

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): David Tall

Article Title: The Gradient of a Graph

Year of publication: 1985

Link to published version:

Publisher statement: None

The Gradient of a Graph

David Tall

Mathematics Education Research Centre

University of Warwick

COVENTRY CV4 7AL

In 'understanding the calculus' ⁴ I suggested that the traditional approach to the subject is not the most appropriate sequence for learning, and outlined an alternative development using the powerful visualizations possible with a microcomputer. In this article I introduce a dynamic interpretation of the gradient of a graph which leads naturally into the notion of differentiation.

The traditional syllabus causes difficulties at the outset by attempting (unsuccessfully) to introduce the general notion of limit before discussing the gradient of a curved graph. The development suggested here starts with the exploration of magnifying graphs then uses a numerical gradient illustrated by moving pictures. Limiting ideas arise implicitly in the course of the exploration, but their full implementation may be postponed to a more appropriate stage. Graphic visualization allows a wider range of examples and non-examples to be considered, so that students learn not only how to calculate a derivative, but also what it means for a function to be differentiable or non-differentiable. The seeds are sown for a level of insight that has hitherto been considered unattainable.

The logical approach

A process of mathematical refinement over the centuries has produced a logical approach to the calculus that is embodied in the opening sequence of topics in the proposed A-level core curriculum:

- The idea of a limit and the derivative defined as a limit.
- The gradient of a tangent as the limit of the gradient of a chord.

In the preface of *Psychology of Learning Mathematics* ¹, Richard Skemp criticizes the purely logical approach in the following words:

Published in *Mathematics Teaching* (1985), 111, 48–52.

It gives only the end-product of mathematical discovery ('This is it: all you have to do is learn it'), and fails to bring about in the learner those processes by which mathematical discoveries are made. It teaches mathematical thought, not mathematical thinking.

To promote mathematical thinking we require a cognitive approach that builds on the learner's current state of knowledge and understanding.

A cognitive approach

It will be assumed that the learner meeting the calculus for the first time is able to do the following:

- manipulate simple algebra (to be able to handle the formulae),
- sketch and visualize the graphs of simple curves (lines, quadratics and simple cubics at least)
- calculate the gradient of a straight line and understand positive and negative gradients using the formula $\frac{y_2 - y_1}{x_2 - x_1}$ for the gradient of the straight line through (x_1, y_1) , (x_2, y_2) .

A student with these capabilities could begin his study by investigating the behaviour of graphs of simple functions under high magnification. Prior to the advent of the computer this would have been a non-trivial task where the technicalities could so easily obscure the essential factors. But modern technology makes it possible to use a computer to draw the graph to allow the mind to concentrate on the essential features. Any adequate graph-drawing package could be used, but the MAGNIFY program in *Graphic Calculus I*⁵ is designed specifically to allow such investigations, exhibiting both the graph and a magnification of a small portion simultaneously. The investigation can be done as a classroom activity, either with small groups of students sharing micros, or with a single computer and socratic dialogue between teacher and pupils. Draw any graph and magnify it to see what happens. (Figure 1.)

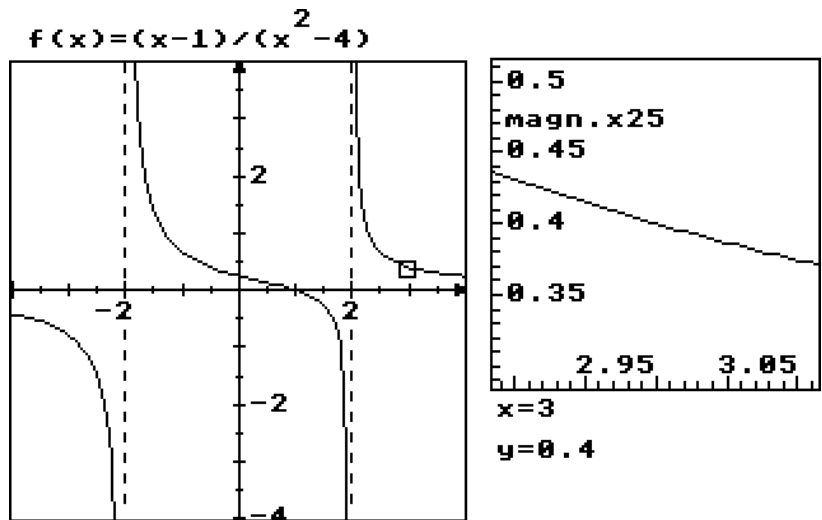


Figure 1 : magnifying a small portion of a graph

Virtually any graph drawn using the standard formulae will look less curved under high magnification. Rational functions, combinations of trigonometric, logarithmic and exponential functions all have this property.

It is a major problem to invent a function that fails to magnify to look straight. Yet in nature such things are common. The edge of a leaf under magnification has tiny irregularities and higher magnification reveals yet smaller ones. Even a plastic ruler, that is supposedly straight, looks irregular under a microscope. Thus the phenomenon of magnifying to 'look straight' is a very special mathematical property and it should be seen as such from the beginning.

Graphs with corners

A small inroad into the 'locally straight' property can be made by drawing graphs with 'corners'. The absolute value function (or modulus) $\text{abs}(x) = |x|$, strips the sign of a number to give the positive value of the pair $x, -x$, so that

$$\text{abs}(1.5)=1.5, \text{abs}(-2.7)=2.7, \text{abs}(0)=0, \text{abs}(-\pi)=\pi.$$

The functions $\text{abs}(x)$, $\text{abs}(x^2-1)$, $\text{abs}(x(x-1)(x-2))$ all have 'corners'. More interesting pictures can be made out of combinations of functions. For instance, $\text{abs}(x^2-1)+x$ (figure 2) has corners at -1 and 1 .

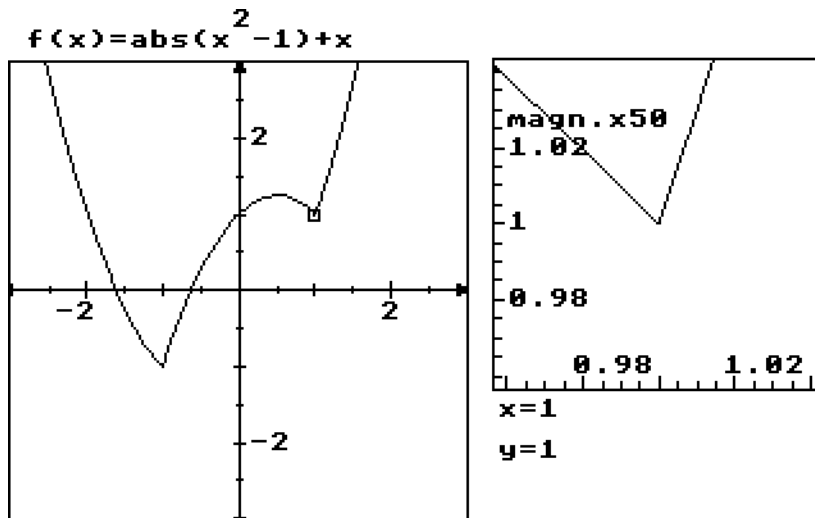


Figure 2 : A graph which is not locally straight

Nowhere differentiable functions

Graphs with ‘corners’ may still have the ‘locally straight’ property everywhere else. A *nowhere differentiable function* is one that nowhere magnifies to look straight. A simple example is the blancmange function³ which was described by the Japanese mathematician Takagi² in 1903. Using a dynamic display, as in the program BLANCMANGE⁵, it is possible to give an idea of this function to mathematicians and non-mathematicians alike. It is much more difficult to appreciate the notions with a static text and selected still pictures. But I shall attempt the task in brief.

The idea is to build up a function in stages. The first stage is a saw-tooth which rises linearly to 1/2 and falls to 0 in every unit interval (figure 3). This is the first approximation $b_1(x)$ to the blancmange function. Next we take a half-size saw-tooth which rises to 1/4 and falls to 0 twice in each unit interval (figure 4).

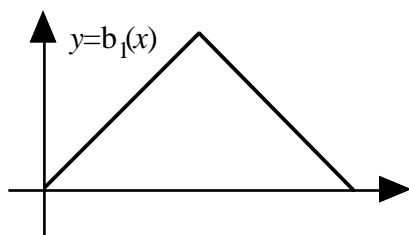


figure 3: a sawtooth (in every unit interval)

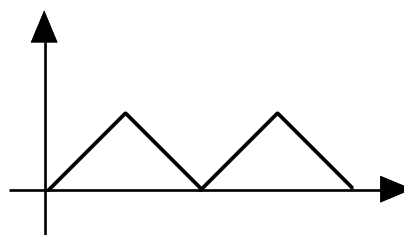


figure 4: a half-size sawtooth

Adding them together gives the second approximation $b_2(x)$ (figure 5).

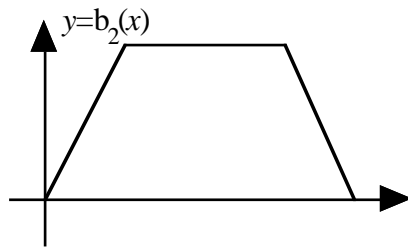


figure 5: the sum of a sawtooth and a half-size sawtooth

We proceed in this manner, each time taking a sawtooth half the previous size and adding it to the approximation to get a better approximation to the blancmange. Figure 6 shows the next stage with a 1/4-size sawtooth added to $b_2(x)$ to give the third approximation $b_3(x)$.

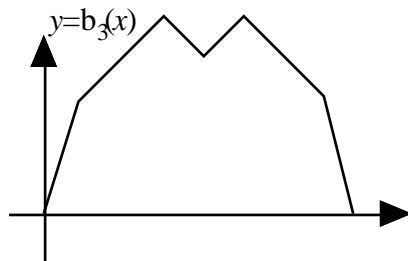


figure 6: adding another (quarter-size) sawtooth to give $b_3(x)$

When drawing this process on a TV monitor, the added saw-teeth soon get so small that the picture stabilizes (after eight or so stages on most modern VDUs). (Figure 7.)

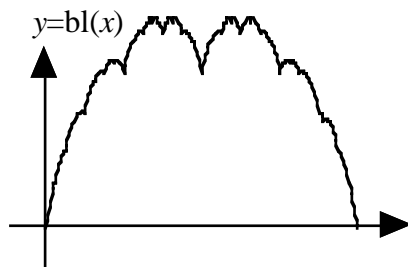


Figure 7 : the blancmange function

We must imagine the process going on forever so that the blancmange function $bl(x)$ is visualized as the sum of an infinite number of teeth. A practical picture only reveals the details of a few teeth, but we can imagine that we can magnify the graph to any desired extent to reveal the finer detail. This is done in the program BLANCMANGE.

When the blancmange function is magnified, a strange thing happens: tiny blancmanges can be seen all over the place!

To understand why this happens, consider what would occur if we left out the first sawtooth and added up the rest. We would start at a half-size sawtooth and successively add sawteeth each half the previous size. The result would be as before, but would give a half-size blancmange.

Adding the first saw-tooth to this sum, reveals the full blancmange as the sum of the first sawtooth and a half-size blancmange.

Similarly, starting the sum at the third (quarter-size) sawtooth and adding all the succeeding saw-teeth gives a quarter-size blancmange. The first two teeth added together give $b_2(x)$; so the full blancmange $bl(x)$ is the sum of $b_2(x)$ and a quarter-size blancmange.

In general

$$bl(x) = b_n(x) + 1/2^n\text{-size blancmange.}$$

Further inspection shows that $b_n(x)$ is made up of straight line segments over each x -interval of length $1/2^n$. Over each interval the $1/2^n$ -size blancmange is added to give the full blancmange function. Figure 8 shows $b_4(x)$ superimposed on the blancmange function. There are tiny blancmanges clearly visible on each horizontal portion of $b_4(x)$ and on the other portions the blancmange is sheared to sit on the line-segment. Thus no tiny portion of the blancmange function will ever look straight under high magnification. It is the supreme non-example. It is *nowhere* differentiable.

The best time to introduce such a function to students is a matter of personal taste for the teacher. Many may consider it too difficult for weaker pupils at an early stage but my inclination would be to approach it experimentally fairly early on using a prepared piece of software such as BLANCMANGE. Though the *technicalities* are difficult, the idea of a wrinkled graph is not. It helps students realize the special nature of those smoother graphs that look straight when small portions are highly magnified. This highly specialized property is the foundation of the differential calculus.

The gradient of a graph

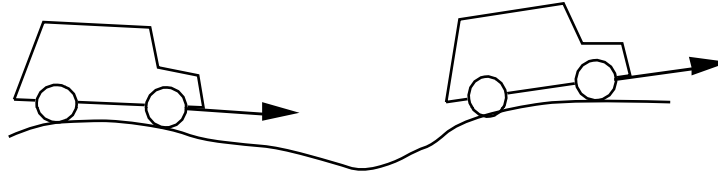


figure 9: changing gradient on a road

How can we measure the gradient of a road? Driving along in a car (figure 9), at any stage the line through the car hub caps makes a measurable angle with the horizontal. The gradient can be measured as the vertical step (from the axis of one hub-cap to the other) divided by the horizontal step. This gives a practical interpretation of the gradient all along the road, no matter how rough the road is. But a car with a different base-length on a very uneven road might give quite a different gradient function, especially a tiny toy car on a typical road. This method of practical gradient calculation is only of value if the gradient does not change much in the car's length.

The same idea can be used on any graph, though it is easier to fix the horizontal distance between two points rather than the base distance between wheels. Consider the graph $y=x^2$ and imagine two points $A=(a,a^2)$, $B=(b,b^2)$ on the graph so that the horizontal distance $c=b-a$ between A,B is fixed. Taking $a=x$, $b=x+c$, measure the gradient of AB as x varies and plot the chord and its gradient using a computer program. (Figure 10 is plotted using the program GRADIENT⁵). What is happening is that the extended chord through the two points $x, x+c$ is plotted as a line and, *simultaneously*, the computer calculates the gradient and plots it as a point. For various values of c we get *practical* calculations of the gradient. For small values of c it may be seen that the chord gradient points lie on a straight line and the approximate equation of the line is $y=2x$. Thus we may conjecture that the *theoretical* gradient of the graph $f(x)=x^2$ is $2x$.

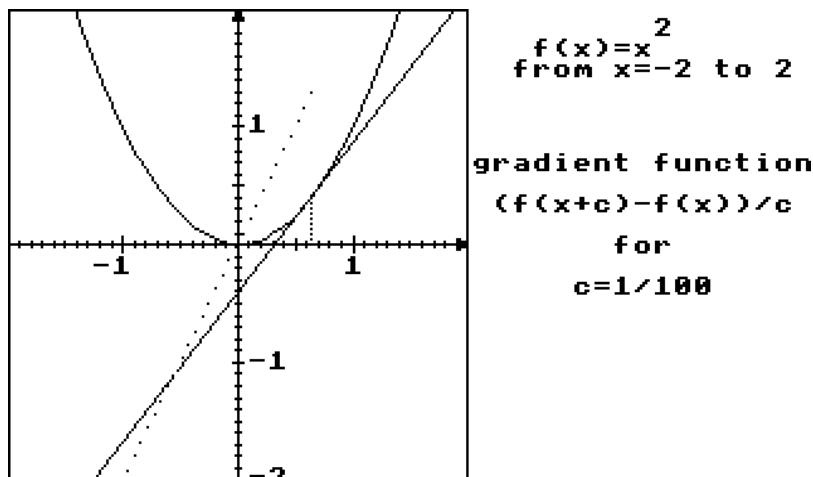


Figure 10 : Building up the gradient of $f(x)=x^2$

This may be confirmed in this particular case by algebra, for the gradient of AB is

$$\frac{b^2-a^2}{b-a} = \frac{(x+c)^2-x^2}{c}$$

which, for $c \neq 0$, simplifies to $2x+c$.

For extremely small values of c this is indistinguishable from $2x$. The derivative $f'(x)$ can be introduced at this stage as the theoretical gradient of the graph. In the case $f(x)=x^2$ we have $f'(x)=2x$.

The roles of practical and theoretical gradient are very different. The practical gradient is calculated numerically using the original function. Every graph has a practical gradient for each value of c . But a graph only has a theoretical gradient at those places where a small magnified portion looks almost straight. The theoretical gradient is then the gradient of this small magnified part.

Discovering the gradient formula for a power

Using a computer program that plots the gradient of the graph $f(x)=x^3$ one may try to estimate what happens to the gradient function when c is fairly small (figure 11). When faced with this gradient graph, students familiar with simple graph-shapes might guess the formula. I've seen various guesses suggested: a common one is x^4 because the graph is a sort of quadratic, but clearly isn't x^2 . Overplotting the graph of x^4 disproves this conjecture and a better guess of $2x^2$ often leads to the gradient of $f(x)=x^3$ being conjectured as $f'(x)=3x^2$.

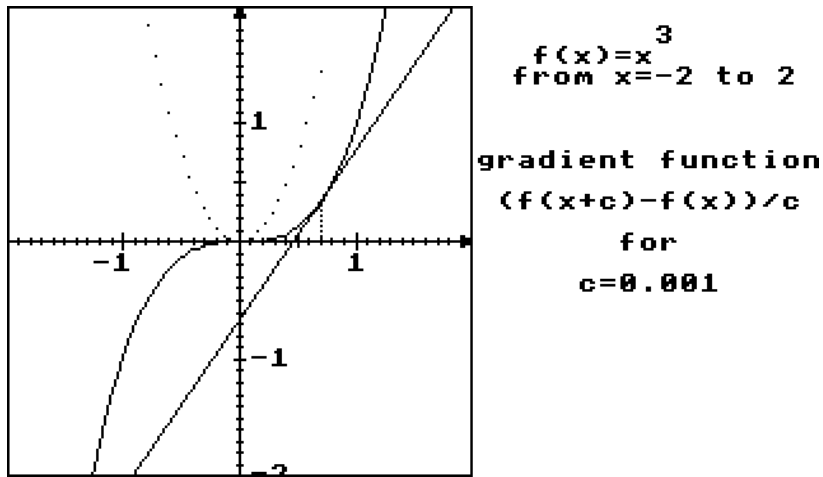


Figure 11 : The gradient of $f(x)=x^3$

At this stage the table with just two results:

$f(x)$	$f'(x)$
x^2	$2x$
x^3	$3x^2$

is usually sufficient for students to suggest the general formula:

$$f(x)=x^n \text{ has derivative } f'(x)=nx^{n-1}.$$

A gradient-drawing program which allows this formula to be typed in and tested can lead to a variety of investigations. Does it give an adequate picture for $n=4$, $n=5$, ... What about $n=27$?

If the students have the appropriate background, one might try negative n , such as $n=-1$, -2 , or fractional n such as $n=1/2$, $n=3/4$ etc. It is fascinating to see the graphs drawn appropriately, though in a few instances the computer may not respond in the expected mathematical way. For instance the mathematical formula $f(x)=x^{1/3}$ for the cube root is defined for all values of x , but computers can often only calculate it for positive values.

Another generalization is to simple polynomials. For instance one might investigate the effect of multiplying by a constant, through considering $2x^2$, or kx^2 for various values of k . Multiplying by a positive constant k may be seen to stretch the y -scale by that constant, leading to a gradient multiplied by the same constant. Thus the gradient of kx^2 is $2kx$.

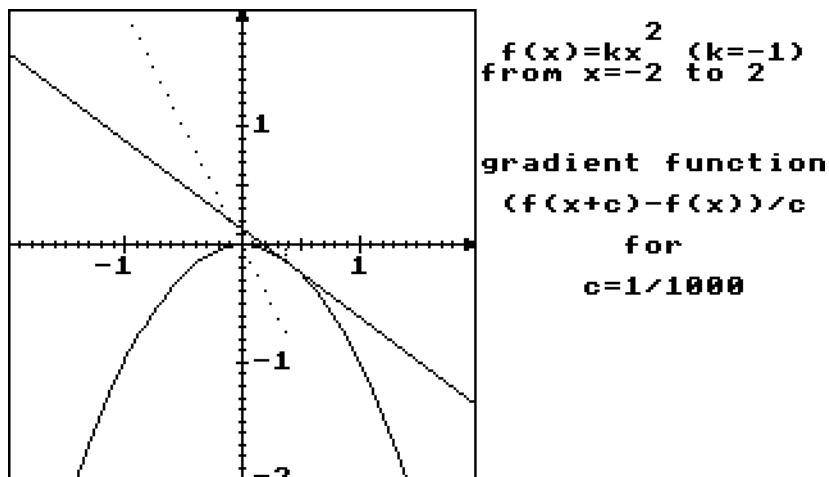


figure 12 : The gradient function for $y = -x^2$

What happens when k is negative? Figure 12 shows the gradient being drawn for the graph of $y = kx^2$ for $k = -1$. The gradient is $-2x$, which is, of course, $2kx$ where $k = -1$. Multiplying by -1 reflects the graph in the x -axis, keeping the x -step the same, but multiplying y -steps by -1 .

By gaining a feeling for the concepts in this way, the more general formula for a polynomial follows naturally. At this stage, even with only one computer in a class, it is possible to accompany class exercises in formal differentiation of polynomials by students taking it in turns to use the computer to check their answers. All that is necessary is to draw the graph of the function and a practical gradient, then to compare this gradient function with the graph of the derivative found by formal differentiation.

Standard formulae

Every year the physics department seems to come to simple harmonic motion before the mathematics department has covered differentiation of trigonometric functions. A gradient drawing program will quickly give a temporary solution to the problem. By drawing the graph of $f(x) = \sin x$ and its gradient, the result clearly suggests that $f'(x) = \cos x$. For $f(x) = \cos x$, the shape of the graph of the gradient quickly suggests $f'(x) = -\sin x$, complete with the minus sign. (Figure 13.)

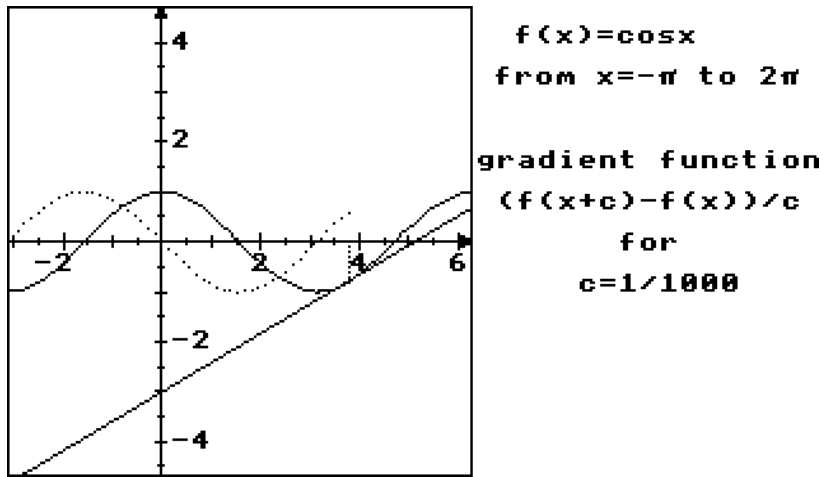


Figure 13 : The gradient function of $\cos x$ is 'sin x upside down'

When I first used the program I was convinced that the derivative of $\tan x$ was too difficult to approach using graphical guesswork. I have been proved wrong on many occasions. Students are far smarter than we give them credit. Looking at the gradient function of $f(x) = \tan x$ gives an interesting graph shape (figure 14). Students with experience in graph sketching can see that the gradient graph is roughly the shape that might occur if the gradient of $\tan x$ were squared, except that it passes through $y=1$ when $\tan x=0$. Hence they add one and suggest the derivative of $f(x) = \tan x$ is $f'(x) = 1 + \tan^2 x$.

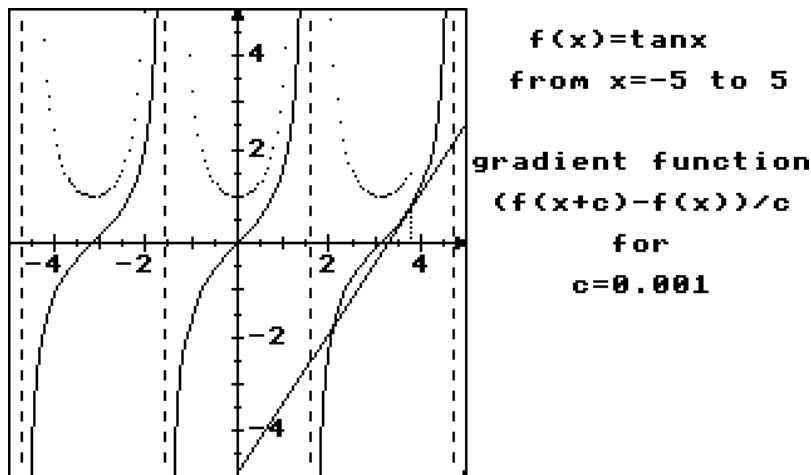


Figure 14 : The gradient of $\tan x$

Another interesting rule to investigate is to find a value of k such that k^x has derivative also k^x . For $k=2$ the derivative is below the original graph, for $k=3$ it is above. Homing into a suitable value plausibly suggests that $k=e$.

The natural logarithm, $\ln(x)$, also leads to amusing conjectures. Students drawing the logarithm and a practical gradient (figure 15) often think that there must be something ‘logarithmic’ about the formula for the derivative. Thus guesses like ‘ $-\ln(x)+1$ ’ or ‘ $(\ln(x))^2+1$ ’ are commonplace. The idea that the derivative of the natural logarithm might be $1/x$ is then quite a revelation.

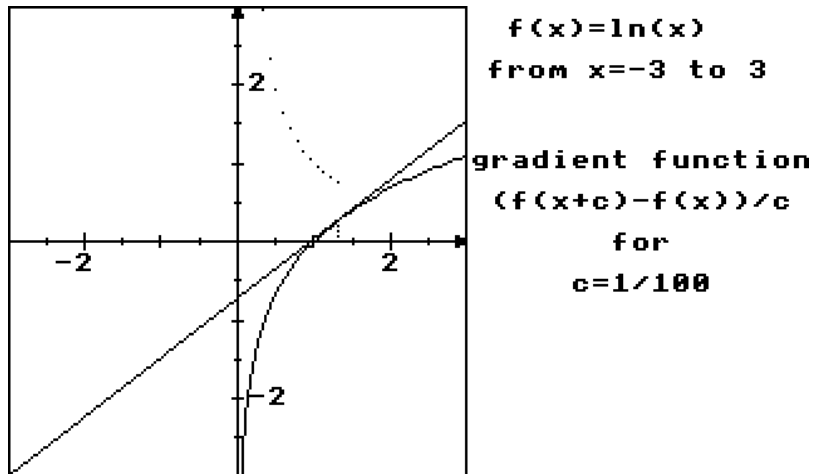


Figure 15 : Guess the formula for the gradient of $y=\ln x$

The picture suggests a clear gap in the theory (usually left unfilled in traditional courses): $\ln(x)$ is defined only for $x > 0$ but the derivative formula $1/x$ is also defined for $x < 0$. What happens to the left of the origin? Investigations by the students, or socratic dialogue guided by the teacher, lead to the function:

$$f(x)=\ln(\text{abs}(x))$$

satisfying $f'(x)=1/x$ for all non-zero x . (Fig 16). Regrettably this picture is only a still snapshot of one instant of the process; it should be seen growing dynamically in all its glory!

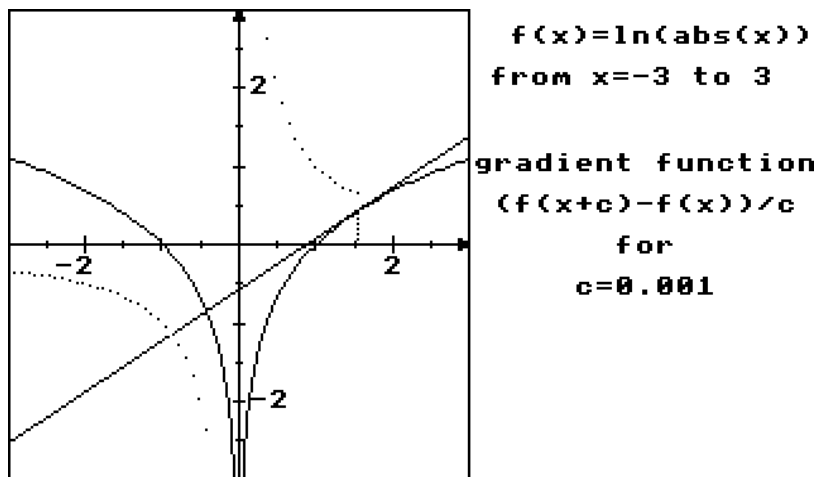


Figure 16 : The function whose gradient is $1/x$ for $x > 0$ and $x < 0$

The gradient as a dynamic process

The experience that a student gets from plotting gradient function enables the gradient of a curve to be visualized as a dynamic process. This comes from observing the extended chord clicking along the curve, simultaneously plotting the gradient as a sequence of points. It becomes possible for the student to scan a curve and see the gradient vary, just as it is to see a cross-section of a mountain road and see the gradient rise and fall. Suppose that the graph in figure 17 is given and one is asked to draw a sketch of its gradient function (assuming that the graph is fairly smooth).

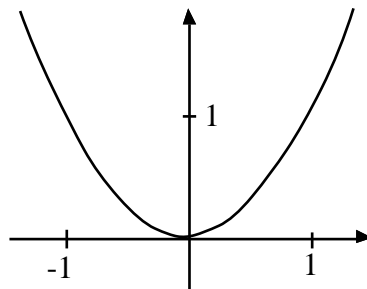


Figure 17 : Sketch the gradient of this curve ...

Someone who has only studied the beginnings of the traditional method must proceed by a sequence of steps:

- 1) The graph looks like $f(x) = x^2$
- 2) I know the formula $f'(x) = 2x$ for this graph

3) I can draw the straight line graph $y=2x$.

A student with experience of gradient plotting could glance along the curve to see:

the gradient is at first large and negative, moving to zero at the origin, then growing large and positive.

With this insight a rough sketch of the gradient graph may be drawn immediately.

For slightly more unusual graphs, such as figure 18, the student with only a traditional background in the calculus may not be able to cope with the gradient function, because an appropriate formula is not obvious. Someone following a graphical approach may see that the gradient is initially large and positive, diminishing to zero for $x=0$ and staying zero thereafter, again being able to make a rough sketch. Thus the student with the graphical approach has a better chance of understanding of the gradient concept and developing of a relational understanding which links formulae, numerical ideas and graphical insight.

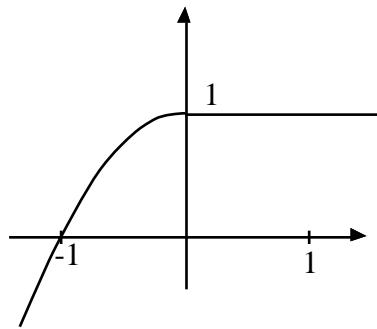


Figure 18: Sketch the gradient of this curve ...

The limit concept

Depending on the inclinations of the teacher and the readiness of the students, the limit concept can be introduced at any suitable point along the route exploring the gradient function. For the more able student this is likely to be fairly early on. But in all cases it is best considered at a time when the need for it is clearly seen. Drawing gradients with a fixed step c gives the impression that the limiting process is rarely necessary, just take $c=1/100$, or $c=1/1000$ and a good pictorial approximation to the derivative results. My own experience suggests that a good time to discuss the necessity for the limit is with the gradient of $f(x)=1/x$. Draw the gradient function for a fairly large size of c and you will see why. (Figure 19.)

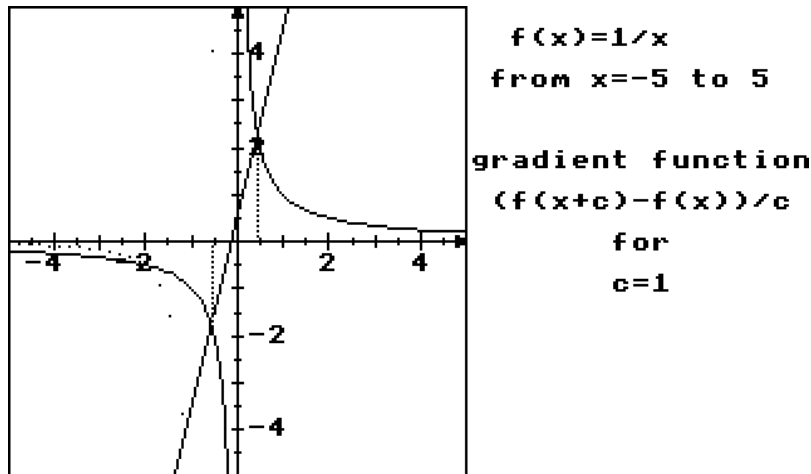


Figure 19 : A bad approximation to the gradient of the graph

The graph of $f(x)=1/x$ is always sloping down and the gradient is clearly always negative. But a fixed value of c hits a position where x is to the left of the origin and $x+c$ is to the right. Thus a false positive gradient is given through joining a point on the left component of the graph to a point on the right. The only solution is to make sure that the value of c chosen is suitably small wherever the gradient is being calculated. For this graph no fixed value of c is suitable. To calculate the gradient near a point x requires $|c| < |x|$ so that x and $x+c$ are at the same side of the origin.

The formula for the practical gradient from x to $x+c$ in this case is

$$\begin{aligned} \frac{f(x+c)-f(x)}{c} &= \frac{1/(x+c)-1/x}{c} \\ &= \frac{x-(x+c)}{cx(x+c)} \\ &= \frac{-1}{x(x+c)} \end{aligned}$$

The limit can be thought of as a dynamic process:

‘as c approaches 0, the gradient approaches the derivative’.

Using an arrow to represent the dynamic activity in this case we have:

$$\text{as } c \rightarrow 0, \text{ so } \frac{-1}{x(x+c)} \rightarrow \frac{-1}{x^2} \text{ .}$$

The derivative $f'(x)=-1/x^2$ gives a true value for the gradient of the graph at every (non-zero) x , but the practical gradient from x to $x+c$ may be grossly inaccurate near the origin. Thus the need to take limits arises naturally at a later stage of the development. It need not be forced meaninglessly into the beginning of the calculus syllabus.

When and how to introduce differentiation

The approach proposed in this article allows an investigation of gradient and the development of the formulae for differentiation to be carried out in the early stages without any formal discussion of the limiting process. The methods are quite sufficient to introduce the concepts to candidates taking calculus in O-level mathematics.

It is known that students have various misconceptions related to limiting processes, I outlined some of them in my previous article ⁴. In the next issue of *Mathematics Teaching* I shall show how to tackle some of these problems using a graphical approach to chords and tangents. By considering the tangent as a vector I shall also explain how the Leibniz notation dy/dx can be considered as a ratio of lengths which are simply the components of the tangent vector (dx,dy) . Yet more fondly held beliefs will come under close scrutiny using new insight given by the microcomputer. Will they stand the test?

References

1. R R Skemp: *The Psychology of Learning Mathematics*, Penguin 1971.
2. T Takagi: A simple example of the continuous function without derivative, *Proc. Phys.-Math. Japan*, 1 (1903) 176-177.
3. D O Tall: The blancmange function, continuous everywhere but differentiable nowhere, *Mathematical Gazette* 66 (1982) 11-22.
4. D O Tall: Understanding the calculus, *Mathematics Teaching*, 110 (1985) 49-53.
5. D O Tall: *Graphic Calculus*, Glentop Publishing, 1986.