

Setting the Calculus Straight

David Tall

The study of the calculus occupies the major part of the pure mathematics A-level. It is full of formulae and procedures for finding derivatives, integrals, solutions of differential equations, and all manner of other horrendous things. But the underlying ideas are extremely simple – so simple that it is only recently that we have realised how simple they are. In this article David Tall shows how new methods of visualizing the ideas of the calculus can set the ideas straight – in a way that makes them easy to imagine.

Introduction

As a reader of *Mathematics Review* at the time when this issue is published it is highly likely that you are just embarking on an extended study of the calculus, or are half-way through the course. Wherever you are, I hope that the ideas presented here illuminate your studies and give them a meaning that allows you to see how the whole theory fits together.

The calculus is traditionally considered to consist of two things: differentiation and integration. To these we must add a third: solving differential equations. Differentiation is about the rate at which things change, integration is about how things grow, and differential equations relate the two together in a way which will soon be made apparent. Let us forget these rather intimidating titles for a moment and get down to fundamental ideas: rate of change and growth. We will later find that the fundamental ideas of the calculus are very simple indeed.

Rate of change

Suppose I am driving a car away from home and I record the distance gone as I drive to give the graph in figure 1. This graph tells a story. First I start in town (A) and travel three or four miles before realising that I have forgotten some papers (at B), so I go back, stop at my home (C), find what I need and set out again (D). Again I travel relatively slowly, until I reach the motorway (E), then I move faster until I hit a contraflow (F) when I slow down. A little later (from G to H) on the graph I stop at a motorway service area for a coffee, then set out at a more reasonable motorway speed. I can tell how far I have gone. In the first hour I travelled about 27 miles away (although I was stupid enough to have to retrace my steps for part of the time). In the second hour I travelled about fifty three miles, including the stop at the service area. Between B and C, if I measure my distance *away* from home, I was actually travelling

in reverse, a *negative* distance, though I promise you I was not stupid enough to do it in reverse gear!

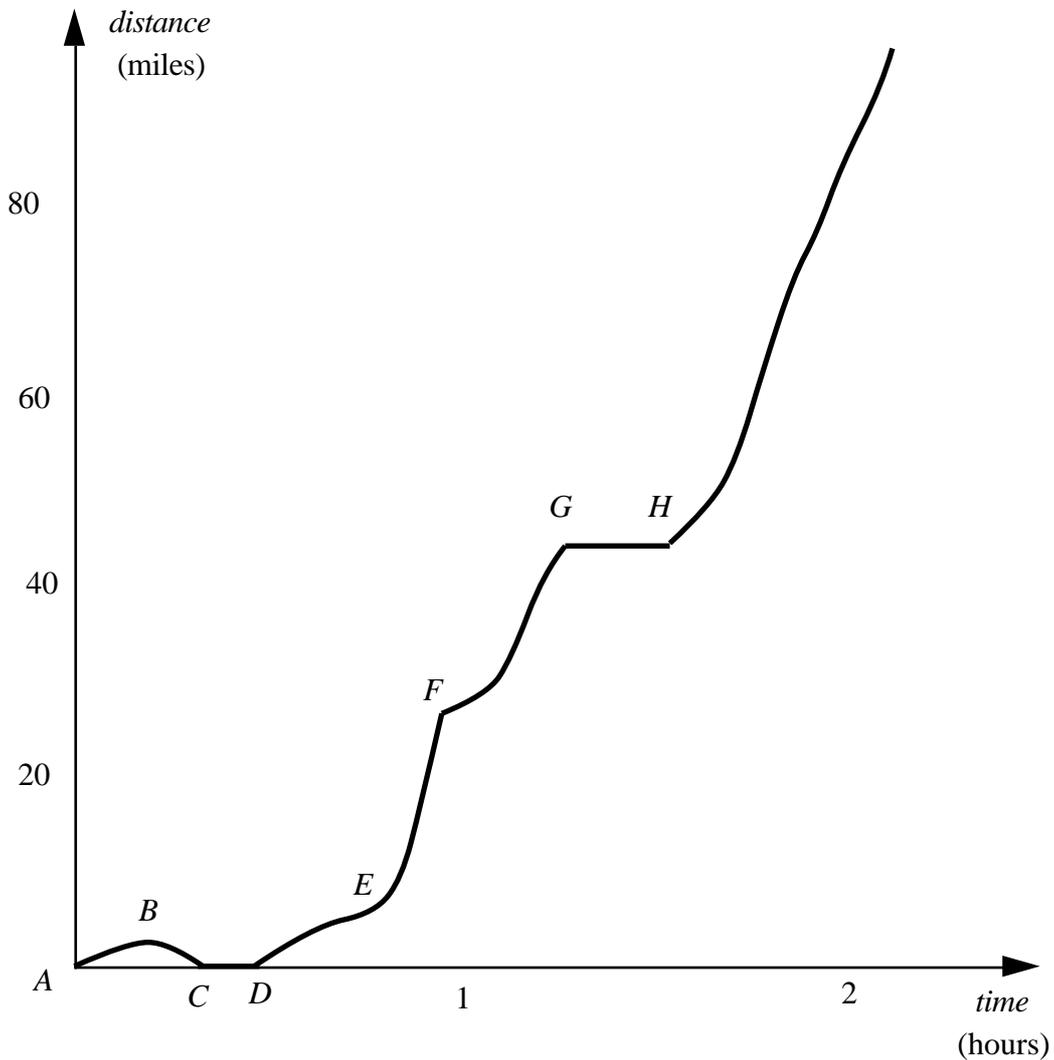


Figure 1 : taking a journey in a car

The question is: how what was my average velocity over any given interval of time. For the first hour the average was 27 miles an hour (measured in forward direction), and in the second hour it was 53 miles per hour. Whilst at the service station my speed was zero, but after leaving the station I travelled about 35 miles in the next half an hour at around the legal maximum of 70 miles per hour. If I want to work out my speed over any time interval, I just calculate the ratio:

$$\frac{\text{change in distance}}{\text{change in time}} .$$

For instance, in the second hour, my average speed was

$$\frac{QR}{PQ}$$

(figure 2.)

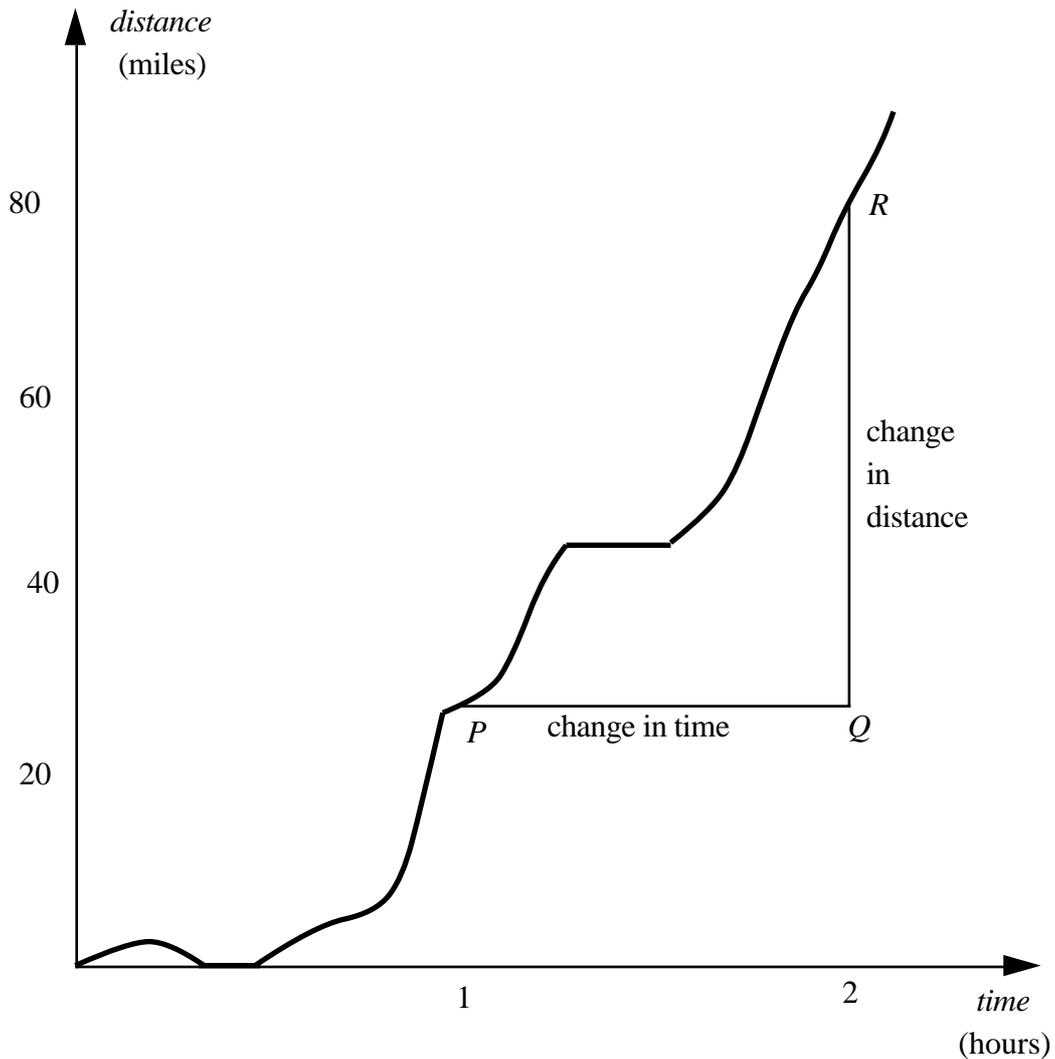


Figure 2 : average speed over the last hour

The problem with working out speed in such a case is that it is varying all the time. If the graph were a straight line, then the speed would be constant. This is the saving grace of the calculus.

Suppose we have a graph like the graph of $y=x^2$. If part of the graph is highly magnified, then it will look almost straight (figure 3). This means that the rate of change of y with respect to x , which is found by measuring

$$\frac{y\text{-change}}{x\text{-change}},$$

will not vary much over the magnified part.

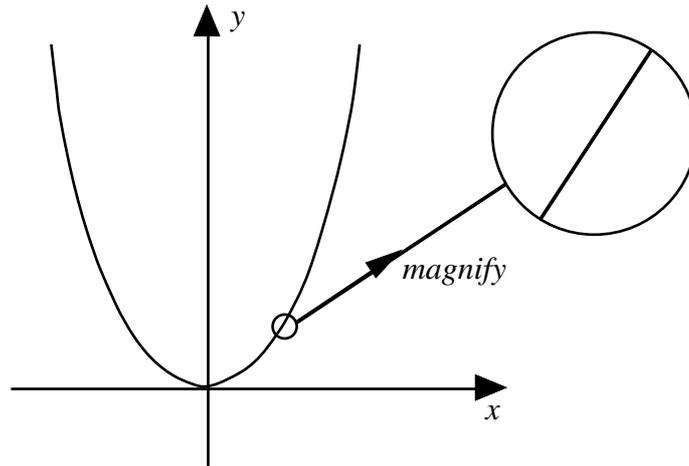


Figure 3 : Magnifying the graph of $y=x^2$

It is this idea that such a curve magnifies to look straight when a small portion is placed under a microscope that makes the calculus possible.

It means that we can scan our eyes along the graph and see various parts of it changing in gradient. Just by *looking*, we can see the gradient to the left being negative and getting less and less steep until the gradient is zero at the origin, then increasing for positive values.

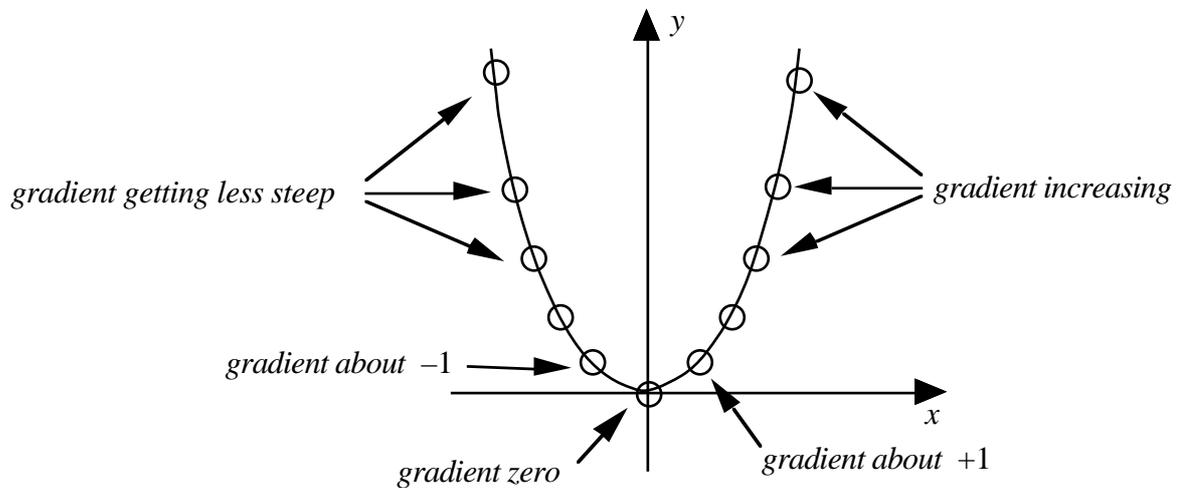


Figure 4: The changing gradient of $y=x^2$

At the point x , we can calculate the rate of change to a nearby point $x+h$, by simply noting that the x -change is h and the y -change is $(x+h)^2-x^2$, so the average rate of change is

$$\frac{\text{y-change}}{\text{x-change}} = \frac{(x+h)^2-x^2}{h}$$

$$= \frac{x^2+2xh+h^2-x^2}{h}$$

$$= \frac{2xh+h^2}{h}$$

and, provided that $h \neq 0$, the gradient is

$$2x+h.$$

It would, of course, be insane to take h equal to zero, as the two points x and $x+h$ would then be the same. But, provided we take h to be very small, the graph is then almost a straight line and the gradient is virtually equal to $2x$. This can be seen in figure 4, where the gradient at $x=-1$ is approximately -2 , at $x=0.5$, the gradient is approximately 1 , and at $x=1$ it is approximately 2 . The gradient simply increases uniformly from large negative numbers through zero and through increasing positive numbers.

The big difference for today's generation is that it is possible to calculate the approximate numerical gradient on a computer or hand calculator and draw the graph of the gradient. For instance, figure 5 shows the gradient of $\sin x$ (with angles measured in radians) and superimposed is a chord through two close points on the graph. The software is plotting the numerical value of the gradient of the chord as a point and then moving on, leaving a trace representing the gradient behind. You can see a few of these calculations yourself by casting your eye along the curve. At the origin, for example, the gradient of the sine curve is 1 . It then decreases, remaining positive, until the sine curve reaches its maximum value when the gradient is zero and then turns negative. This will help confirm to you that the computer is truly drawing the gradient curve of $\sin x$. Anyone with a mite of knowledge is likely to suggest that the dotted gradient curve looks like $\cos x$.

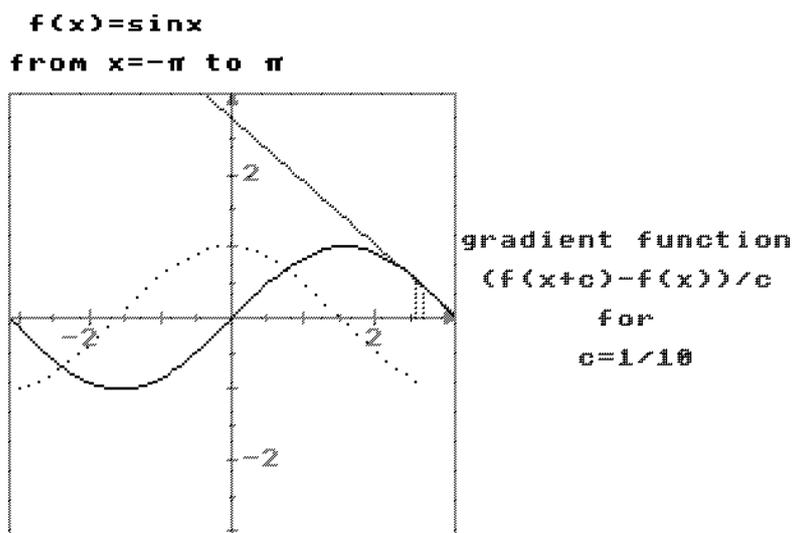


Figure 5: The gradient curve for $\sin x$ looks like $\cos x$

“Differentiation” is just the calculation of the symbolism for the gradient function. The latter is called the *derivative* of the original function. If $y=f(x)$, the derivative is denoted by $f'(x)$. We have already guessed that the derivative of $f(x)=x^2$ is $f'(x)=2x$, and of $f(x)=\sin x$ is $f'(x)=\cos x$. It is not difficult to guess a few more simple cases either. But once you have the idea that the derivative is just the gradient function, this sets you on the path to understanding what is happening in differentiation. The derivative is determined visually by looking along the graph and seeing the changing gradient. This will not get you far if you wish to work out the gradient of more complicated functions. The techniques of differentiation are developed to enable you to work out the gradients of more complicated functions from simpler ones that you already know and this will take you a few months of study and practice.

The Leibniz Notation

Leibniz was one of the co-inventors of the calculus in the seventeenth century, along with Sir Isaac Newton. Leibniz had his own notation for the calculus and wrote $\frac{dy}{dx}$ for the gradient. He actually thought of dx as representing the change in x and dy as being the corresponding change up to the tangent at x (figure 6).

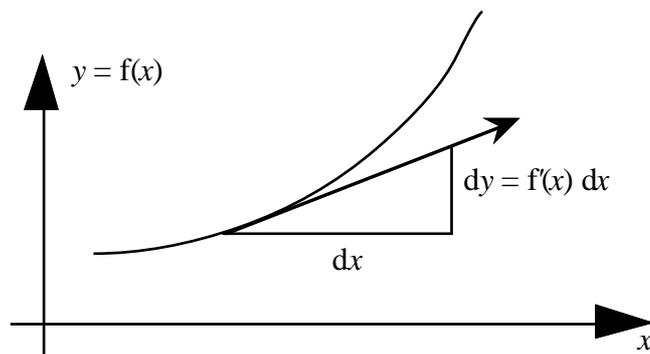


figure 6: dx and dy as components of the tangent vector

The two notations share the relationship

$$f'(x) = \frac{dy}{dx}$$

and, by thinking of dx and dy as being the components of the tangent vector, we see that they are related by the equation

$$dy = f'(x) dx.$$

Undoing differentiation

The reverse of differentiation is knowing the gradient $\frac{dy}{dx}$, to find the original function $y=f(x)$. The idea is again extremely simple. Because the graph is differentiable, in the small a tiny portion approximates to a straight line segment and, if we know the gradient of the segment, then we can draw it. For instance, if we have

$$\frac{dy}{dx} = \frac{1}{2} y.$$

which tells us that the gradient of the original function through (x,y) is $\frac{1}{2} y$, then we can draw a line segment of gradient $\frac{1}{2} y$. By visually sticking such line segments end to end we can build up an approximate solution curve. Figure 7 shows the *Solution Sketcher* software written for the new 16-19 A-level with a line segment drawn through $x=1.5, y=2$, where the gradient is therefore $\frac{1}{2} y=1$.

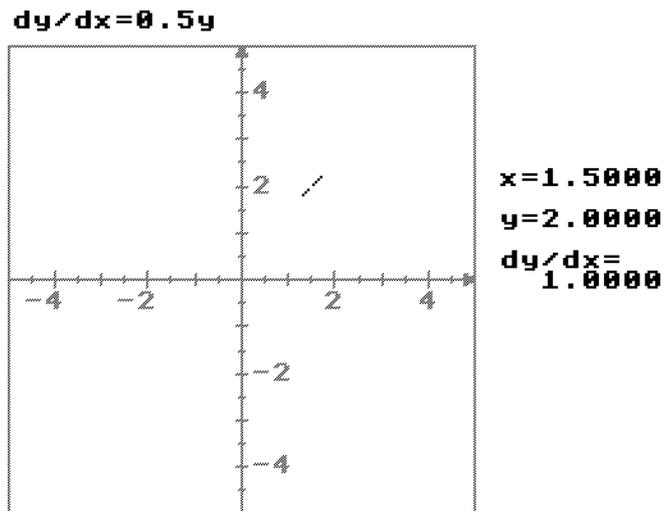


Figure 7 : part of a curve of gradient $\frac{1}{2}y$

By simply moving the segment around the computer screen under software control and leaving a trace of such curves fitting end to end an approximate solution curve can be built. Figure 8 shows such a curve and an array of line segments showing the directions of other possible solution curves. Through each point in the plane there is a unique solution of the differential equation

$$\frac{dy}{dx} = \frac{1}{2}y.$$

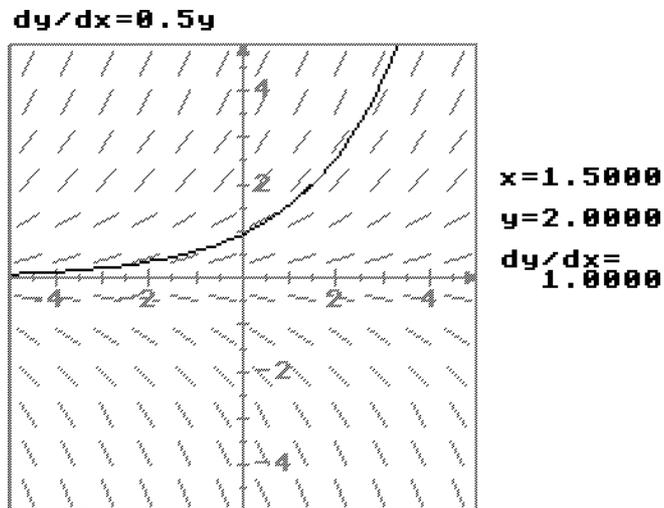


Figure 8 : A solution curve and an array of other segments with the appropriate gradient

This is the essential idea of a differential equation. It is an equation which specifies the gradient of a graph and, provided it does this unambiguously at every point, there is a unique solution of the differential equation through every point in the plane. There is thus a *whole family of solutions to a given differential equation.*

Cumulative growth

The final ingredient of the calculus is given by the cumulative growth of a function. The most straightforward example is to take a function and calculate the growing area under the graph. Say one might take a graph as in figure 9 and, by some method or other, work out the area from a fixed point a to a variable point x . The area will then be a function $A(x)$.

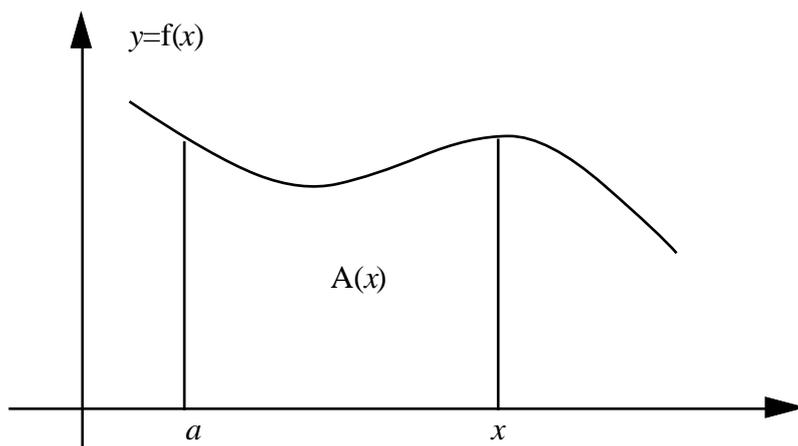


Figure 9 : The area $A(x)$ under the curve from a to x

The area from $x=a$ to $x=b$ is easily calculated approximately on a computer by simply chopping up the interval from a to b into small lengths which, in the Leibniz notation will be denoted by dx , and then adding together the rectangles of height $f(x)$ and width dx (figure 10).

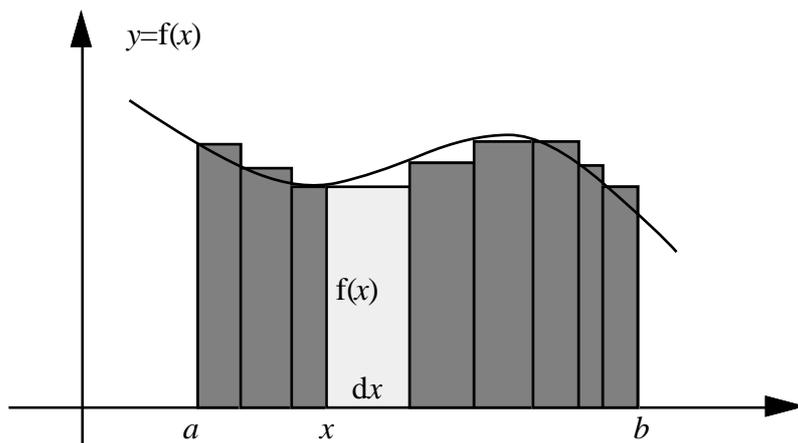


Figure 10 : The Leibniz sum

By taking thinner and thinner strips the Leibniz sum is likely to give better and better approximations to the area.

Leibniz denoted the area by

$$\int_a^b f(x) dx$$

where the elongated S denoted the first letter of the Latin word Summa, for sum. This Leibniz notation should therefore be read “the sum from a to b of $f(x) dx$ ”. It is also called “the *integral* from a to b of $f(x) dx$ ”. The process of working out the sum is called *integration*. You can think of integration as adding up of all the little bits to make up the whole, “integrating” them, as it were into a cumulative total, in the usual sense of the verb “to integrate”.

For more than three centuries Leibniz has been misunderstood. Old (i.e. pre-1991) textbooks often accused Leibniz as having thought up a good notation with a very imprecise meaning and the poor chap has been given an awful press, particularly in chauvinistic British publications. His salvation is at hand. We have simply been looking at the wrong diagram for three hundred years. This amazing misconception is now being set right for the first time in these pages of *Mathematics Review*.

We should not look at figure 10 to see what is going on. We should suppose instead that there is a function $I(x)$ whose derivative is $I'(x)=f(x)$ and instead draw the corresponding graph of $I(x)$ (figure 11).

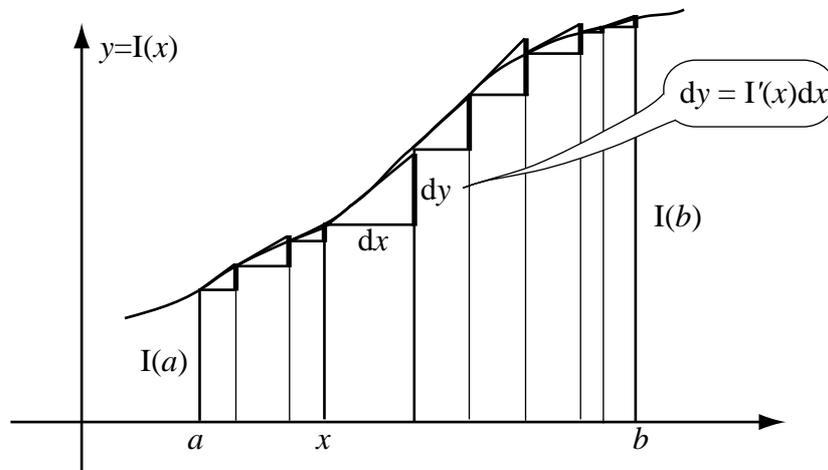


Figure 11 : The Leibniz sum as a sum of vertical line segments

This picture has the same subdivision of the interval from a to b into subintervals. As we saw earlier in figure 6, the value of $I'(x) dx$ is equal to the vertical distance to the tangent. Because $f(x)=I'(x)$, the Leibniz sum

$$\int_a^b f(x) dx = \int_a^b I'(x) dx$$

is the sum of these vertical segments.

The picture as it stands is not very helpful. But now we must see it through our new spectacles and set Leibniz's reputation straight. Figure 12 shows the same idea with a large number of very thin strips.

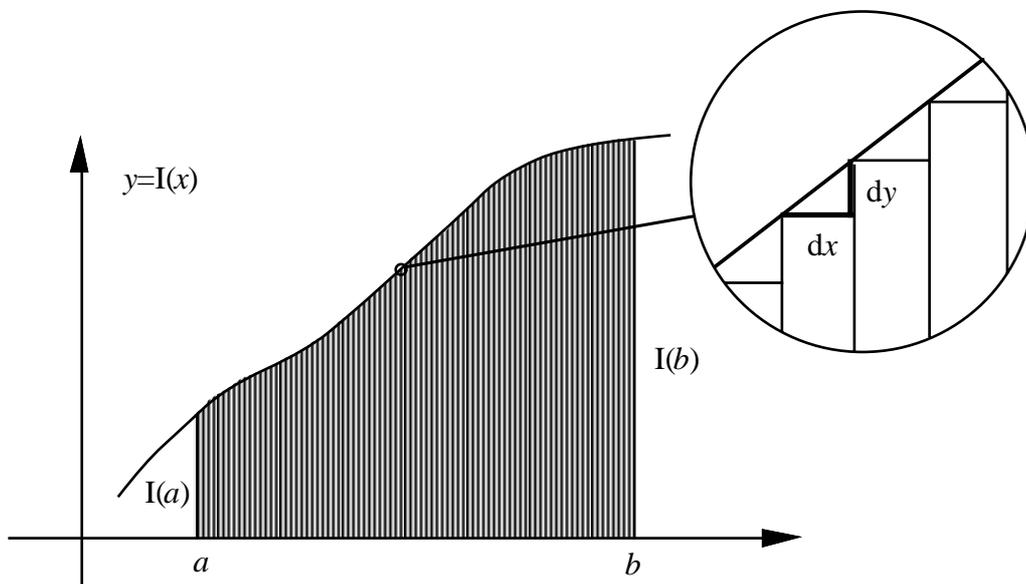


Figure 12: The Leibniz sum as the sum of the risers to the curve $y=I(x)$

Because the graph of $y=I(x)$ has a derivative, under a microscope it will look straight. So now the vertical steps are approximately equal to staircase risers. Adding together all these risers when the strips are very small gives the total sum to be the rise from $x=a$ to $x=b$, which is

$$I(b)-I(a).$$

Thus, when a very large number of strips are taken, the value of the Leibniz sum is

$$\int_a^b f(x) dx = I(b)-I(a).$$

This result is called the *Fundamental Theorem of the Calculus*. It says that, if you wish to calculate the integral

$$\int_a^b f(x) dx$$

then all you need do is to find a function $I(x)$ such that $I'(x)=f(x)$ and then the integral is just $I(b)-I(a)$.

For instance, to calculate

$$\int_4^5 2x dx$$

since $I(x)=x^2$ satisfies $I'(x)=x$, then the integral

$$\int_4^5 2x \, dx = I(5)-I(4) = 5^2-4^2 = 25-16 = 9.$$

You may find this a breath-taking step to make, but closer scrutiny will show that the argument can be made really watertight and that is one of the items you should study in more detail during your two years of calculus.

Thus the calculus gives powerful and immensely simple ways of computing rates of change (derivatives) and growth (integrals). Through the Fundamental Theorem of the Calculus it shows that these are related in a surprisingly simple way. It will take two years of study at A-level to fill in all the details and develop all the symbolic techniques to calculate derivatives, integrals and solutions of differential equations. But the pictures shown in this article demonstrate the nature of the fundamental ideas in a manner akin to the original vision of Leibniz, married to modern computer techniques to give graphic representations of the ideas. I hope they will help you see how simple the ideas are at heart, in a subtle way which sets the calculus straight.

Further Reading

Graphic Calculus I, II, III: Books and computer software for the BBC, Master, Archimedes, Nimbus PC and X-series computers from Glentop Press, Unit 11, Stirling Industrial Centre, Stirling Way, Boreham Wood, WD6 2BT.

Real Functions & Graphs (including the *Solution Sketcher*): Computer software for the same computers from Cambridge University Press.

David Tall is author of *Graphic Calculus* and of the 16-19 Software designed to visualize ideas in the new 16-19 A-level approach to the subject. He is currently Reader in Mathematics Education at Warwick University, though admits his title is a misnomer because the endless work on computer software is annually requiring him to have stronger spectacles.