slowly as part and parcel of general scientific enquiry, but mainly as something growing out of what exceptionally gifted experimentalists actually did rather than as a body of opinion guiding scientific investigation. The history of statistics is, therefore, closely tied to that of the sciences in general. That statistics has developed so slowly simply reflects the fact that scientific knowledge is hard to come by.

Superficial acquaintance with scientific facts and procedures is so commonplace these days that it is hard for us to appreciate how difficult it was to arrive at them. Indeed it is only too easy to find something both sad and comic in the theories of earlier times and to dismiss them out of hand as little more than evidence of our forebears' naivety. It seems strange today, for example, that it could have been seriously held that all substances were made up of the four elements: earth, air, fire and water. But this idea, inherited from Empodocles of Agrigentum (495 - 435 B.C.) and adopted by Aristotle (384 - 322 B.C.) was so much part of the public domain that one of Shakespeare's characters could say in *Twelfth Night*, without fear of ridicule, *'Does not our life consist of the four elements?'* and, indeed, this was the prevailing view of scientists until the work of Robert Boyle (1627 - 1692) and the publication in 1661 of his best known work, *The Sceptical Chymist*. Yet even Boyle thought that all matter was built up of water and in adopting an atomic viewpoint like that of the Greek philosopher Democritus (460 - 360 B.C.) he supposed that water was the substance of the basic atoms from which all matter was composed. This supposition would not have seemed so strange to Boyle's contemporaries since it was then already accepted that plants were made of water and Boyle himself had repeated an earlier experiment which seemed to establish that fact beyond dispute.

By the 17th century the belief that water had the power of transmuting itself into the various plant substances seemed to be based on conclusive experimental evidence. The experiment in question was originally suggested by Nicolas of Cusa (1401 - 1464) in his *De Staticis Experimentis* but seems to have been first carried out by the Belgian chemist, Jan Baptiste van Helmont (1577 - 1644), an account of it being published after his death when in 1652 his son collected his manuscripts and published them in Amsterdam as *Ortus Medicinae*. This is van Helmont's account of what he did:

> "I took an earthen vessel in which I put 200lb of soil dried in an oven, then I moistened it with rain water and pressed hard into it a shoot of willow weighing 5lb. After exactly five years the tree that had grown weighed 169lb and about 3oz. But the tree had never received anything but rain water or distilled water to moisten the soil (when this was necessary) and the vessel remained full of soil which was still tightly packed; lest any dust from outside should have got into the soil it was covered with a sheet of iron coated with tin but perforated with many holes. I did not take the weight of the leaves that fell in the autumn. In the end I dried the soil once more and got the same 200lb that I started with, less about 2oz. Therefore the 164lb of wood, bark and root arose from the water alone."

Sir John Russell (1873 - 1965), an outstanding agricultural scientist, said of this account:

> "It is a model of scientific communication: terse, clear, omitting nothing essential. Life would be much easier for scientists and their students if more were like it. Everything is right except the conclusion, and that is wrong because van Helmont did not know, and for more than a century nobody knew, that a gas present in the air, carbon dioxide, took part in the process and supplied the carbon which formed a large part of the growth material."†

For over 30 years Russell was director of the famous Rothamsted Experimental Station where, during the 1920's, Sir Ronald Fisher (1890 - 1962), eminent both as a geneticist and the greatest

---

† Sir John Russell, *A History of Agricultural Science in Great Britain*, George Allen & Unwin Ltd, London, 1966.

statistician of modern times, engaged in the pioneering work on agricultural experimentation which was to lay the foundations of modern statistical practice.

As noted above, Boyle repeated van Helmont's experiment. He obtained much the same result but went even further and argued that since worms and insects arise spontaneously from the decay of plants they too must be produced by the transmutation of water.

There is an important lesson to be learnt from this experiment: *viz.* that it is always necessary to allow for the possibility that currently unknown factors might also account for what we observe. This does not mean, however, that we cannot design other experiments to check on that possibility and, perhaps, reveal something about the nature of those unknown factors. For example, it was not too long before John Woodward (1665 - 1728) performed another experiment which cast doubt on the view that plants were composed simply of water in spite of the seemingly incontrovertible evidence of Helmont and Boyle. Woodward grew sprigs of mint in three *different* environments: in distilled water, in water from Hyde Park conduit and in the Hyde Park water shaken up with earth. After 56 days he found that the sprigs had made the following gains in weight in grains.

	Distilled water	Hyde Park conduit	Hyde Park conduit shaken with soil
Gain in weight	41	139	284
Weight of water transpired	8803	13 140	14 956
Water transpired per unit of gain	215	95	53

Woodward argued that if the plant substances had been made up solely of congealed water, and not due to something in the water that had come from the soil, then the gains in weight should have been proportional to the amounts of water transpired. Since this was not so he concluded that growth was due to "a certain peculiar terrestrial matter" and went on to say:

> "It hath been shown that there is a considerable quantity of this matter contain'd both in rain, river and spring water; that the greatest part of the fluid mass that ascents up into the plants does not settle or abide there, but passes through the pores of them and exhales up into the Atmosphere; that a

great part of the terrestial matter, mixt with
the water, passes up into the plant along with
it; and that the plant is more or less augmented
in proportion as the water contains a greater
or smaller quantity of that matter. From all
of which we may reasonably infer, that Earth,
and not Water, is the matter that constitutes
vegetation."†

In this experiment Woodward made the crucial step of recognizing the importance of varying experimental conditions to allow for a corresponding variation in other factors and the detection of their effects, should they exist. He was, of course, still a long way from knowing the role played by atmospheric carbon dioxide in plant growth but, nevertheless, by varying experimental conditions he was able to discover that some factor other than the amount of water did have a role to play. The experiment is also noteworthy because it is an early example of the use of an experiment to disprove a previously held hypothesis and embodies essentially the same argument that Fisher was to develop in the 1920's under the guise of the so-called significance test of a null hypothesis. In this case the null hypothesis is that gain in plant weight is proportional to weight of water transpired independently of what other matter is in the water. Woodward's data clearly runs counter to this hypothesis and indicates instead that gain in weight increases with the amount of other matter in the water.

---

† John Woodward, *Philosophical Transactions of the Royal Society*, 21 (1699), pp. 193-227.

∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

LETTER FROM DAVID DOWE (GEELONG GRAMMAR SCHOOL, GEELONG)

One of the editors (GBP) wrote to David Dowe saying that he thought that computing "is going to have an increasingly major effect on mathematics. It is certainly going to enhance the importance of mathematical analysis [this word has a technical meaning within mathematics which is not intended here] in applications of mathematics. It has given a tremendous impetus to thinking about how results are proved [so as] to answer the question what can be calculated: ...".

David Dowe wrote back (and we give some extracts from his letter):

"I am afraid that I cannot come near agreeing as regards modern (computer) technology.

To me, Pure Mathematics (Mathematics) is a theoretical science, which has conception and understanding as its aims. The mathematician takes nothing for granted (save his axioms).

Isaac Newton was not prepared to take for granted that an apple should fall from a tree. Why didn't it travel sideways? It was this analysis of the 'obvious' taking nothing for granted, that lead Newton to the 'Law of Gravity'. The same can be said of Einstein and relativity. Neither Newton nor Einstein was prepared to take the ideas of everyone else for granted. Why should we take the 'ideas' of the calculator for granted? As far as I am concerned, this is not what Pure Mathematics is all about. Pure Mathematicians are the sort of people who feel compelled to prove what might be regarded by the layman as 'intuitively obvious'. That these same people should take for granted the 'read-outs' of an electronic device (trusting that all its circuits are correctly joined, etc.) is indeed a great irony.

Furthermore, ... I feel that the ability of my fellow students to perform simple arithmetical computation has been greatly impaired by continually relying upon their calculator to answer the question at hand ... . I am not exaggerating as I inform you of these: (1) A Chemistry Student (Form V) was given a problem asking how many moles of substance X are there in 1·08g of X, if the atomic weight of X is 108g. The student, having been brought up to take fright at decimal points and numbers larger than 100, promptly brought out his friend, who saved the day. (2) (You won't believe this!) A student was faced with the problem of 10 ÷ 2. He probably would've tried to break the calculator habit if he hadn't been afraid of suffering withdrawal symptoms. Upon ingeniously pressing the correct buttons, he obtained an answer of 4·9999 (etc., I guess!) He wrote this down ... .

Now that the VUSEB have been brilliant enough to allow calculators (non-programmable, they say) to be used in HSC ... all students are to have calculators, and those who find it difficult to work with powers of 10 should spend an extra 10 or 20 dollars and buy one which works in standard form (what will science think of next!?) ... .

I sincerely feel that calculators, computers (call 'em what you like) are not really serving (Pure) Mathematics.

If I asked you to prove the Four-Color Map Theorem, could you tell me honestly that you understood the proof, could re-construct it (or be sure that the computing was correct)?"

Do you agree with David Dowe? Please let us have your views on computing and calculators.

∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞

"If I have seen farther than Descartes, it is by standing on the shoulders of giants."

Isaac Newton

∞ ∞ ∞

"All things began in order, so shall they end, and so shall they begin again; according to the ordainer of order and mystical mathematics of the city of heaven."

Sir Thomas Browne: *The Garden of Cyrus*, 1658

∞ ∞ ∞

A mathematician asks his engineer friend, "How do you boil water if you have an empty kettle and an unlit gas stove?" "Very easily," replies his friend, "Fill the kettle, light the gas and put the water on to boil." "Good," said the mathematician. "Now solve this problem: the gas is on and the kettle is filled. How do you boil the water?" "No difficulty," replied the engineer. "Just put the kettle on the range." "Better," said the mathematician, "to turn off the gas, empty the kettle and so arrive at the first problem, which we know how to solve."

∞ ∞ ∞

"Perhaps, looking back to 1977, our successors may claim that it was then we first appreciated that the linking computers to the ubiquitous cathode-ray tube in every sitting room made possible not merely childish TV games of ping-pong, but also the ultimate demise of letter writing, of the coinage (why use money or banks when a computer-link through one's TV set will do just as well?), of newspapers, and of schools and school teachers.

Magnus Pyke, *New Scientist*, January 5, 1978

∞