Function

Volume 17 Part 5

October 1993

S   C   S'
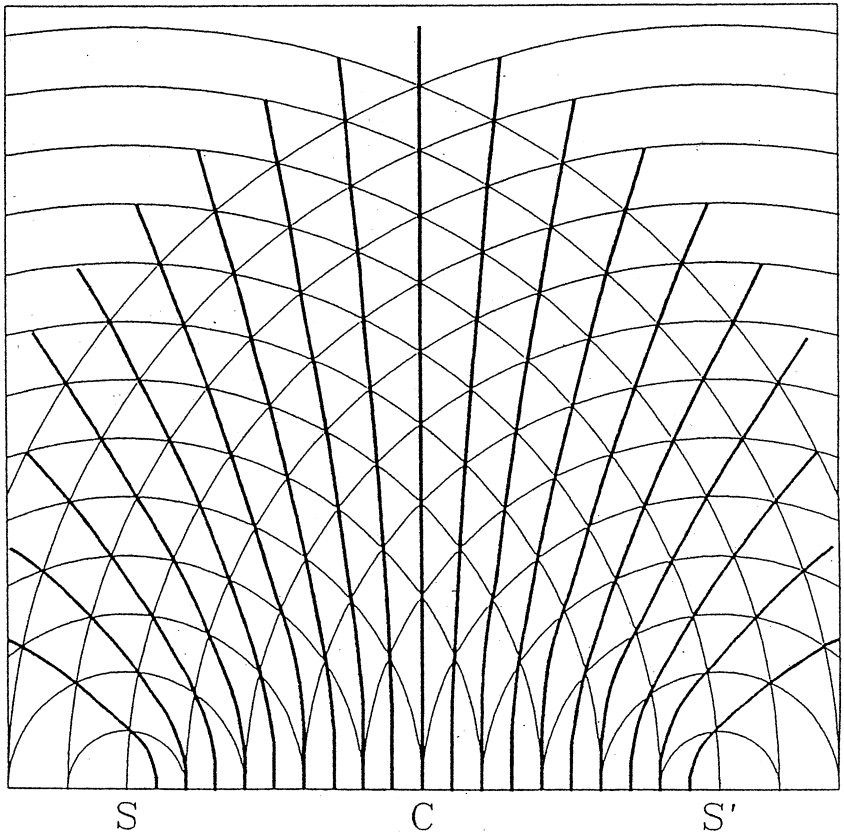
A SCHOOL MATHEMATICS MAGAZINE

FUNCTION

CONTENTS

* * * * *

THE FRONT COVER

Bert Bolton, University of Melbourne

Our cover diagram for this issue is a good example where Physics and Mathematics are brought together in some calculations. It shows the pattern of waves in a plane issuing from two sources. This can be represented in a simple experiment by two rods dipping together into the surface of water and is often the subject of a demonstration in Physics teaching. We will return to this demonstration at the end of the article and show how to do it for yourself.

This piece of Physics was first discussed by Thomas Young in 1802 and his correct explanation of the pattern using light waves was one of the experiments that gave evidence for the fact that light is propagated by waves. In the seventeenth century, Isaac Newton had been a supporter of the idea that light is propagated by particles, or, as they were called in those days, corpuscles; his influence on science was so strong that for over a hundred years his ideas were accepted, even though they could not explain some of Newton's own observations, viz. the coloured circles known as "Newton's Rings". It is worth emphasizing that we are only considering the propagation of light through space or through a homogeneous medium. If we have to ask questions about the interaction of light with matter at an atomic level, we now know that energy is transferred by discrete amounts of energy. To understand the interaction of light with matter we need to use the Quantum Theory.

But we needn't go that far in talking about Thomas Young's classical ideas of wave propagation. In the cover diagram, the two points S and S' are the sources of the two waves that spread outwards. Each of the points S and S' is known mathematically as a focus and the two together as the foci. To keep the physical picture in mind, consider S and S' as the points where two thin rods dip together into the water at regular intervals of time. Each wave spreads out over the water as a pulse, a very narrow region with a disturbance on the surface. We call each circular disturbance a "wavefront" and they appear in the diagram at unit distance apart. The diagram has been drawn with their radii 1, 2, 3, etc. units. In the original diagram, the unit was 1 centimetre (cm) and the distance SS' was 10 cms. Only the top half is given; the bottom half is the mirror image in the line through S and S'. Maintaining the physical picture of the pulses on water, the wave-fronts eventually overlap and give a double amplitude. There are 14 wave-fronts, from S to S'. Beyond these 14 wave-fronts there are no interactions.

The diagram can be interpreted as a "snap-shot" of the wave-fronts at one particular time. But time advances continuously and the intersections with double amplitudes are moving through space and fall on the dotted curves. We now want to find the shape of these dotted curves. (The intersections do not have quite double amplitudes as the waves are damped, but we can ignore that for the moment.)
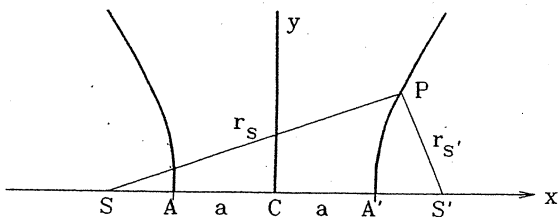

Figure 1

In Figure 1, P is a point on one of the dotted curves, C is the midpoint between S and S' and A, A' are the points where the two branches of the dotted curve meet the horizontal line. It is conventional to call $AC = CA' = a$. The physical condition defining the dotted curve is that

$$r_s - r_{s'} = \text{constant}. \tag{1}$$

When this constant is positive we get the right-hand dotted branch and when the constant has the same absolute value but is negative, we get the left-hand branch. The value of this constant can be given by noting that A' lies on it so that

$$r_s - r_{s'} = SA' - S'A'.$$

There is symmetry about the line through C and at right angles to SS' and

$$S'A' = SA.$$

Thus

$$r_s - r_{s'} = SA' - SA = AA' = 2a. \tag{2}$$

The equation (2) defines the shape of the dotted curve as a hyperbola but the discussion is probably more familiar in terms of the co-ordinates (x, y) of the point P. The origin of co-ordinates is the point C, the positive x-axis is the horizontal line CS' and the y-axis is the vertical line through C at right angles to CS'. It is conventional to call the co-ordinates of S $(-ae, 0)$ and of S' $(ae, 0)$. e is a constant for a given dotted line and known as the eccentricity; we shall discuss it later.

Using Pythagoras' Theorem, Equation (2) can be written as

$$[(x + ae)^2 + y^2]^{1/2} - [(x - ae)^2 + y^2]^{1/2} = 2a.$$

Then

$$(x + ae)^2 + y^2 = \{[(x - ae)^2 + y^2]^{1/2} + 2a\}^2,$$

which gives on expansion and simplification

$$a - ex = [(x - ae)^2 + y^2]^{1/2}.$$

Squaring again and rewriting the result gives

$$\frac{x^2}{a^2} + \frac{y^2}{a^2(1-e^2)} = 1.$$

There are many conventions for the curves such as the one we are examining and we usually write $b^2 = a^2(e^2 - 1)$ to give

$$x^2/a^2 - y^2/b^2 = 1. \tag{3}$$

Notice the two symmetries that we have already met; Equation 3 is unchanged when x is replaced by $-x$ and also when y is replaced by $-y$.

Before continuing with the mathematical analysis of (3), we should appreciate how the cover picture allows us to understand how Thomas Young defined experiments to reveal the

wave nature of light. Young allowed the light waves from two sources S and S' to fall on a white screen parallel to the line SS'. The pattern of the intensity on this screen would be a series of large intensities nearly regularly spaced in the centre, but with spacing increasing towards the outside in both directions. When the distance SS' is much larger than the wavelength of the light waves (usually so) the region on the screen in which the regular spacings occur is wide. This regular spacing of intense positions allows the wavelength of light to be calculated. The wavelength of the yellow light from flame containing sodium is about 0.5 of a micrometre (a millionth of a metre) and the distance SS' is about a millimetre, so the condition is well satisfied to get the regularly spaced lines of intensity on the screen. They are called "interference fringes" or "Young's fringes".

Returning to the Mathematics, let us consider the behaviour of Equation (3) for values of x and y much larger than a and b. The individual terms $\frac{x^2}{a^2}$, $\frac{y^2}{b^2}$ are each much larger than 1, and must almost cancel one another out. If therefore we neglect the 1, we have

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$$

or

$$y/x = \pm(b/a) = \pm(e^2 - 1)^{1/2}, \text{ a constant.}$$

This represents two straight through the origin C and they are known as the asymptotic behaviour of the hyperbola for large x, y. They can be seen on the cover diagram by looking at the dotted curves, which become straight lines in the asymptotic limit. Notice that the equations give the two asymptotic straight lines for the two branches.

The word hyperbola comes from two Greek words meaning "an excessive throw". When a ball is ordinarily thrown in the air it returns to the earth; ignoring air resistance the curve is a parabola. But if the hyperbola is considered to be the path of a projectile it has the asymptotic behaviour of "going off to infinity" as a straight line. To the Greeks this must have seemed an exaggerated (or excessive) thing to happen. We now know that a piece of matter, such as an asteroid or a non-returning comet, moving fast enough and coming towards a star (such as our sun) or a planet will be attracted towards it, swing round it and go off into space again. The asymptotic straight lines that we have seen arising from Equation (3) are indeed the straight lines that a material particle describes when not acted on by a force, which is precisely an expression of Newton's first law of motion. The American satellite or space-craft "Voyager" was guided from planet to planet on such hyperbolic paths and is now on its way outwards from the solar system on its final asymptotic straight line.

The rectangular hyperbola ($e = 2^{1/2}$) has properties that are displayed readily in the co-ordinate system based on the asymptotic lines which are at 45° to the x- and y-axes. Figure 3 is a sketch showing the x- and y-axes and the two asymptotic lines, shown dotted, through the origin which are at 45° to the x- and y-axes. The asymptotic lines define the x'- and y'-axes which can be seen to be given by a clockwise rotation through 45° from the x- and y-axes. The point P can now have co-ordinates (x, y) or (x', y') and the relationship between them can be seen to be

$$x = y'c + x's \left.\vphantom{\begin{array}{c}a\\b\end{array}}\right\} \quad (4)$$
$$y = y's - x's$$

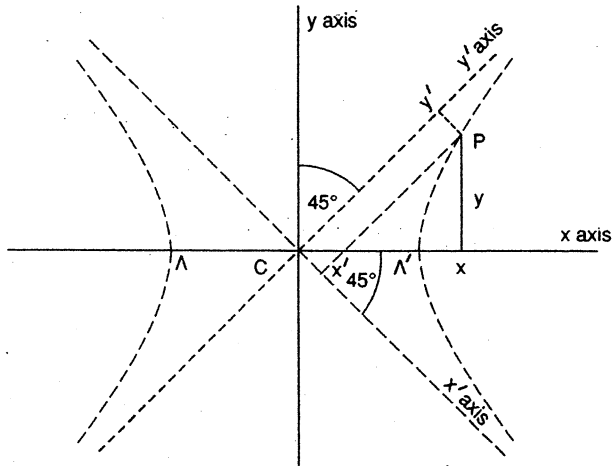**Figure 2**

where $c = \cos 45°$ and $s = \sin 45°$. The values of c and s are both $1/2^{1/2}$ but it is worth writing the equations in general terms so that for any clockwise rotation of the axes from x to x' by an angle θ, these equations (4) are still valid with $c = \cos \theta$, $s = \sin \theta$.

We can now substitute for x and y in Equation (3) which for $e = 2^{1/2}$ becomes

$$x^2 - y^2 = a^2.$$

Using the trigonometrical relationship $c^2 + s^2 = 1$, we get

$$x'y' = a^2/4cs = a^2/2 = \text{constant.} \quad (5)$$

This equation shows that when x' tends to zero, y' tends to infinity, thus defining a "singularity"; this is another expression of the asymptotic behaviour of the hyperbola. The same occurs when y' tends to zero. Expression (5) is well suited to computation.

Finally, it is easy to demonstrate on water the pattern in the Cover Diagram. Fill a hand-basin with water and hold two sharpened pencils firmly in one hand with their points about 10 cm apart. Dip them together into the water somewhere near the middle of the basin at about 3 times per second. The pencil points dip into the water at the foci. A few minutes' practice, to keep your hand steady and dipping rate constant, and the hyperbolae become clear between the two foci with the circular waves outside. You will soon learn how to do it and you will be able to vary the distance SS' and the rate of dipping to see what happens. Water waves are damped and the hyperbolic pattern can only be seen near the foci S and S'.

There are several good books on various mathematical curves and they nearly all include the hyperbola. The one I like, especially for its fine sketches of curves, is

LOCKWOOD, E.H., *A Book of Curves* (1976), Cambridge University Press.

Useful catalogues of curves and their properties are

ARCHIBALD, R.C., *Encyclopedia Britannica* under Special Curves.

YATES, R.C., *Curves and their properties* (1947), Michigan.

Try your local library. There are many older books on drawing curves and computer programs are easy to write. As an exercise, try the equation for the ellipse

$$x^2/a^2 + y^2/b^2 = 1.$$

Notice the difference between this and equation (3). There are close links between the ellipse and the hyperbola. Try the values for (a, b) as $(1, 5)$, $(2, 4)$, $(3, 3)$, $(4, 2)$, $(5, 1)$.

OLYMPIAD NEWS

Hans Lausch, Monash University

The 34th International Olympiad was held in Istanbul from August 13-24. Six questions were posed, originating (in order) from Ireland, the United Kingdom, Finland, Macedonia, Germany and the Netherlands.

For each of the six team members per country and each of the six questions asked, a maximum of seven points was possible. Thus the maximum score per team was 252. The top performers were

1.	P.R. China	215	8.	Hungary	143
2.	Germany	189	9.	Vietnam	138
3.	Bulgaria	178	10.	Czech Republic	132
4.	Russia	177	11.	Romania	128
5.	Taiwan	162	12.	Slovakia	126
6.	Iran	153	13.	**Australia**	**125**
7.	USA	151			

Australia finished ahead of all other Commonwealth countries, and scored its best ever medal tally since joining the IMO in 1980. Our results were:

Gold medal:	Anthony Henderson (year 12), Sydney Grammar School, NSW
Silver medals:	William Hart (12), Elizabeth College, Tasmania Frank Calegari (12), Melbourne Church of England Grammar School, Victoria
Bronze medals:	Rupert McCallum (12), North Sydney Boys' High School, NSW William Hawkins (11), Canberra Grammar School, ACT Simon Schwarz (12), Moriah College, NSW.

[Dr Lausch supplied copies of the questions, but space limitations have prevented us from printing them.]

MATHEMATICS AND GLIDING

Gordon Hookings, University of Auckland

The sport of gliding involves Mathematics in many ways. Firstly, of course, the designer of the aircraft uses Aerodynamics in deciding the aerofoil characteristics to adopt, e.g. the best lift to drag ratio (which also corresponds to the best achievable glide slope in calm air; see Figure 1), and the speed at which that is achieved. Then there are structural considerations and centre of gravity calculations and many design parameters to be fixed. An extensive course in Aerodynamics is really required to cover these topics adequately as, of course, a glider is simply an aircraft without a motor. At least this simplifies the vector diagram representing the glider in flight. There are
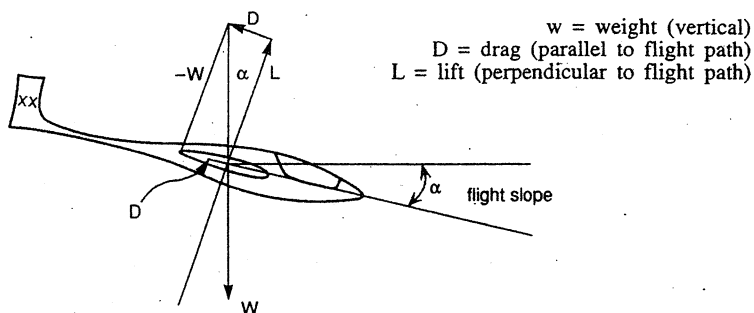

w = weight (vertical)
D = drag (parallel to flight path)
L = lift (perpendicular to flight path)

Figure 1

only two forces involved: the weight acting vertically downwards and the resultant aerodynamic force, a composition of lift L and drag D. If the latter is equal and opposite to the weight, then the motion is unaccelerated. From Figure 1, $\tan \alpha = \dfrac{D}{L}$. A typical number would be $D/L_{\max} = 1{:}45$ or, as glider pilots like to express it, $L/D = 45$, which is interpreted as "for each metre you do down, you go 45 m forward". This exercise can be repeated for different speeds, giving a range of values for L/D.

It should be realized that the glider is generally descending relative to the local air-mass, but if that air is rising at a greater speed than the rate of sink of the glider, then the aircraft can soar.

Once a pilot has mastered the elementary controls, the next step is to find and utilise rising currents of air. Most often these are caused by convection which gives rise to what the glider pilots call "thermals", or "chimneys" of relatively hot air. However, air rises mechanically as wind sweeps over hills to provide "slope soaring" (also indulged in by seagulls), and a very potent source of lift is in "lee waves", downwind of mountain ranges.

Every glider flight is an exercise in optimisation – the pilot tries to maximise duration, height or speed across country. The route to be followed is adjusted according to the pilot's judgement as to where the strongest lift will be found. A deviation from the shortest path is justified if, for example, the overall result is a faster speed. As well as a choice of heading, the pilot can select the velocity through the air. As the forward speed increases, so does the sinking speed and the curve relating these is called the "polar" by glider pilots, which shows the combined result of repeating the analysis of Fig. 1 for different speeds. A typical polar is shown in Figure 2.
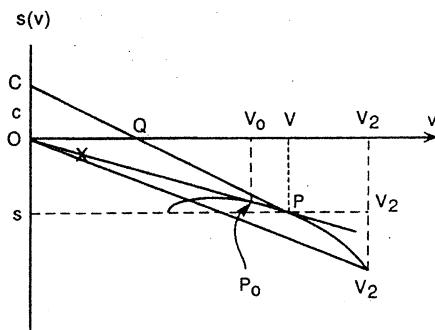

Figure 2

For each speed, e.g. V_2, the $\angle V_2 OP_2$ represents the angle α of Figure 1. It is then clear that the best achievable glide slope (minimum value for α) corresponds to the tangent to the "polar" from O, and P_0 represents that point at speed V_0.

An illustration of using the polar to maximise cross-country speed on the assumption that the general airmass is neither rising nor descending and there is no wind is as follows.

Suppose the glider is flown at speed v for a time t. It loses height at the sinking speed $s(v)$ and consequently the height loss is $H = s(v)t$. Assuming that this height is regained in a thermal at a rate c, the time taken to return to the original level is H/c. During this combined operation the distance covered over the ground is vt. Thus the cross-country speed is $X = vt/(t + H/c) = v/(1 + s(v)/c) = vc/(c + s(v))$.

Note that this result may be written as $X/c = v/(c + s(v)) = \tan \angle QCO$ (see Figure 2).

To maximise the cross-country speed one differentiates with respect to v and finds that the maximum occurs when $s'(v) = [c + s(v)]/v$. A very simple interpretation of this result is that to obtain the optimum speed to fly, one draws the tangent to the polar from the point C on the vertical axis (where $OC = c$). If P is the point of tangency its abscissa represents the optimum speed to fly, its ordinate represents the corresponding sinking speed and OQ represents the optimum cross-country speed.

An interesting conclusion is that to achieve the maximum cross-country speed, the pilot must fly *faster* than the best L/D speed, perhaps against intuition.

The above analysis needs modification to take account of head- or tail-winds and to cover the cases when the air is sinking or ascending; also to allow for a different all-up weight which means a change in wing-loading.

If the wing loading is altered one can calculate a revised polar to sufficient accuracy by expanding or contracting the whole polar linearly from the origin. It can be proved that to carry out this operation, one uses a scaling factor equal to the square root of the ratio of the new weight to the old weight as shown in Figure 3.
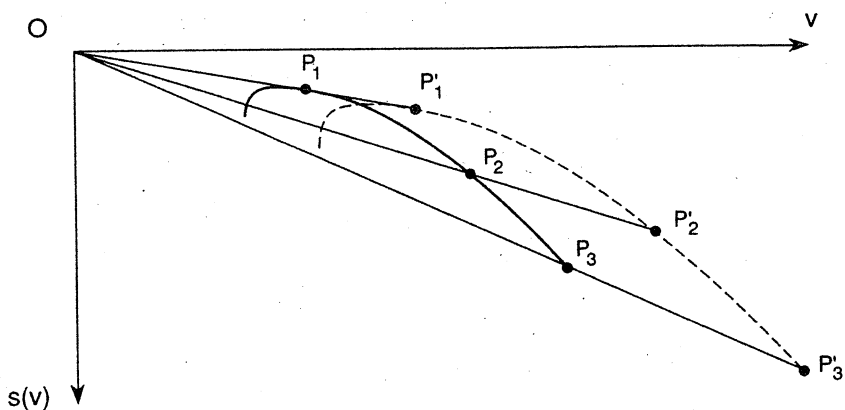

Figure 3

If the sailplane is to achieve a very high altitude, one needs to take into consideration the decrease in air density and pressure. The aircraft must fly faster in order to produce the requisite aerodynamic force. It will then also sink faster. These effects can be shown on the polar by simply changing the coordinates of each point in the ratio of the standard air density to the actual air density. But since the air-speed indicator is also affected in the same way, the pilot does not need consciously to take any action but should realise that the actual speed is greater than that indicated.

To get the best glide-ratio in relation to the ground allowing for a head- or a tail-wind, assuming zero vertical air motion, one treats the polar of Figure 2 as valid at all times relative to the (sea-level) air mass. But a head- or tail-wind will slow the airplane down or speed it up by an amount equal to the magnitude of the wind velocity. This means the polar must be displaced as implied in Figure 4. O_t represents the origin for a tailwind, v the speed and $s(v)$ the rate of sink, as previously. O_h would be the origin for a headwind. The pilot, though, would still fly in relation to the surrounding air, i.e. "his" polar would be that with origin at O.
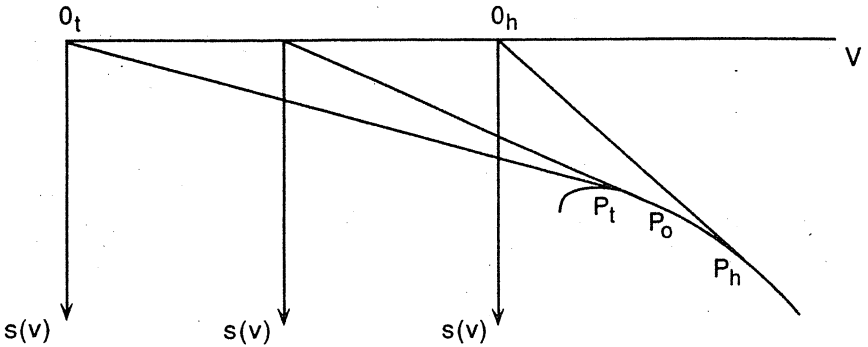
Figure 4

Since the tangent analysis is still valid for these modified conditions, it can be graphically seen that when there is a tailwind component, the speed to fly would be slower than for no wind conditions (P_t *versus* P_0), while the opposite is true for a headwind (P_h).

In a similar way for a rising or sinking general airmass (as opposed to the rising air in a thermal), the sailplane polar must be displaced upwards or downwards by the appropriate amount and the best glide speeds re-determined. Figure 5 illustrates the procedure by showing origin O_{si} for the case of a sinking airmass and origin O_r for a rising airmass.

Figure 5

For an airmass sinking at 1 m/s the best glide speed indicated is 120 km/h, which gives a glide ratio only slightly better than 16:1.

On the other hand, if the airmass is rising at 0.58 m/s the tangent from O is horizontal and indicates that at a speed of 73 km/h it is possible to have zero altitude loss.
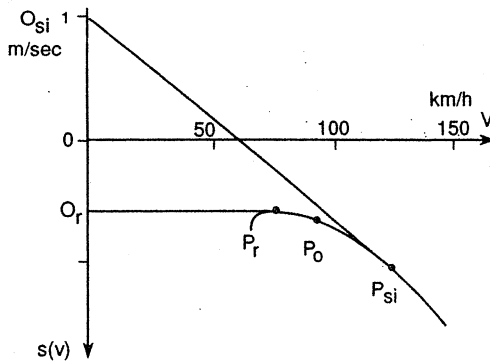
It is again graphically evident that in general, if the glider is flying in sinking air conditions, the optimum result will be achieved by flying faster than in still air (P_{si} *versus* P_o), and the opposite is true for rising air conditions (P_r). This also contradicts intuition to some extent.

From these examples it is clear that every vertical airmass movement provides a correct speed-to-fly and these can be displayed by means of a movable ring that can be rotated about the variometer (rate of climb or descent meter).

Instead of the geometric approach used above by means of tangents to the polar of Figure 2, it is possible to obtain a reasonably accurate formula for the rate of sink, s, as a quadratic function of the speed v.

Assume $s(v) = av^2 + bv + c$. Appropriate values for a, b and c can be calculated by inserting in the equation values of $s(v)$, say s_1, s_2, s_3 corresponding to the "best-glide" speed, 180 km/h, and an intermediate speed.

It may be shown that these lead to

$$a = \frac{(v_2 - v_3)(s_1 - s_3) + (v_3 - v_1)(s_2 - s_3)}{v_1^2(v_2 - v_3) + v_2^2(v_3 - v_1) + v_3^2(v_1 - v_2)}$$

$$b = \frac{s_2 - s_3 - a(v_2^2 - v_3^2)}{v_2 - v_3}$$

$$c = s_3 - av_3^2 - bv_3.$$

Such expressions enable on-board computers to provide answers to many of the questions that a competitive glider pilot wishes to know, most importantly, the optimum speed-to-fly conditions prevailing *at that time*.

The first person to utilise such an analysis was Dr Paul MacCready, who won a world gliding championship and also became the designer of the man-powered aircraft that crossed the English channel from the United Kingdom to France.

The glider pilot does not necessarily have to know or understand this analysis, which can be complicated much further, as more and more considerations are taken into account. In fact, most of them don't. However, the conclusions drawn from this mathematical/ geometric analysis are simple to understand and follow (with the aid of on-board instruments), and without doubt determine the flying patterns and the decision-making processes of glider pilots throughout the world.

* * * * *

CONTINUITY

Mark Kisin, Princeton University, U.S.A.

What do we mean by a continuous function? Most people would agree that the function in Figure 1 is continuous, but that the function in Figure 2 is not. It seems that the "problem" with the second function occurs at 1/2, where the function "jumps".

What precisely do we mean by "jumps"? To answer this question look at Figure 3, and the graph $y = f(x)$ marked there. Suppose that $f(1/2) = 1$, as shown in Figure 3. Now imagine that the variable x approaches 1/2 from the right as shown in Figure 3. It is clear from the diagram that $f(x)$ will never exceed 1, while x is greater than 1/2 (i.e. approaching 1/2 from the right).

What is the situation for a continuous function?

In Figure 4, $y = g(x)$ is shown, and $g(1/2) = 1$. It is clear that as x approaches 1/2 from the right, $g(x)$ gets as close to 1 as we please. Similarly if x approaches 1/2 from the left, $g(x)$ gets as close to 1 as we please.

In summary, when $f(x)$ "jumped" at 1/2, there were always $f(x)$ values that were not close to $f(1/2)$, even if we restricted the x values to be as close to 1/2 as we liked.

When $g(x)$ was continuous (i.e. did not jump at 1/2), we could get $g(x)$ as close to $g(1/2)$ as we pleased by making x sufficiently close to 1/2.
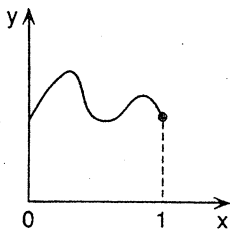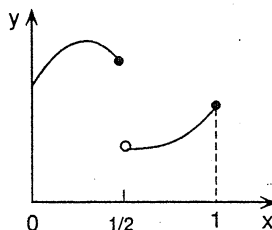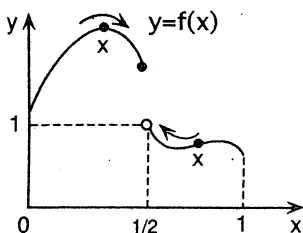

Figure 1
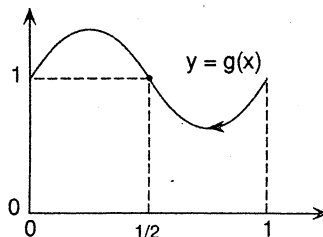

Figure 2

Figure 3

Figure 4

We use the above observations as motivation for the following definition: A function f is continuous at x_0 if $f(x)$ is as close to $f(x_0)$ as we like for all values of x that are sufficiently close to x_0.

So if we pick any (very small) number $\varepsilon > 0$, then $f(x)$ will always be within ε of $f(x_0)$ (i.e. $|f(x) - f(x_0)| < \varepsilon$), provided that x is sufficiently close to x_0 (i.e. $|x - x_0| < \delta$, for some positive number δ). It should be noted that we may require x to be very close to x_0 (i.e. δ may be very small) but once $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$.

Putting these ε's and δ's back into our definition, we get the following: A function f is continuous at x_0 if for any $\varepsilon > 0$ there is some $\delta > 0$, so that if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < e$.

Note that above we assumed that f was defined at x_0, and also "around" x_0 (i.e. in some neighbourhood of x_0). In what follows we will assume that all the functions with which we are working are defined on some sufficiently large interval (a, b) (open at both ends), and we will be working inside this interval. The reader may take the point of view that if we write down an expression like $f(x_0)$ or $f(x)$ or $f(1/2)$ then we are assuming that f is defined at x_0 or x or $1/2$ respectively.

We should have confidence that our definition will agree with our intuitive idea of a continuous function because we have already checked that according to this definition f (in Figure 3) is discontinuous (not continuous) at $1/2$, and that g (in Figure 4) is continuous at $1/2$. (You may wish to recheck this at this stage.)

Many "everyday" functions like x^2, $x + 1$, $\sin x$, e^x, etc. are continuous at all points on the real number line. So far we have looked only at functions which are continuous either everywhere or are discontinuous only at one point. Using our definition of continuity we are now in a position to look at some "stranger" functions.

Example 1

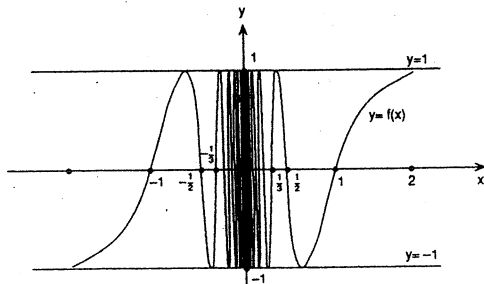$$f(x) = \begin{cases} \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Figure 5

As x approaches 0, $1/x$ gets larger, and does so increasingly quickly, so as x approaches 0, $\sin(1/x)$ oscillates more and more frequently. So inside any interval $(-\delta, \delta)$ around 0 there will be values of x for which $f(x) = 1$ and others for which $f(x) = -1$. This means that if we take ε to be, say, $1/2$, then no matter how small we take $\delta > 0$, there will always be values of x such that

i) $\qquad |x - 0| = |x| < \delta$, but $\qquad$ ii) $\quad |f(x) - f(0)| = |f(x)| = 1 > \varepsilon.$

So using our definition of continuity, we see that $f(x)$ is not continuous at 0, even though it does not appear to "jump" at 0.

Example 2

$$f(x) = \begin{cases} 1 & |x| \geq 1 \\ 1/2 & 1/2 \leq |x| < 1 \\ 1/4 & 1/4 \leq |x| < 1/2 \\ \vdots \\ 2^{-n} & 2^{-n} \leq |x| < 2^{-n+1} \end{cases}$$

$f(0) = 0$
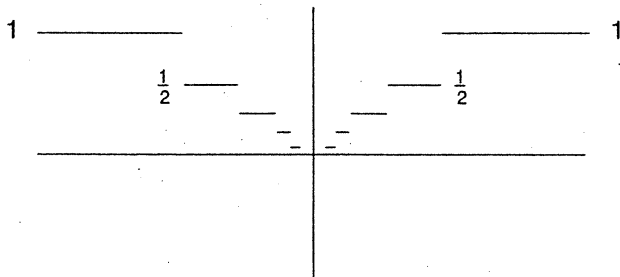
Figure 6

This function is continuous at 0, since provided we take x sufficiently close to 0, $f(x)$ will be as close to 0 as we like.

To check this strictly against our definition of continuity, choose $\varepsilon > 0$. Then $2^{-n} < \varepsilon \leq 2^{-n+1}$ for some positive integer n. Since from the definition of f, if $|x - 0| = |x| < 2^{-n+1}$, then $|f(x) - f(0)| = |f(x)| < 2^{-n}$, we can put $\delta = 2^{-n+1}$, and see that if $|x - 0| < 2^{-n+1} = \delta$, then $|f(x) - f(0)| < 2^{-n} < \varepsilon$. So f is continuous at 0.

Example 3.

$$f(x) = \begin{cases} x & \text{if } x \text{ is rational} \\ 0 & \text{otherwise} \end{cases}$$

This function is not continuous anywhere except at 0. We leave it to the reader to check this, using the definition of continuity given above.

When we look at the above three examples, we see that functions can behave in very surprising ways, so it is not sufficient to rely just on our intuition when dealing with notions of continuity.

* * * * *

LETTER TO THE EDITOR

On p. 127 of the August 1992 edition of *Function* there is a brief but very interesting report on Armstrong numbers. This is an area that I have given some attention to recently and I enclose my search for such numbers up to and including $m = 127$.

You will see that there are 22 entries for $3 \leq m \leq 9$ (not the 26 mentioned in the article). I wonder if I have missed some or whether the source article in *Pythagoras* included some spurious entries.

I also wonder if Armstrong numbers are widely known or referred to by this designation. I recently came across a delightful little book: *The Penguin Dictionary of Curious and Interesting Numbers* by David Wells (1988). This contains several such numbers, but there is no reference to "Armstrong". In fact, on p. 201, A_{39} is described as "the longest known pluperfect digital invariant in base ten". Possibly there are other references to Armstrong numbers but under a different title.

Order (m)	Armstrong numbers of order m
1.	All 1-digit numbers
2.	None
3.	153 ; 370 ; 371 ; 407
4.	1 634 ; 8 208 ; 9 474
5.	54 748 ; 92 727 ; 93 084
6.	548 834
7.	1 741 725 ; 4 210 818 ; 9 800 817 ; 9 926 315
8.	24 678 050 ; 24 678 051 ; 88 593 477
9.	146 511 208 ; 472 335 975 ; 534 494 836 ; 912 985 153
10.	4 679 307 774
11.	32 164 049 650 ; 32 164 049 651 ; 40 028 394 225 ; 42 678 290 603 44 708 635 679 ; 49 388 550 606 ; 82 693 916 578 ; 94 204 591 914
12.	None
13.	None
14.	28 116 440 335 967
15.	None
16.	4 338 281 769 391 370 ; 4 338 281 769 391 371
17.	21 897 142 587 612 075 ; 35 641 594 208 964 132 35 875 699 062 250 035

S.I.B. Ayeni
12/46 Arthur St., South Yarra

HISTORY OF MATHEMATICS

EDITOR: M.A.B. DEAKIN

"Unreasonable" Numbers

The fact that our word "ratio" (now understood as a quotient of two numbers) and our word "reason" are linguistically related has a long history. The Latin word for "reason" in fact is *ratio* and this is where we get our modern English word "rational", meaning "reasonable" in everyday speech.

But in Mathematics, we use the word "rational" in a technical sense, a number being rational if it is expressible as a fraction formed in dividing one integer by another.

This dichotomy of meaning clearly goes back to a time when the two concepts were seen as closely related. The concepts *to think* (to reason, to ratiocinate) and *to count* (to measure in ratios, see below) were seen as connected. Some authorities even suggest that this connection lies very deep indeed: pre-dating the early Indo-European language from which Latin, Greek and (later) English have developed. The connection is very evident in what we can piece together of the work of *Pythagoras of Samos* (after whom Pythagoras' Theorem is named).

Very little is really known of Pythagoras, but we may briefly summarise what is thought to be fact, sifting it from myth and legend. He is believed to have lived from about 560 B.C. to about 480 B.C., to have travelled widely in the ancient world (around the eastern end of the Mediterranean) and to have founded a school (or perhaps more accurately a form of religious order) in which Mathematics played a leading rôle.

His followers were either *akousmatikoi* (pupils who had to sit in silence and listen) or *mathematikoi* (learned ones who could engage in debate and discussion). The latter word is the source of our term "mathematics".

The Mathematics studied by the Pythagoreans is believed to have included a theory of musical harmony. A vibrating string will produce a note (*A*, say) if it vibrates so many (in this case 440) times per second. If it vibrates at twice that speed, we get the note (*A*) an octave higher, and if we vibrate the string at, in this example, 660 times per second, we get the note *E* lying between the two. If *A* and *E* are sounded together, we obtain a pleasing chord. Similarly with other proportions, provided those proportions may be expressed as simple integer ratios[†] like 2:1 or 3:2 (in the example above). Thus harmony is seen to depend on simple arithmetic ratios.

The Pythagoreans were almost certainly aware, too, of Pythagoras' Theorem, at least in special cases. Now if we take (say) a right-angled triangle whose short sides are each of length 1, then the hypotenuse will have a length of $\sqrt{2}$.

This isn't quite how the Pythagoreans are likely to have seen the matter, however. For us, lengths have to be expressed in terms of some sort of unit and we have standard units like the centimeter which we use when we wish to express a measurement of length. The Pythagoreans, however, seem to have had a rather more complicated notion – according

---

[†] In real life things are actually more complicated than this. See *Function, Vol.* 10, *Part 4, pp. 13-19.*

to which one not so much measured a single length, but rather compared lengths in pairs. The hope was that, like the musical harmonies, the lengths could be given as a pair of integers.

So, if, for example, we had lengths in a ratio 7:5, it meant that some unit length could be found that went seven times exactly into the first length and five times exactly into the second.[†] In such an event, we say the two lengths are "commensurable". That is to say, they may both be measured with a common unit.

Now, the side of a square and its diagonal (essentially the simple right-angled triangle outlined above) are in fact not commensurable. This is equivalent to the statement that $\sqrt{2}$ is *irrational* – i.e. not a rational number, or, in today's terms, not expressible in the form p/q, where p, q are integers.

This fact was certainly known in antiquity, and was very possibly discovered by the Pythagoreans. The first reliable account we have of the discovery comes from Aristotle who lived from 384 B.C. to 322 B.C., i.e. some 170 years after Pythagoras. In a book we now know as the *Prior Analytics*, Aristotle wrote

> " ... if the diagonal of a square is taken to be commensurable [with the side of the square], odd numbers are equal to even ones".

This is thought to be a reference to a proof of the irrationality of $\sqrt{2}$ which we write as follows in today's notation.

> Suppose $\sqrt{2} = p/q$, and that p/q is expressed in its lowest terms, so that p, q have no common factor. Then
>
> $$p^2 = 2q^2 \qquad (*)$$
>
> so that p^2 is an even number. It then follows[††] that p itself is even. Because p, q have no common factor, q is thus odd. But if p is even then $p = 2n$ (say) and Equation (*) now tells us that
>
> $$q^2 = 2n^2$$
>
> so that (as before with p), q must be even, when we have just found it to be odd!

> Thus $\sqrt{2}$ is irrational.

There is another proof of $\sqrt{2}$'s irrationality that I personally prefer. It proceeds as above and writes Equation (*) as

$$2 = p^2/q^2. \qquad (**)$$

[†] Quite remarkably, this ancient concept was revived by those who in the 1960's and 1970's marketed the "Cuisenaire rods", also known as "Numbers in Colour", for use in School Mathematics.

[††] This is an application of the theorem listed as IX.12 in Euclid's *Elements* and discussed in this column in *Vol. 17, Part 3*.

We now suppose p to be decomposed into its prime factors and similarly q. Because p/q is expressed in its lowest terms, no prime on the top row of (**) can cancel one on the bottom. Thus the integer 2 equals the fraction p^2/q^2, and this is possible only if $q = 1$, i.e. $\sqrt{2}$ is an integer, and clearly it isn't.

My reason for preferring this proof is that it very easily generalises. By the same argument we may prove the theorem:

If m is a positive integer which is not the n-th power of an integer then $m^{1/n}$ is irrational.

The other proof may also be so generalised, but most students find the path to the generalised version a little harder. I may also prefer the second proof because it was the one I first encountered (in Year 10, in Barnard & Child's *A New Algebra*, first published in 1912). Modern mathematical educators tend, however, to favour the first proof because it makes no reference to decomposition into primes.

It has been speculated that because the Pythagoreans were interested in ratios of integers, the discovery of irrational numbers was disturbing to them. Much has been written on this point; some of it scholarly speculation, some of it unadulterated garbage. The plain truth is that we know very little about who discovered irrational numbers or when the discovery took place.

Nor is it completely clear that $\sqrt{2}$ was the irrational number involved. A case has been made that it could have been τ (or φ, as in some notations), the Golden Ratio given by

$$\tau = \tfrac{1}{2}(1 + \sqrt{5}).$$

This irrational number is related to the geometry of the regular pentagon and the star-shaped pentagram formed from its diagonals.

An outline of a geometric proof of the irrationality of τ was given in *Function*, *Vol. 16, Part 5*, pp. 135-137. This simplifies an earlier pentagonal version that could (perhaps) have been discovered by the Pythagoreans, who showed interest in the properties of the pentagon and the pentagram.

Later in history, much later in fact, other sources of irrational numbers also emerged. If we consider the sequence

$$(1 + \tfrac{1}{n})^n$$

as n gets larger and larger, we find that the values tend ever closer to a number now known as e. We find

$$e = 2.71828\ 18284\ 59045\ ...\ .$$

e is the base for the system of natural logarithms that has many advantages over the common (base 10) logarithms, once so widely taught.

It is not difficult to prove that e is irrational, although it would take up too much space to give the details here.

e is in fact a basic constant of our number system, and so for that matter is π. (Although readers of *Function* will probably have encountered π only in a geometric context.) π too is irrational, but the proof takes some sophistication. In fact, no proof was known until 1761 when the German mathematician Lambert gave one. Later (1882)

another German, Lindemann, extended this to prove that the ancient Greek problem of "Squaring the circle" was impossible.[†]

It was not till 1947 that an "elementary" proof of π's irrationality was discovered. The American mathematician Ivan Niven was able to produce a proof using only (relatively) elementary concepts from the calculus. He rather generously acknowledged ideas from the 19th-Century French mathematician Hermite, but the proof is clearly Niven's.

We will not present it here – although elementary, it is beyond the scope of *Function*, and again it is also rather lengthy. It has been presented in various forms, but perhaps the best is one that begins with a prior theorem.

If r is rational and r ≠ 0, then cos r is irrational.

[Remember, when interpreting this result, that r is in *radians*, not in degrees!]

This is the theorem whose proof we omit, but the irrationality of π follows immediately from it, because if π were rational, then $\cos \pi$ would be irrational; but $\cos \pi = -1$.

The number system has other constants also besides e and π. Perhaps the next best known of these is Euler's constant (see *Function*, Vol. 17, Part 2, p. 37). Euler's constant is usually represented by the Greek letter γ and it is the limit as n gets larger and larger of

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - ln\ n$$

where $ln\ n$ is the natural (base e) logarithm of n. It is not known whether γ is irrational. This is one of the outstanding problems of Mathematics.

There are others too: $e + \pi$ and $e\pi$ are both thought to be irrational (as is γ), but no proof is known.

Nowadays, we build both rational and irrational numbers into a single system of "real numbers". Some, actually a very tiny minority, have the property that they may be expressed as exact ratios of integers. These are "rational" and we should now think of this only as representing this special mathematical property. To modern eyes, they are no more "reasonable" than (say) $\sqrt{2}$ is. Probably the Pythagoreans saw things differently, but Mathematics has progressed a long way since then.

The unsolved problem of (e.g.) $e + \pi$ is a technical one in Mathematics. Its solution would tell us whether this number is rational, not whether it was "reasonable"!

Further Reading: Much has been written on the subject, not all of it good. However, a *very* good (probably the best) general book on the History of Mathematics is Morris Kline's *Mathematical Thought from Ancient to Modern Times*. There is a lot (more than most readers will want) about the Pythagoreans in W.B. Knorr's *The Evolution of the Euclidean Elements*, and the ancient discovery of irrationals is discussed in J.N. Crossley's *The Emergence of Number*. The more modern material is covered in Ivan Niven's *Irrational Numbers*. The life of Pythagoras is discussed in the *Dictionary of Scientific Biography*.

[†] That is to say, it is not possible, by classical Euclidean ruler and compass constructions, to form a straight line equal in length to the circumference of a given circle.

COMPUTERS AND COMPUTING

EDITOR: CRISTINA VARSAVSKY

Getting The Message Through

Communication is an important aspect of our every-day life. If we look around us we find that telephones, radios, televisions and fax machines are everywhere. Their function is to send messages to and from all over the world. Similarly, in the computer world, huge amounts of information are sent from the computer memory to a disk or *vice versa*, between computers in a local area network, and within worldwide networks such as Internet. All this talk, either between people or computers, heavily relies on the communication channels used to transmit the information. No matter how sophisticated the technology used for transmission is, any channel is prone to error which could be introduced through a variety of means such as hardware failure or interference. Any disturbance in the channel can corrupt the information being transmitted, that is, the information received may not be what was sent.

This article is to give you an idea of the mathematics involved in the different existing techniques to protect against the errors that may occur in the transmission of information. We will not be concerned about the secrecy of the message, as in the article "A Security Matter" (*Function*, April 1993). In that article we introduced cryptology; the idea was to code a message such that only the receiver would be able to understand it, and even if the message was intercepted by the "enemy" it would be hard, if not impossible, to decipher it. The goal here is different: we just want *to transmit the message correctly*, there is no need to hide it from anyone.

Coding is basically about translating information from one form to another. A typical example is the set of postcodes we have to represent cities; for example, the postcode 3000 is the numerical equivalent to the city of Melbourne. These codes may be electronically scanned at the distribution centres to make the whole delivery system more efficient. Similarly we have codes for our telephone numbers which identify clusters of switching networks. In the retail industry, black vertical lines are used to identify products, making stock and accounting processes as well as our lives much easier. The Morse code once used in telegraphy is another example of translating information, using an alphabet constructed with dots, dashes and spaces. Blind people are certainly grateful to those who came up with the idea of coding books using the Braille code.

Coding, in the sense we will see it here, goes beyond that. It is important to translate the information to a more convenient form, but it is also important to get the message through. The coded information is transmitted and checked at the receiving end, which decides whether to accept the received information or correct it. In either case this is translated back to the original form or what the receiver (an electronic device) "thinks" was the original form.

The most used form for encoding is the binary form, where the information is translated into strings of zeros and ones. This could be done in many ways, depending on the type of information; if our information is, for example, in numerical or alphabetical form, the universal ASCII code is handy. For example, the ASCII code for the capital letter "A" is the decimal number 65 or its equivalent binary string 1000001. So we can match letters and symbols with strings of zeros and ones. For instance, the binary version of the word HELLO is

1001000100010110011001001101001111

As a message can become too long to be handled all at once, the most practical thing to do is to take blocks of shorter strings of a given length, say three or four bits at a time.

Assuming our information is already in binary form, the next step is to find ways of ensuring that each block or *word* is delivered correctly. Let us have a look at some of the ways this may be handled.

Repetition codes

Sometimes we use this method (especially teachers!) when we want to make sure that the message was received correctly: we repeat it several times. This same principle is applied in coding. In this case the message is decomposed into its individual bits, and each of them is repeated several times, say three. Thus 1 is encoded as 111 and 0 as 000. In our example, HELLO will be transmitted as

1110000000111000000000011100000000001110001111110000001111110000001110000000111111
0000001110000001111111111111

The decoding procedure is pretty simple: take groups of three bits, decode 111 to 1 and 000 to 0. This is all nice but as we said before, errors may occur in the transmission and each group of three bits may not entirely consist of only ones or zeros. The receiver may get strings like 101 or 010. Since the occurrence of one error is more likely than the occurrence of two errors, we can proceed as follows: if the received word has more zeros than ones we decode it to 000 and if the received word has more ones than zeros then we assume that the transmitted word was 111.

For example, the following message was received, and we were told that a 3-repetition code was used:

101111001100000111

First we split the message in words of 3 bits: 101, 111, 001, 100, 000, 111. An error has occurred in the first, third and fourth words, and the message is corrected to

111111000000000111

and decoded as 110001. If the same message had been sent and received as

100111011100000111

we would have corrected it to

000111111000000111

and decoded as 011001, which is incorrect. This happened because in the first and third 3-bit strings two errors occurred. Our decoding system was able to detect them but not to correct them. If three errors occur, our system would not be able to detect it because 000 with three errors becomes 111 and *vice versa*, a legitimate word or *code word* as it is called in this field. In summary, this is a *double-error-detecting* and *single-error-correcting* code, because it can detect at most two errors, and can correct properly only when exactly one error occurs.

[**Exercise**: consider the 5-repetition code, that is, each bit is repeated 5 times. What is the maximum number of errors that this code can detect? How many can be corrected?]

Parity check codes

Another simple code consists in adding a check digit to the information string: 1 if the number of ones in the string is odd and 0 if it is even. Therefore each code word has an even number of ones. For example, if working with two-bit strings we encode

> 00 as 000
> 01 as 011
> 10 as 101
> 11 as 110

appending 1 or 0 to the original string, depending on the number of ones. Thus the set of code words is {000, 011, 101, 110}. The receiver accepts a word as correct if it belongs to this set, or in other words, if the number of ones is even. Otherwise, if the number of ones is odd, the receiver can be only sure that at least one error has occurred (it might have been three). If two errors occur in the transmission it would be impossible to detect them with this scheme because the number of ones would be even. What about correction? What if 111 is received? We know an error has occurred because we have an odd number of digits, but did it occur in the first, second or third digit? We cannot say. This is then a code that cannot correct errors, although it can detect one error.

This code is used, for example, when communicating with another computer via a modem. If you have ever used a modem you may well be aware that sometimes there is noise in the telephone line used as the transmission channel, resulting in the presence of funny symbols in the middle of the word, which make the message a bit hard to read. In this sort of transmission a 7-bit parity check is usually used. This means that seven of the eight bits are information bits[†], the eighth is the parity check. When only one error occurs the software controlling the transmission will detect it and ask for retransmission (and consequently slow down the communication). When the line is very noisy, more than one error can occur: if the number of errors is even the software cannot detect it and the word (actually letter or symbol) is decoded incorrectly.

Hamming Codes

Let us consider again repetition codes as described above. Using common sense we notice that the more times we repeat the bits 1 or 0, the more errors can be detected. This is because the two code words, a string of only ones and a string of only zeros, differ one from another in all the bits, making the "distance" between code words the largest. It is also true that a longer repetition code will decrease efficiency as we have to transmit the same information several times. More complicated codes can be constructed with better efficiency, while still having a set of code words sufficiently widely scattered that would permit some correction, if necessary. We will illustrate one of those, which belongs to the family of so-called Hamming codes.[††] It can be used to encode 4-bit strings into 7-bit strings, and is a 1-error correcting 2-error detecting code. The following table shows all the 4-bit strings and their encoded versions.

[†] This also explains why not all the symbols can be sent when using a 7-parity check: the largest number that can be written with seven digits is 1111111 which corresponds to the decimal 127. Only 127 ASCII characters can be handled in this way.

[††] Richard W. Hamming pioneered the study of error-detecting and error-correcting codes in 1950.

Word Code

```
0000  0000000
0001  0001111
0010  0010011
0100  0100110
1000  1000101
1100  1100011
1010  1010110
1001  1001010
0110  0110101
0101  0101001
0011  0011100
1110  1110000
1011  1011001
1101  1101100
0111  0111010
1111  1111111
```

This code then has 16 code words, one for each 4-bit string. What if an error occurs in the transmission? It could occur in any of the seven bits, so for each code word we have associated seven 7-bit strings which differ from the code words in only one position. Here is the list of the code words and their 1-bit variations.

Word	Code word		Variations of code words in only one bit					
0000	0000000	1000000	0100000	0010000	0001000	0000100	0000010	0000001
0001	0001111	1001111	0101111	0011111	0000111	0001011	0001101	0001110
0010	0010011	1010011	0110011	0000011	0011011	0010111	0010001	0010010
0100	0100110	1100110	0000110	0110110	0101110	0100010	0100100	0100111
1000	1000101	0000101	1100101	1010101	1001101	1000001	1000111	1000100
1100	1100011	0100011	1000011	1110011	1101011	1100111	1100001	1100010
1010	1010110	0010110	1110110	1000110	1011110	1010010	1010100	1010111
1001	1001010	0001010	1101010	1011010	1000010	1001110	1001000	1001011
0110	0110101	1110101	0010101	0100101	0111101	0110001	0110111	0110100
0101	0101001	1101001	0001001	0111001	0100001	0101101	0101011	0101000
0011	0011100	1011100	0111100	0001100	0010100	0011000	0011110	0011101
1110	1110000	0110000	1010000	1100000	1111000	1110100	1110010	1110001
1011	1011000	0011000	1111000	1001000	1010000	1011100	1011010	1011001
1101	1101101	0101101	1001101	1111101	1100101	1101001	1101111	1101100
0111	0111010	1111010	0011010	0101010	0110010	0111110	0111000	0111011
1111	1111111	0111111	1011111	1101111	1110111	1111011	1111101	1111110

Table 1

We obtained 128 7-bit strings. Are they all distinct? Believe me, they are! So by changing only one bit in each code word, we generated all possible 7-bit words. Then we can use the above table for decoding and correcting if at most one error occurs. Take for example the case where the word 1010000 is received. This is not a code word, but since 1010000 is in the same row as the code word 1110000 (fifth from bottom), we decide that one error occurred in the second bit and we decode it as 1110.

You may be wondering how that magic set of code words generated all 7-bit strings, or why we chose those code words in the first place. You may also question the efficiency of this code, as each received word has to be found in the table, and that may take a fair

bit of time. Although the theory behind this code is a bit complicated, a few things can be pointed out here.

Firstly, there is no need to have the code words listed. They are actually obtained through the following matrix H.

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

To encode, for instance, the word 0011, we consider that word as a 1×4 matrix and multiply it by H using the following addition rules[†]

$$1 + 1 = 0$$
$$0 + 1 = 1$$
$$1 + 0 = 1$$
$$0 + 0 = 0$$

In short, the sum of an even number of ones is 0, and the sum of an odd number of ones is 1. Now, multiplication of a 1×4 matrix by a 4×7 matrix will give a 1×7 matrix. [0011]×H is obtained by multiplying the word 0011 by each row of H, bit by bit:

$$\text{1st bit} = 0{\times}1 + 0{\times}0 + 1{\times}0 + 1{\times}0 = 0$$
$$\text{2nd bit} = 0{\times}0 + 0{\times}1 + 1{\times}0 + 1{\times}0 = 0$$
$$\text{3rd bit} = 0{\times}0 + 0{\times}0 + 1{\times}1 + 1{\times}0 = 1$$
$$\text{4th bit} = 0{\times}0 + 0{\times}0 + 1{\times}0 + 1{\times}1 = 1$$
$$\text{5th bit} = 0{\times}1 + 0{\times}1 + 1{\times}0 + 1{\times}1 = 1$$
$$\text{6th bit} = 0{\times}0 + 0{\times}1 + 1{\times}1 + 1{\times}1 = 0$$
$$\text{7th bit} = 0{\times}1 + 0{\times}0 + 1{\times}1 + 1{\times}1 = 0$$

In conclusion, the process of encoding a 4-bit word consists simply in multiplying that word by H.

Secondly, decoding is also done through a matrix which is related to H. Observe that the first four columns of H for the 4×4 identity matrix with ones on the diagonal and zeros elsewhere. Crossing out that identity matrix, we are left with

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

which we augment in the vertical sense with a 3×3 identity matrix, obtaining

---

[†] This is addition modulo 2 as described in the previous article titled "Modular Arithmetic keeps the Numbers Small" (*Function*, February 1993).

$$G = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Observe first that if we multiply any code word by G we obtain [000]. Try, for example, with 0001111:

$$0 \times 1 + 0 \times 1 + 0 \times 0 + 1 \times 1 + 1 \times 1 + 1 \times 0 + 1 \times 0 = 0.$$

You can check the rest. It is also true, but we will not show it here, that all the words that result in 000 when multiplying by G are code words.

Decoding is done as follows: multiply the received word by G, if the product gives [000] then the received word is a code word and we accept it. If it does not, an error has occurred and we need to find out in what position. Let us call the received word B. Since B is not a code word, $B \times G$ does not result in [000]. If one error occurred in the transmission, B differs in one bit with the transmitted word A we are trying to get back. Then we can write

$$A = B + C$$

where C is the error. Now C has a single 1 and the rest of the bits are 0. Then using the distributive law for matrix multiplication, we have

$$A \times G = (B + C) \times G = B \times G + C \times G.$$

Since A is a code word, $A \times G = 0$, and then $B \times G = -C \times G$. Remember that we are working with a particular arithmetic where $1 + 1 = 0$; consequently $1 = -1$. Hence

$$B \times G = C \times G.$$

Therefore C is a word with a single 1 such that when multiplying G gives the same result as $B \times G$. Now, there are not many words with a single 1, we can multiply G with each of them and see which is the one with that characteristic.

Say, for example, that the word 0000011 was received. Multiply [0000011] and G:

$$[0000011] \times G = [011].$$

Since the product is not [000], we need to find C. Let us set the table for all possible Cs:

C	$C \times G$
0000001	001
0000010	010
0000100	100
0001000	111
0010000	011
0100000	110
1000000	101

Table 2

You must have observed that the pattern is pretty simple: as a single 1 is present, each multiplication by C gives the row corresponding to the position of 1. Now, 011 comes from 0010000, therefore we decode 0000011 as 0000011 + 0010000 = 0010011 which corresponds to 0010.

[Exercise: Pick up any other word in Table 1, and decode it using this procedure.]

What if we receive the word 1001111? Performing the product

$$[1001111] \times G = [101]$$

we see that 1001111 is not a code word, but neither is 101 listed in the second column in Table 2. This means that more than one error occurred in the transmission. Trying out with the different 7-bit strings with two ones in it we find that $[0110000] \times G = [101]$, and also $[0001010] \times G = [101]$. This indicates that the error might have occurred in the second and third bits as well as in the fourth and sixth bits, but we cannot decide with certainty. Therefore this Hamming code is capable of correcting a word when only one error occurred, but if two errors occurred they will be detected but not corrected.

Different matrices with different dimensions will generate new codes. In any case, there is no need to construct a table like Table 1, coding and decoding will only involve multiplications by the appropriate matrices.

Coding theory is an interesting area with an increasing range of applications, and it is still in the making. Its foundation is given by algebraic structures that are beyond the level of this magazine, but I hope this simple introduction gives you at least some understanding of the existing techniques to get the message through.

PROBLEMS AND SOLUTIONS

NOTE ON PROBLEMS 14.4.7, 14.4.8, 14.4.9

These three problems came from the "very hard" section of the Kürschák competition (Budapest). We subsequently published a fourth problem from the same source (Problem 15.3.1) and this *was* solved. So, although no solutions are yet to hand, we live in hope that some may be forthcoming.

SOLUTION TO PROBLEMS 15.1.1 and 15.1.2

We begin with 15.1.2 which asked in how many ways may four distinct integers a, b, c, d be chosen from the set $\{1, 2, ..., n\}$ (where n is even) in such a way that $a + c = b + d$. Hans Lausch sent us the following solution.

Without loss of generality, let $a < b < d < c$. For each pair a, c the number of quadruples depends only on $c - a$.

$c - a$	# pairs (c, c)	# quadruples for each pair (a, c)
1	$n - 1$	0
2	$n - 2$	0
3	$n - 3$	1
4	$n - 4$	1
.	.	.
.	.	.
.	.	.
$n - 3$	3	$(n/2) - 2$
$n - 2$	2	$(n/2) - 2$
$n - 1$	1	$(n/2) - 1$

Hence the total number of quadruples is given by

$$(2n - 7)1 + (2n - 11)2 + ... + 5\left[\frac{n}{2} - 2\right] + \left[\frac{n}{2} - 1\right]$$

$$= \sum_{i=1}^{\frac{n}{2}-2} (2n - 3 - 4i)i + \left[\frac{n}{2} - 1\right]$$

$$= (2n - 3) \sum_{i=1}^{\frac{n}{2}-2} i - 4 \sum_{i=1}^{\frac{n}{2}-2} i^2 + \left[\frac{n}{2} - 1\right]$$

$$= (2n - 3)\frac{(\frac{n}{2}-2)(\frac{n}{2}-1)}{2} - \frac{4}{6}\left[\frac{n}{2} - 2\right]\left[\frac{n}{2} - 1\right](n - 3) + \left[\frac{n}{2} - 1\right]$$

(using standard summation formulae)

$$= \frac{n - 2}{24}\{3(2n - 3)(n - 4) - 4(n - 4)(n - 3) + 12\}$$

$$= \frac{n - 2}{24}(2n^2 - 5n)$$

$$= \frac{n(n - 2)(2n - 5)}{24} .$$

Problem 15.1.1 began with the set $\{1, 2, ..., n\}$ as in 15.1.2 and asked in how many ways four numbers could be chosen from this set in such a way as to form a rectangle in which a circle may be inscribed. See Figure 1.

We have $AE = AH$, etc. Let $a = DA$, $b = AB$, $c = BC$ and $d = CD$. Then readily we show that

$$a + c = b + d$$

and so the problem is the same as the previous one if n is even.

If n is odd then $n - 1$ is even. We thus have already $\frac{1}{24}(n-1)(n-3)$ ways (found by replacing n by $n-1$ in the answer to the previous problem. The last choice, n, adds new ways, and following the details of the previous problem, we begin with the knowledge that for these extra ways $c = n$.

We thus want triples (a, b, d) such that

$$a < b < d < n,$$

Figure 1

and counting as in the previous problem (we omit the details) we find $\frac{1}{4}(n - 1)(n - 3)$ such new ways. After some algebra, we find a total of
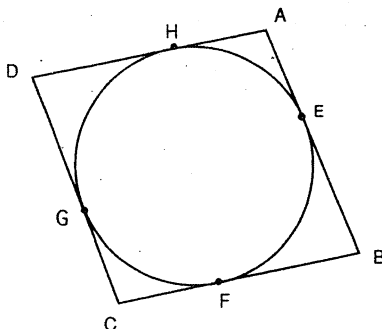
$$\frac{1}{24}[n(n-2)(2n-5) - 3]$$

ways.

This is equivalent to the answer required. [We note and belatedly apologise for a minor misprint in the answer we quoted for Problem 15.1.2.]

SOLUTION TO PROBLEM 15.2.2 (sent in by Keith Anker, Monash University).

Consult Figure 2. AD, BE and CF are the angle bisectors of $\triangle ABC$, with D lying on BC, etc.. We want to find all triangles ABC for which DF bisects $\angle BFC$ and to say what the angles of the triangle are if (also) $\angle AFC = 2\angle BAC$.
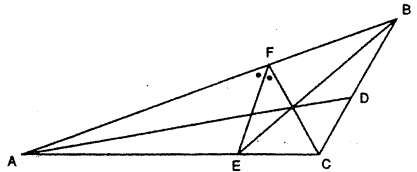
Figure 2

Let $AF = c_1$, $FB = c_2$, $BD = a_1$, etc.. around the triangle. Write $CF = f$. Now apply the sine rule to $\triangle BCF$.

$$\frac{f}{c_2} = \frac{\sin B}{\sin\frac{C}{2}} = \frac{\sin B}{\sin\frac{C}{2}} \cdot \frac{2 \cos \frac{C}{2}}{2 \cos \frac{C}{2}} = \frac{2 \sin B \, \cos \frac{C}{2}}{\sin C}$$

which by the sine rule (for $\triangle ABC$) is $(2b \cos \frac{C}{2})/c$.

Also, since CF bisects $\angle ACB$, $c_2/c_1 = a/b$. This means that

$$c_2/c = a/(a + b).$$

But now

$$f = \frac{f}{c_2} \cdot \frac{c_2}{c} \cdot \frac{c}{1} = \frac{2b \cos \frac{C}{2}}{c} \cdot \frac{a}{a+b} \cdot \frac{c}{1} = \frac{2ab \cos \frac{C}{2}}{a+b} \,.$$

Thus

$$\cos \frac{C}{2} = \frac{f(a+b)}{2ab} \,.$$

Now apply the cosine rule to $\triangle ABC$:

$$\cos C = \frac{a^2+b^2-c^2}{2ab} \,.$$

But also

$$\cos C = 2 \cos^2 \frac{C}{2} - 1.$$

We thus find

$$\frac{f^2(a+b)^2}{2a^2b^2} = \frac{a^2+b^2-c^2}{2ab} - 1$$

which simplifies to

$$f^2 = 4abs(s-c)/(a+b)^2, \tag{1}$$

where $s = \frac{1}{2}(a + b + c)$.

If now we assume that FD bisects $\angle BFC$, then we also have

$$\frac{f}{c_2} = \frac{a_2}{a_1} = \frac{b}{c},$$

from which

$$f = \frac{b}{c} \cdot \frac{ac}{a+b} = \frac{ab}{a+b}. \tag{2}$$

We now have in Equations (1), (2), two different formulae for f. Equating them, we find (after some algebra) that

$$c^2 = a^2 + ab + b^2.$$

By the cosine rule, this gives $C = 120°$.

Another way to specify the triangle (though the details are omitted) is to show that in this case FE bisects $\angle AFC$.

To solve the second part is relatively easy once we know the angle C. We find that under the additional hypothesis $A = 20°$, $B = 40°$.

SOLUTION TO PROBLEM 15.4.3 (from Ken Evans, Dromana).

How many triangles are there in Figure 3 (at right)?

Begin with the largest and end with the smallest and thus see that the number is

$$1 + 2 + 3 + ... + 8 = 36.$$
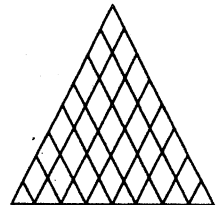
Figure 3

SOLUTION TO PROBLEM 15.4.4 (also from Ken Evans).

The problem read:
You do not know how many there are.
Putting groups of three together, two are left;
Five and five together, three are left;
Grouping seven and seven, two are left again.
Tell how many there are.
To solve the problem, let x be the number of objects. Then

$$x = 3m + 2 \qquad (1)$$
$$x = 5n + 3 \qquad (2)$$
$$x = 7p + 2 \qquad (3)$$

for some natural numbers m, n, p.

Now from Equations (1), (2)

$$3m + 2 = 5n + 3$$

so $3m - 5n = 1$.

This is an example of a so-called Diophantine equation and it is known that if a, b are relatively prime, the Diophantine Equation $am + bn = 1$ has the general solution $m = M - bt$, $n = N + at$ for integral t and M, N some particular solution.

Our equation has the obvious solution $m = 2$, $n = 1$, so in general

$$m = 2 + 5t, \ n = 1 + 3t$$

where t is a non-negative integer (as $m > 0$, $n > 0$).

Similarly, from Equation (2), (3)

$$n = 4 + 7u, \ p = 3 + 5u$$

for a non-negative integer u.

Now

$$n = 1 + 3t = 4 + 7u$$

and so another Diophantine equation results. This has the solution

$$t = 8 + 7v, \ u = 3 + 3v \quad (v \geq -1)$$

for some integer v. We may now put p in terms of v and so determine x. We find

$$x = 128 + 105v.$$

So we have, if $v = -1$, $x = 23$

if $v = 0$, $\quad x = 128$

if $v = 1$, $\quad$ x = 233, etc.

SOLUTION TO PROBLEM 15.4.6 (also from Ken Evans).

The problem asked for those odd positive integers n which are such that

$$f(x) = (x+1)^n - x^n - 1$$

and
$$h(x) = (x+1)^{n-1} - x^{n-1}$$

have common zeros which include the roots of $x^2 + x + 1 = 0$.

These roots are $\frac{1}{2}(-1 \pm i\sqrt{3})$, where $i^2 = -1$, and since $(x-1)(x^2+x+1) = x^3 - 1$ they both satisfy the equation $x^3 = 1$.

The roots of $x^3 - x + 1 = 0$ similarly satisfy $x^3 = -1$.

Now every positive integer n may be expressed as either $n = 3m$, $n = 3m+1$ or $n = 3m+2$, where m is a non-negative integer. As we are told that n is odd, we need only consider the cases:

$$n = 3m \qquad \text{where } m \text{ is odd} \qquad (1)$$
$$n = 3m + 1 \qquad \text{where } m \text{ is even} \qquad (2)$$
and $\qquad n = 3m + 2 \qquad \text{where } m \text{ is odd.} \qquad (3)$

In Case (1), $f\left[\tfrac{1}{2}(-1 \pm i\sqrt{3})\right] = (-1)^m - (1)^m - 1 = -3 \neq 0$.

In Case (2), $f\left[\tfrac{1}{2}(-1 \pm i\sqrt{3})\right] = \tfrac{1}{2}(1 \pm i\sqrt{3}) - \tfrac{1}{2}(-1 \pm i\sqrt{3}) - 1 = 0$.

In Case (3), $f\left[\tfrac{1}{2}(-1 \pm i\sqrt{3})\right] = -\tfrac{1}{2}(-1 \pm i\sqrt{3}) - \tfrac{1}{2}(-1 \pm i\sqrt{3}) - 1 = 0$.

We thus reject Case (1).

In Case (2), $h\left[\tfrac{1}{2}(-1 \pm i\sqrt{3})\right] = (-1)^m - 1^m = 0$.

In Case (3), $h\left[\tfrac{1}{2}(-1 \pm i\sqrt{3})\right] = -\tfrac{1}{2}(1 \pm i\sqrt{3}) - \tfrac{1}{2}(1 \pm i\sqrt{3}) \neq 0$.

Thus only Case (2) remains and

$$n = 3m + 1 = 3(2p) + 1 \quad \text{(say)} = 6p + 1.$$

SOLUTION TO PROBLEM 15.4.7 (also from Ken Evans).

We are to prove that the polynomial

$$f_n(x) = x^n\sin \alpha - x \sin(n\alpha) + \sin[(n-1)\alpha]$$

is exactly divisible by

$$h(x) = x^2 - 2x \cos \alpha + 1$$

where α is real and n an integer greater than 1.

For $\cos \alpha + i \sin \alpha$, write $cis\ \alpha$. Then

$$h(x) = (x - cis\ \alpha)(x - cis(-\alpha)).$$

But $f_n(cis\ \alpha) = 0$ by the use of de Moivre's theorem

$$(cis\ \alpha)^n = cis(n\alpha)$$

and so by the Remainder Theorem $x - cis\ \alpha$ divides $f_n(x)$. Similarly $x - cis(-\alpha)$ divides $f_n(x)$ and so the result is proved.

INDEX TO VOLUME 17

INDEX TO VOLUME 17

* * * * *