On Differentiation II

On the rules of differentiation

By Chris Fenwick

110 Uploaded on 02/11/2022

fenwick_chris@hotmail.com, uczlcfe@ucl.ac.uk

Table of Content

1.1 Introduction

Up to this point we have seen the definition of the derivative and how to use it to differentiate simply functions like x^n. But what happens when we want to differentiate more than one function at once? For example, what is the derivative of the sum of two or more functions? What is the derivative when functions are multiplied together or divided by one another ? In other words, we now have to see what happens when we apply the definition of the derivative on sums, products, and divisions.

We know how arithmetic operation work: addition is addition, subtraction is subtraction, multiplication is multiplication, division is division, and powering is powering. We also know how to use logs as a way of simplifying arithmetic.

But $\frac{d}{dx}$ is a totally new operation. It isn't just based on the division of differences. It is based on taking the limit of this division. And it is this limiting process which creates the brand new operation of differentiation. Without it we would just be doing the division of two differences.

What all this implies is that we have to find out how $\frac{d}{dx}$ works with our standard operations of addition, subtraction, multiplication, division, powering, and rooting. This is not as obvious as it may seem, as we will see in some of the following sections. And even if some of these operations do seem obvious we still need to justify them mathematically.

1.2 The multiplication by a constant rule for differentiation

We know from the power rule that if $f(x) = x^2$, $df/dx = 2x$. But what about the situation when a function is multiplied by a constant? In other words if we have $f(x) = k.x^2$, where k is a real number, will doing

$$\frac{d}{dx}(kf(x))$$

give the same answer as if we do

$$k\frac{df(x)}{dx} \text{ ?}$$

For example, if $f(x) = x^2$ we have

$$\frac{df}{dx} = 2x.$$

If we then have $g(x) = 3x^2$ we will have

$$\frac{dg}{dx} = \frac{d}{dx}(3x^2),$$

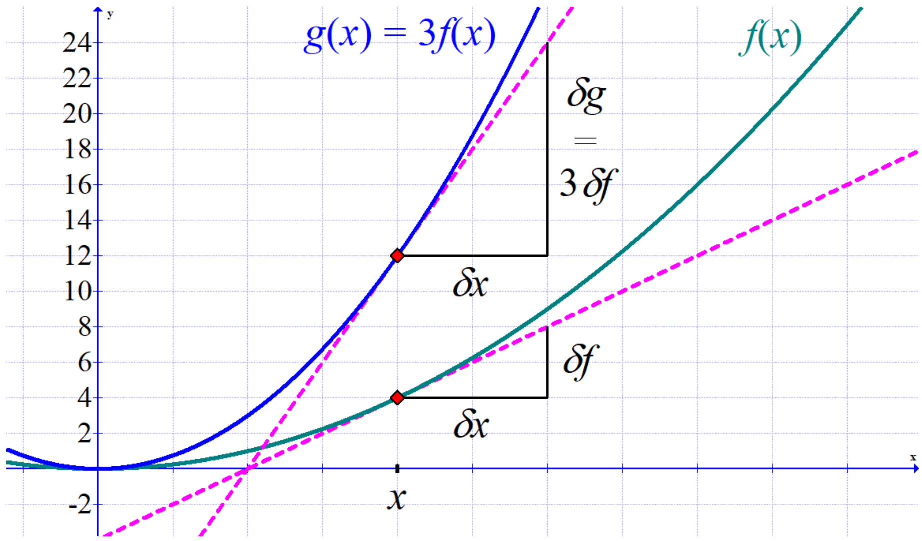
but will this give us

$$\frac{dg}{dx} = 3\frac{df}{dx} = 3(2x) = 6x\ ?$$

The answer is yes.

1.2.1 An informal proof of the multiplication by a constant rule for differentiation

To see why the previous answer is correct consider a function $f(x)$. Suppose we now multiply this function by 3: $g(x) = 3f(x)$. In doing this we are simply magnifying the y value of $f(x)$. Geometrically speaking this means we are staying at the same x position but our y position is stretched (or shrunk depending on the function) in the vertical direction, as shown below:



$$\frac{\delta f}{\delta x} = \frac{4}{3} \quad \text{and} \quad 3\frac{\delta f}{\delta x} = 3 \times \frac{4}{3} = \frac{12}{3} = \frac{\delta g}{\delta x}$$

Note that, compared to δf, δg is the magnified (or shrunk) change in height of $g(x)$ so

$$\frac{\delta g}{\delta x} = 3\frac{\delta f}{\delta x}.$$

Hence as $\delta x \to 0$

$$\frac{dg}{dx} = \frac{d}{dx}(3f) = 3\frac{df}{dx},$$

and in general the derivative of any constant k times $f(x)$ is the same as the constant k times derivative of $f(x)$:

$$\frac{d}{dx}(kf) = k\frac{df}{dx}.$$

1.2.2 A formal proof of the multiplication by a constant rule for differentiation

Again, the way to prove this properly is to go back to first principles, i.e. to use the definition of the derivative. Therefore, given a function $f(x)$ the derivative of $k.f(x)$ is given by

$$\frac{d}{dx}(k.f(x)) = \lim_{\delta x \to 0} \frac{k.f(x+\delta x) - k.f(x)}{\delta x}, \quad (1)$$

$$= k.\lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x}, \quad (2)$$

$$= k.\frac{df(x)}{dx}, \quad (3)$$

Expression (1) is about applying the definition of the derivative to $k.f(x)$. From (1) we do whatever algebra is needed to recover the difference between $f(x + \delta x)$ and $f(x)$. At this stage we are able to factor out constant k because the limiting process will have no effect on it. This gives expression (2). Expression (3) comes from applying the limit in (2).

1.3 The addition and subtraction rule for differentiation

Let us say that we have two numbers, 2 and 3. How do we add them? We add them simply by adding them! In other words we simply do 2 + 3 to get the answer 5. It also happens to be the case that we can do 3 + 2 to get the same answer 5. But will this property of addition (or subtraction) also apply to the derivative of the sum or difference of two functions $f(x)$ and $g(x)$?

In other words if we do

$$\frac{d}{dx}(f \pm g)$$

will we get the same answer as if we do

$$\frac{df}{dx} \pm \frac{dg}{dx}\ ?$$

For example, if $f(x) = x^3$ and $g(x) = -3x^{1/2}$ we have

$$\frac{df}{dx} = 3x^2 \qquad \text{and} \qquad \frac{dg}{dx} = -\frac{3}{2}x^{-\frac{1}{2}}$$

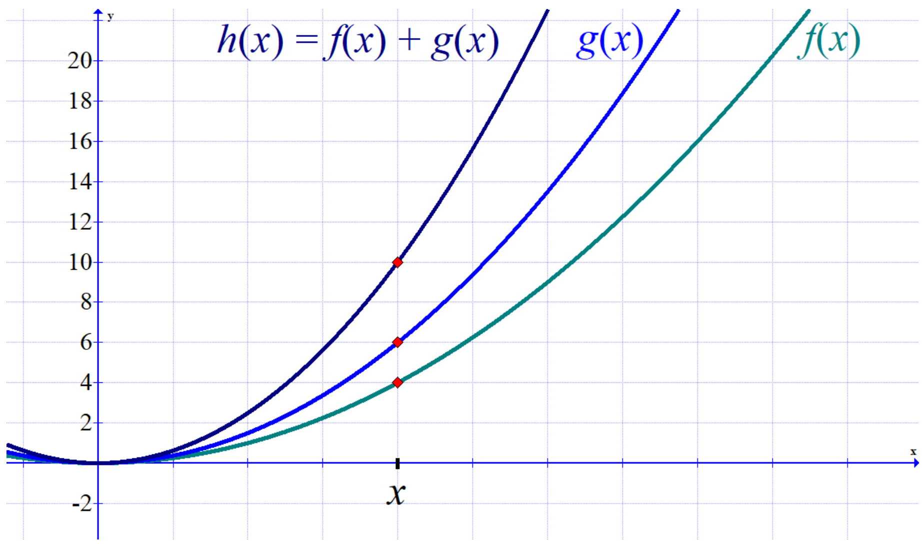
from which

$$\frac{df}{dx} \pm \frac{dg}{dx} = 3x^2 - \frac{3}{2}x^{-\frac{1}{2}} \quad .$$

But will we get the same answer if we differentiate $f(x) + g(x) = x^3 - 3x^{1/2}$? The answer at the moment is that we don't know, and we can't assume it will be so. The reason we can't assume it is because we are dealing with a brand new mathematical 'object', and we don't yet know how it behaves mathematically with respect to arithmetic operations between functions.

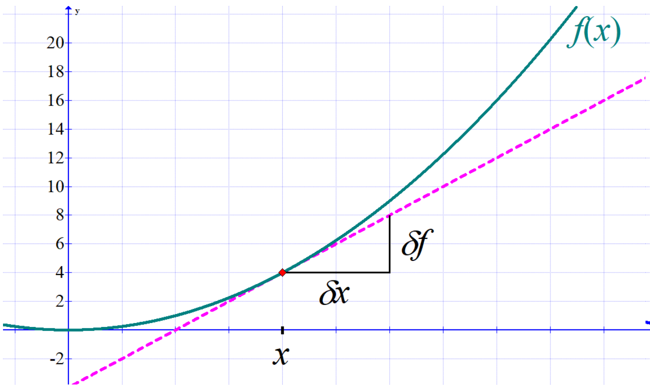
1.3.1 An informal derivation of the addition and subtraction rules for derivatives

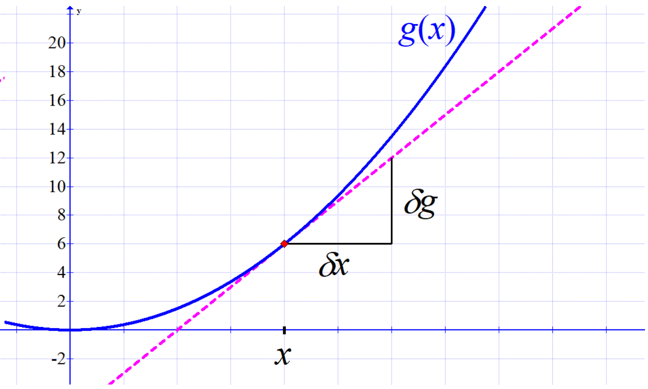
Consider two function $f(x)$ and $g(x)$. We can add them to get $h(x) = f(x) + g(x)$. When we add two functions we are simply adding their respective y values at a given x value. Geometrically speaking this means we are staying at the same x position but our y position is increasing (or decreasing depending on the functions) in the vertical direction, as shown below:



f evaluated at x gives 4; g evaluated at x gives 6; h evaluated at x gives f + g = 10;

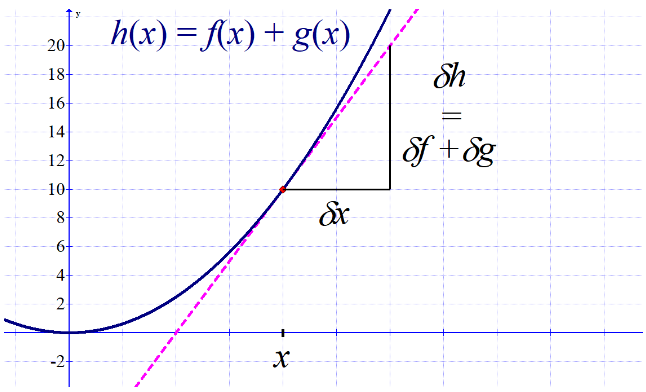
How does this affect the gradient of $h(x)$ at x given the gradients of $f(x)$ and $g(x)$ at x? Well, the horizontal displacement from x to $x + \delta x$ remains the same for all three functions. Only the vertical displacement changes, and this is just the addition of the two separate vertical displacements of $f(x)$ and $g(x)$.





$$\frac{\delta f}{\delta x} = \frac{4}{3}$$

$$\frac{\delta g}{\delta x} = \frac{6}{3}$$



$$\frac{\delta h}{\delta x} = \frac{10}{3}$$

Note that δh is the combined height of $f(x)$ and $g(x)$, so δh can unofficially be expressed as $\delta(f + g)$. Hence $\delta h/\delta x$ can be expressed in terms of combinations of δf and δg as follows:

$$\frac{\delta h}{\delta x} = \frac{\delta(f + g)}{\delta x} = \frac{\delta f}{\delta x} + \frac{\delta g}{\delta x}\ .$$

What this expression is telling us is that the change in height of the sum of the functions is the same as the sum of the separate changes in height of $f(x)$ and $g(x)$.

Therefore as $\delta x \to 0$ the derivative of the sum of $f(x)$ and $g(x)$ is the same as the sum of the separate derivatives:

$$\frac{dh}{dx} = \frac{d}{dx}(f + g) = \frac{df}{dx} + \frac{dg}{dx}.$$

The same approach can be used for subtraction. If $f(x) = h(x) - g(x)$ then we end up with

$$\frac{\delta f}{\delta x} = \frac{\delta h}{\delta x} - \frac{\delta g}{\delta x} = \frac{\delta(h - g)}{\delta x},$$

and as $\delta x \to 0$ we get

$$\frac{df}{dx} = \frac{d(h - g)}{dx} = \frac{dh}{dx} - \frac{dg}{dx}.$$

1.3.2 A formal proof of the addition and subtraction rules for derivatives

In order to prove the addition and subtraction rules properly we have to go back to the "beginning": we go back to the definition of the derivative.

Therefore, given two functions $f(x)$ and $g(x)$ the derivative of $f(x) \pm g(x)$ is given by

$$\frac{d}{dx}(f(x) \pm g(x)) = \lim_{\delta x \to 0} \frac{\{f(x + \delta x) \pm g(x + \delta x)\} - \{f(x) - g(x)\}}{\delta x}, \qquad (4)$$

$$= \lim_{\delta x \to 0} \frac{\{f(x + \delta x) - f(x)\} \pm \{g(x + \delta x) - g(x)\}}{\delta x}, \qquad (5)$$

$$= \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x} \pm \lim_{\delta x \to 0} \frac{g(x + \delta x) - g(x)}{\delta x}. \qquad (6)$$

Expression (4) is about applying the definition of the derivative to $f(x) \pm g(x)$. From (4) we do whatever algebra is needed to recover a form we know. In this case we manipulate the numerator to get the difference between $f(x + \delta x)$ and $f(x)$, and the difference between $g(x + \delta x)$ and $g(x)$. This gives expression (5). Expression (6) is about applying the limits separately to both differences so that we can identify the individual derivatives.

Looking at (6) we see that these are none other than the definition of df/dx and dg/dx. Hence

$$\frac{d}{dx}(f(x) \pm g(x)) = \frac{df}{dx} \pm \frac{dg}{dx} \qquad (7)$$

In other words, the derivative of a sum of terms is the sum of the separate derivatives.

Along with the result of section 1.3 this means that for two functions $f(x)$ and $g(x)$ we have

$$\frac{d}{dx}(k_1 f(x) \pm k_2 g(x)) = k_1 \frac{df}{dx} \pm k_2 \frac{dg}{dx} \qquad (8)$$

for any two constants k_1 and k_2.

Examples

1) If $f(x) = 4x^2 - 3x + 100$ then

$$\begin{aligned} \frac{df}{dx} &= \frac{d}{dx}(4x^2 - 3x + 100)\,, \\ &= \frac{d}{dx}(4x^2) + \frac{d}{dx}(-3x) + \frac{d}{dx}(100)\,, \\ &= 8x - 3. \end{aligned}$$

2) If $y = \sqrt{x} - 1/x + 1/x^2$ then

$$\begin{aligned} \frac{dy}{dx} &= \frac{d}{dx}\left(x^{1/2} - x^{-1} + x^{-2}\right), \\ &= \frac{d}{dx}\left(x^{1/2}\right) - \frac{d}{dx}(x^{-1}) + \frac{d}{dx}(x^{-2})\,, \\ &= \frac{1}{2}x^{-1/2} + x^{-2} - 2.x^{-3}\,. \end{aligned}$$

3) If $x = (t^2 - 7t)/t^3$ then

$$\begin{aligned} x &= (t^2 - 7t).t^{-3} = t^{-1} - 7t^{-2}\,. \\ \frac{dx}{dt} &= \frac{d}{dx}(t^{-1} - 7t^{-2})\,, \\ &= \frac{d}{dx}\left(x^{1/2}\right) - \frac{d}{dx}(x^{-1}) + \frac{d}{dx}(x^{-2})\,, \\ &= \frac{1}{2}x^{-1/2} + x^{-2} - 2.x^{-3}\,. \end{aligned}$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.

1.4 On the product rule for differentiation: $d(u.v)/dx$

Consider the two numbers, 2 and 3. How do we find their product? We find their product by simply multiplying them together! In other words we do 2×3 to get the answer 6. It also happens to be the case that we can do 3×2 to get the same answer 6. But does this property of multiplication also apply to the derivative of the product of two functions?

Suppose we have a function $f(x)$ made up of the product of two terms $u(x)$ and $v(x)$, Then $f(x) = uv$. If we differentiate $f(x)$ as

$$\frac{d}{dx}(uv).$$

will we get the same answer as if we do

$$\frac{du}{dx}\times\frac{du}{dx}\ ?$$

To answer this consider $f(x) = x^3$. Let us split this into two separate functions $u(x) = x$ and $v(x) = x^2$. Let us now find the derivative of the product $u.v = x.x^2$. We know the derivative of $f(x)$ is

$$\frac{d}{dx}\big(f(x)\big) = 3x^2.$$

But let us see what happens if we keep the two functions separate. Is it then simply a matter of multiplying the separate derivatives? No:

$$\frac{d}{dx}(u.v) = \frac{d}{dx}(x.x^2) = \frac{d}{dx}(x).\frac{d}{dx}(x^2) = 1\times 2x = 2x$$

which is not the correct answer. So here we see for the first time that the usual distributive process of an arithmetic operation does not carry over to differentiation.

Suppose, however, that we differentiate as follow:

$$\frac{d}{dx}(x.x^2) = x.\frac{d}{dx}(x^2) + x^2\frac{d}{dx}(x).$$

What I am doing here is to alternate between differentiating and not differentiating either term. In other words I am doing

$$\frac{d}{dx}(first\times second) = first.\frac{d}{dx}(second) + second.\frac{d}{dx}(first).$$

When we do this we get

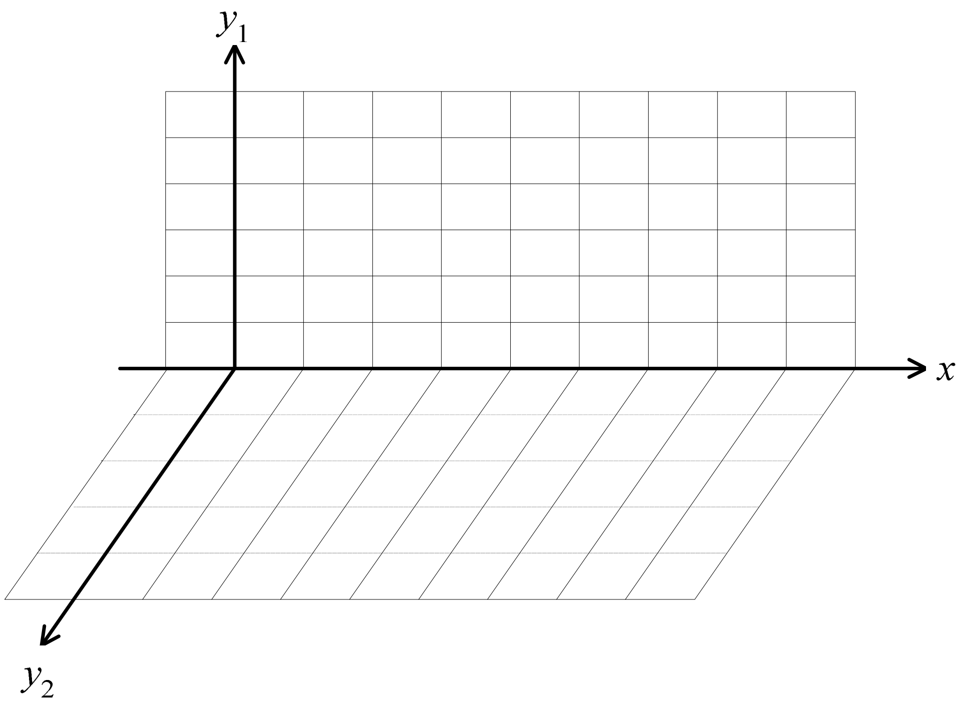
$$\frac{d}{dx}(x.x^2) = x.(2x) + x^2.1 = 3x^2$$

which is the answer we want. Was this a coincidence? No.

1.4.1 An informal derivation of the product rule

To see why the above process works we need to go back to the definition of the derivative, and prove it from 1st principles. Before we do this we will go through a conceptual approach to getting the rule for differentiating a product of two functions. This conceptual approach will involve us studying areas set up by the functions $u(x)$ and v(x). At the beginning it may not be obvious why we consider areas, but I will try to make this as clear as possible as we go through the ideas below.

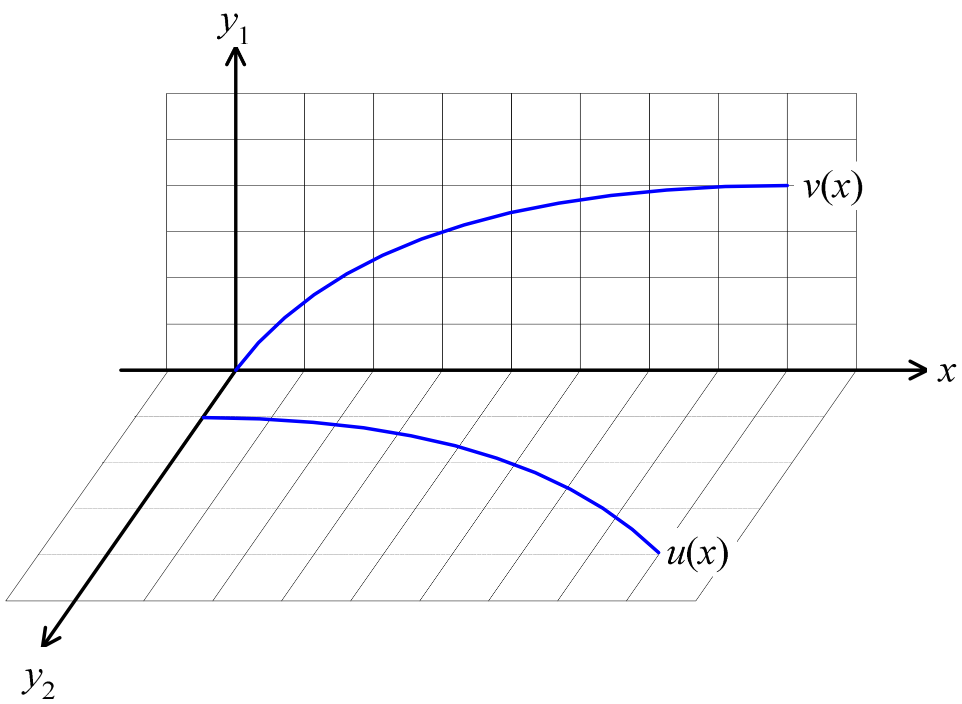
In order to understand the concept of the product rule for derivatives we will need to set up a new coordinate system. As such, consider two standard (x, y) coordinate systems, one called (x, y_1) and the other called (x, y_2), where both of these coordinate systems share the x-axis, and such that these two coordinate systems are perpendicular to each other:



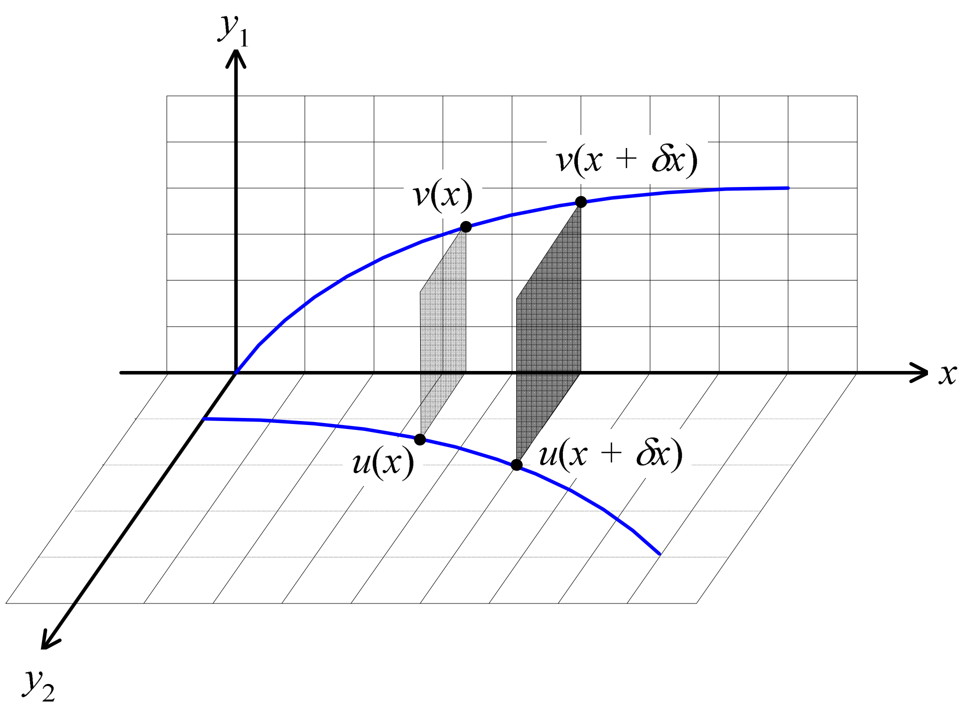
(All coordinate system diagrams adapted from

http://jwilson.coe.uga.edu/EMAT6680/Horst/product/product.html, the original source being *Derivative of a product: Exploration 2 of concepts of calculus for middle school teachers*, 1990, *Middle School Math Project*, The Math Learning Center, Salem, OR, USA).

Onto (x, y_1) we draw $v(x)$, and onto (x, y_2) we draw $u(x)$.



At any point x draw a line from the x-axis to $u(x)$ and to $v(x)$. The product $f(x) = u(x).v(x)$ can then be represented geometrically as a rectangle (shown in light grey below) so that $f(x)$ represents an area. We can then draw another rectangle at some point $x + \delta x$ (shown in dark grey below). This represents the area $u(x + \delta x).v(x + \delta x)$:



To make things easier to visualise let us look at the above rectangles from two different perspectives. In one perspective we look at the two rectangles from an angle, and in the other perspective we superpose the two rectangles and look at them face-on in 2-D.

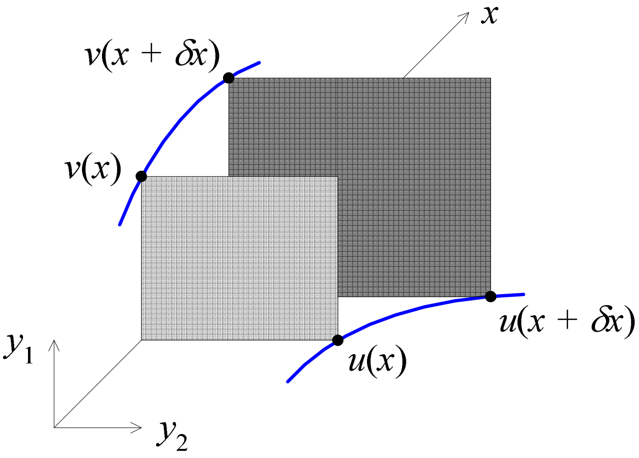


diagram 1

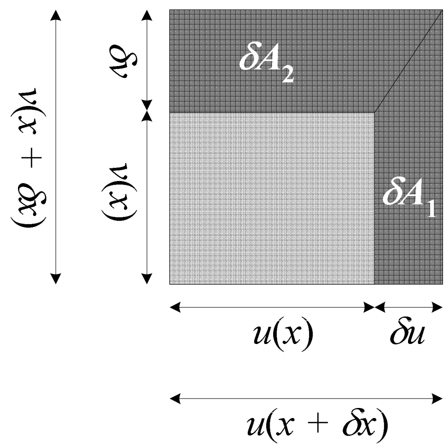


diagram 2

Now imagine, if you can, $u(x + \delta x)$ approaching $u(x)$ as $\delta x \rightarrow 0$, but $v(x + \delta x)$ staying where it is.

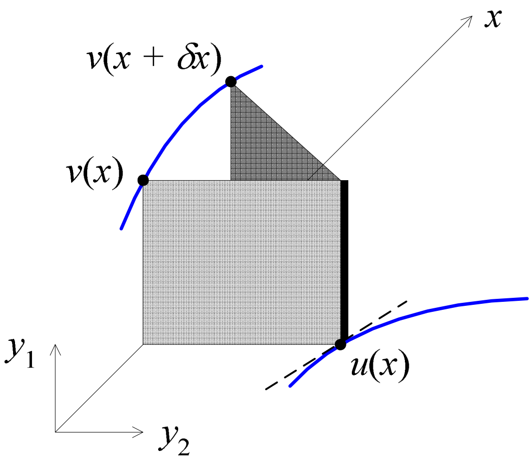


diagram 3

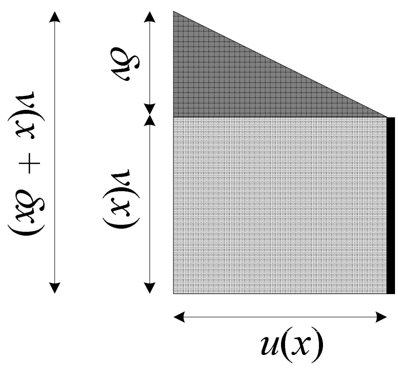


diagram 4

In diagram 2 we have $\delta A_1 \approx v(x + \delta x).\delta u$ since the thinner the vertical dark grey strip, the more it resembles a rectangle. If we now divide δA_1 by δx we get

$$\frac{\delta A_1}{\delta x} \approx v(x + \delta x)\frac{\delta u}{\delta x}.$$

As $\delta x \rightarrow 0$ the secant that would be going through $u(x)$ and $u(x + \delta x)$ becomes a tangent at $u(x)$ and area δA_1 becomes infinitely thin (this being represented by the thick black vertical line).

Hence as $\delta x \to 0$ we then get

$$\frac{dA_1}{dx} = v(x)\frac{du}{dx}.$$

Now imagine $v(x + \delta x)$ approaching $v(x)$ as $\delta x \to 0$, but $u(x + \delta x)$ staying where it is:

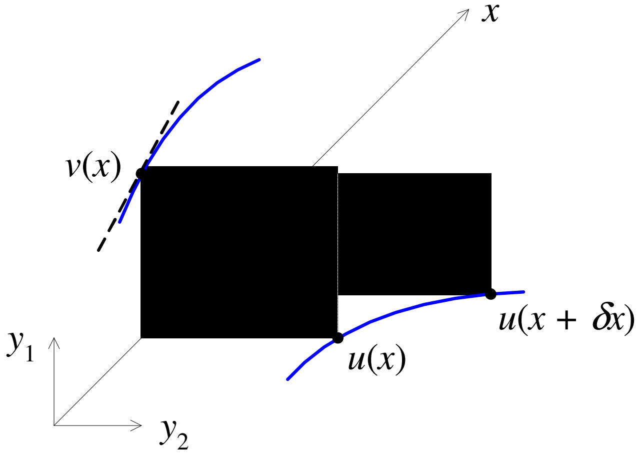


diagram 4

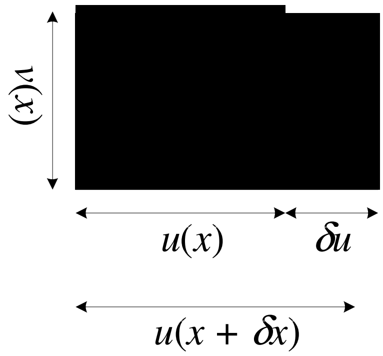


diagram 6

Since, in diagram 2, we have $\delta A_2 \approx u(x+\delta x).\delta v$ (again by the strip being approximately a rectangle), if we divide δA_2 by δx we get

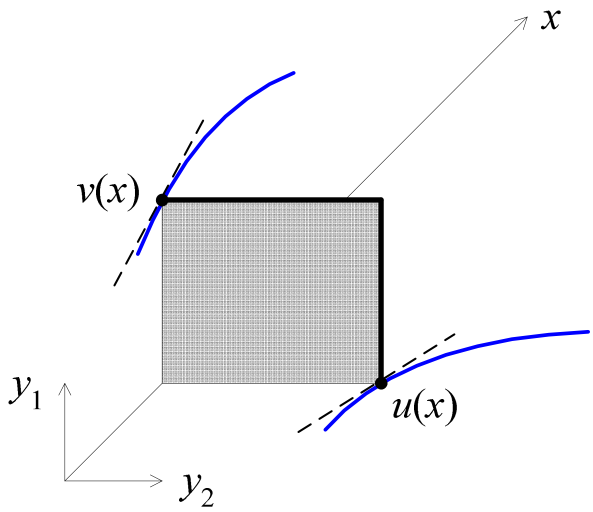
$$\frac{\delta A_2}{\delta x} \approx u(x+\delta x)\frac{\delta v}{\delta x}.$$

As $\delta x \to 0$ the secant that would be going through $v(x)$ and $v(x + \delta x)$ becomes a tangent at $v(x)$ and area δA_2 becomes infinitely thin (this being represented by the thick black horizontal line).

Hence as $\delta x \to 0$ we then get

$$\frac{dA_2}{dx} = u(x)\frac{dv}{dx}.$$

However, as $\delta x \to 0$ both $u(x + \delta x)$ and $v(x + \delta x)$ are moving respectively towards $u(x)$ and $v(x)$ at the same time. In other words, the dark grey rectangle shrinks back to the original size of the light grey rectangle, and we recover the original area $A(x)$:

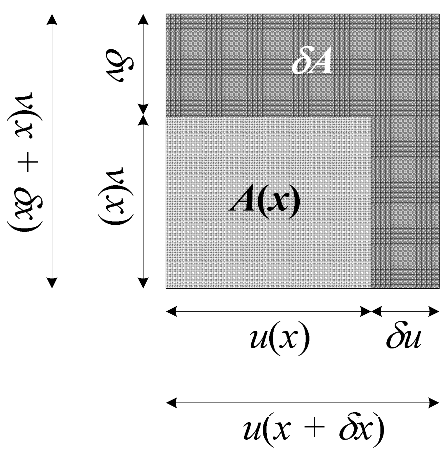


The combined effects of these movements makes the derivative of $f(x) = u(x).v(x)$ a combination of infinitely small areas $u.dv/dx$ and $v.du/dx$, where each area contains a tangent (i.e. derivative) effect at $u(x)$ and $v(x)$.

So in treating the product $u(x).v(x)$ as an area, expanding this area/product by a small but still finite amount, and then shrinking this area/product to one infinitely close to the original area we have been able to uncover the fact that there are tangents associated with these areas, and thence arrive at the derivative of $u(x).v(x)$.

Let us now return to the superposed face-on diagram of the two rectangular areas. The small (light grey) rectangle represents the area $A(x) = u(x).v(x)$ and the big (dark grey) rectangle represents the area $A(x + \delta x) = u(x + \delta x).v(x + \delta x)$ or $A(x) + \delta A = (u(x) + \delta u).(v(x) + \delta v)$.

We now need to study what happens to certain areas in this diagram, as well as what happens to relevant secants through $u(x)$ and $u(x + \delta x)$, and $v(x)$ and $v(x + \delta x)$, as $\delta x \rightarrow 0$.



We want to find out the rate of change of area as δx approaches 0. This will then tell us the rate of change of $A(x)$ at the point x itself. Since rate of change is given mathematically by taking the limit of $\delta A/\delta x$ we first need to find an expression for δA, then divide by δx and then apply the limit.

We can find expressions for δA with the help of the diagrams below. Here we chop up the dark grey part of the diagram in several different ways, each one leading to a different expression for δA:

δA = big rectangle – small rectangle

$= u(x+\delta x).\,v(x+\delta x) - u(x).\,v(x)$

δA = top dark grey rectangle
+ vertical dark grey rectangle

$= \delta v.\,u(x) + \delta u.\,v(x+\delta x)$

δA = top dark grey rectangle
+ vertical dark grey rectangle
+ small white rectangle

$= \delta v.\,u(x) + \delta u.\,v(x) + \delta u.\,\delta v$

δA = top dark grey rectangle
+ vertical dark grey rectangle
– small corner black rectangle
(overlap from the top and
vertical rectangles)

$= \delta v.\,u(x+\delta x) + \delta u.\,v(x+\delta x)$
$-\delta u.\,\delta v$

Table 1

Choosing the simplest of the four expressions we then divide δA by δx to give $\delta A/\delta x$, after which we take limits to get dA/dx. Hence, using $\delta A = \delta v.u(x) + \delta u.v(x+\delta x)$ leads to

$$\frac{\delta A}{\delta x} = \frac{\delta v.u(x) + \delta u.v(x+\delta x)}{\delta x},$$

$$= \frac{\delta v}{\delta x}.u(x) + \frac{\delta u}{\delta x}.v(x+\delta x).$$

Since A represents $u(x).v(x)$ we can replace it by our original notation $f(x)$. Then $\delta x \to 0$ we get

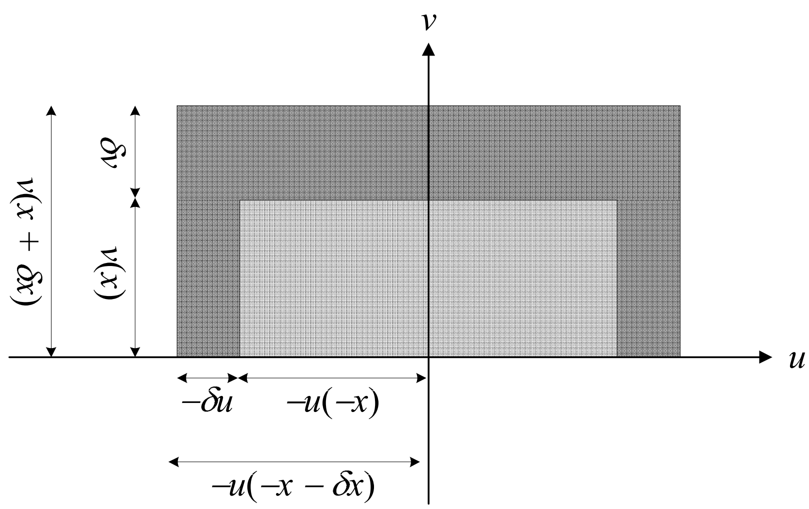
$$\frac{df}{dx} = u(x)\frac{dv}{dx} + v(x)\frac{du}{dx}. \quad (9)$$

In terms of the process of differentiation this means that when differentiating the product of two functions we keep the first function as is and differentiate the second function, and then keep the second function as is and differentiate the first function, and then add the results.

1.4.2 A note on the proof of the product rule of the previous section

In the previous section I showed the use of one of the expressions of table 1 as a way of helping us derive the product rule. However, the expressions in table 1 are only true in the case where the function is said to be positive and increasing (i.e. $f(x)$ lies in quadrant 1 and gets bigger as x get bigger). If the function were decreasing and/or negative and/or in another quadrant new expressions would have to be developed, and even then they would results in a correct product rule.

To see one example of this consider $v(x)$ being positive but $u(x)$ being negative. The geometric representation of the two squares would then look like this:



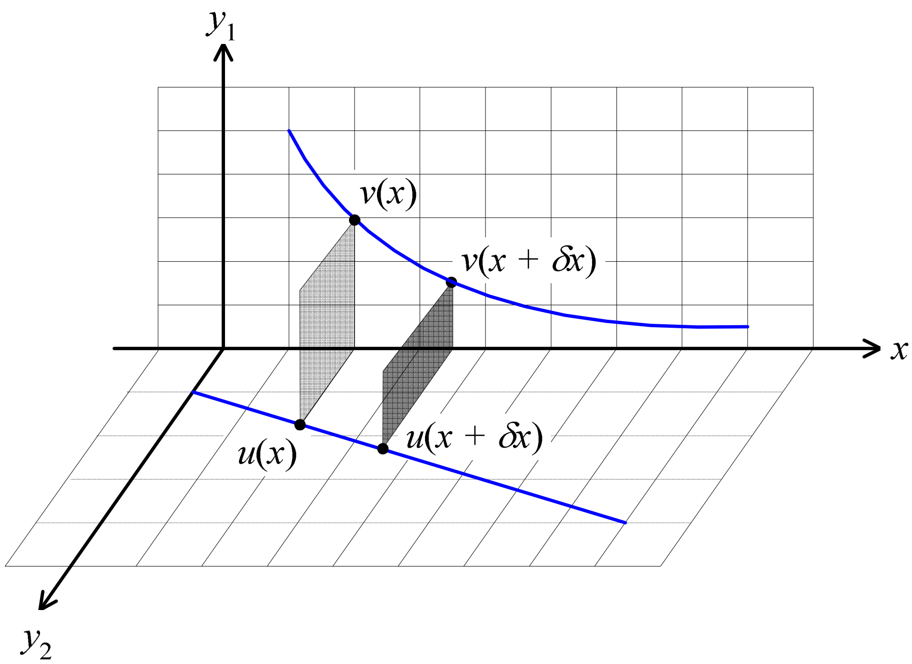
For this situation $\delta A = \delta v.\left(-u(-x)\right) - \delta u.\, v(x+\delta x)$. Dividing by δx and letting $\delta x \to 0$ we get

$$\frac{dA}{dx} = -u(-x)\frac{dv}{dx} - \; v(x)\frac{du}{dx}.$$

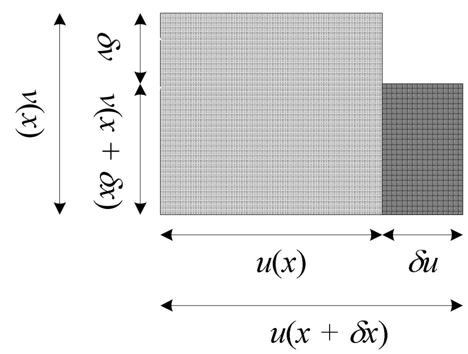
which is not the correct version of the product rule.

The other issue with the proof of the product rule shown in the previous section is that diagram 1 only shows the case when *f*(*x*) = *u*(*x*).*v*(*x*) is an increasing function (i.e. *f*(*x*) gets bigger and bigger as *x* gets bigger and bigger). It is because of this that we have the blue rectangle being bigger than the pink rectangle. But what if the blue rectangle was smaller than the pink rectangle? What if, when we move from *x* to *x* + δx we end up with a smaller rectangle than we originaly had?

As an example consider the diagram below where *v*(*x*) is decreasing as *x* increases:



Supperposing the two rectangles we have



In this case we have lost some area along $v(x)$ and gained some area along $u(x)$. So out net change in area δA is given by

$$\delta A = -\delta v. u(x) + \delta u. v(x + \delta x)\ .$$

Dividing by δx and letting $\delta x \to 0$ we get

$$\frac{dA}{dx} = -u(x)\frac{dv}{dx} + \ v(x)\frac{du}{dx},$$

which is not the correct version of the product rule.

1.4.3 The formal proof of the product rule

Although I have said that (9) is the correct product rule how do we know this? And how do we go about deriving/proving the product rule that applies to all functions, whether increasing or decreasing, positive or negative?

Well, we always start from 1st principles. Then we use whatever algebra necessary which doesn't rely on whether the separate parts of $f(x)$ are increasing/decreasing or positive/ negative.

Therefore we start as such:

$$\frac{d}{dx}\big(u(x). v(x)\big) = \lim_{\delta x \to 0} \frac{u(x + \delta x). v(x + \delta x) - u(x). v(x)}{\delta x}\ . \qquad (10)$$

We now need to do some relevant algebra so that we can recover the differences $u(x + \delta x) - u(x)$ and $v(x + \delta x) - v(x)$, because we know that when we divide these differences by δx and apply the limit (i.e. let $\delta x \to 0$) we get a derivative.

To recover these differences we do some algebra rather than some geometry. Specifically we add and subtract $v(x).u(x + \delta x)$ to the numerator of (10). Doing this will allow us to recover the differences $u(x + \delta x) - u(x)$ and $v(x + \delta x) - v(x)$ which will give us the derivatives we want. Hence

$$\frac{d}{dx}\big(u(x). v(x)\big) = \lim_{\delta x \to 0} \frac{u(x + \delta x). v(x + \delta x) - v(x)u(x + \delta x) + v(x)u(x + \delta x) - u(x). v(x)}{\delta x}.$$

The advantage of doing this over the geometric way is that we have placed no restrictions on what direction (increasing/decreasing or positive/negative) $u(x)$ or $v(x)$ should take. In this case the formula we end up with will apply to all $u(x)$ or $v(x)$ whatever their shape and direction.

Now, grouping terms so as to get the aformentioned differences we get

$$\frac{d}{dx}\big(u(x).v(x)\big) = \lim_{\delta x\to 0} u(x+\delta x)\frac{v(x+\delta x)-v(x)}{\delta x} + v(x)\lim_{\delta x\to 0}\frac{u(x+\delta x)-u(x)}{\delta x}.$$

As $\delta x \to 0$, so $u(x + \delta x) \to u(x)$, we get the product rule

$$\frac{d}{dx}\big(u(x).v(x)\big) = u(x)\frac{dv(x)}{dx} + v(x)\frac{du(x)}{dx}. \tag{11}$$

So now we see that $d(u.v)/dx \neq (du/dx)(dv/dx)$, and why it is important to 'go back to the beginning' and prove everything from first principles.

And to say again, the proof above shows that we end up with a numerator which is derived purely algebraically (via the addition and subtraction of a suitable term) rather than geometrically (i.e. the increase of decrease of an area).

The process of the product rule can be represented visually as any one of the diagrams below:

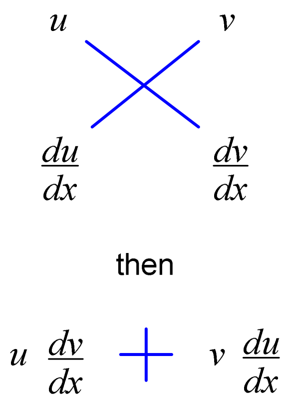


diagram 1

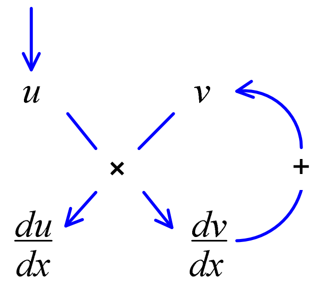


diagram 2

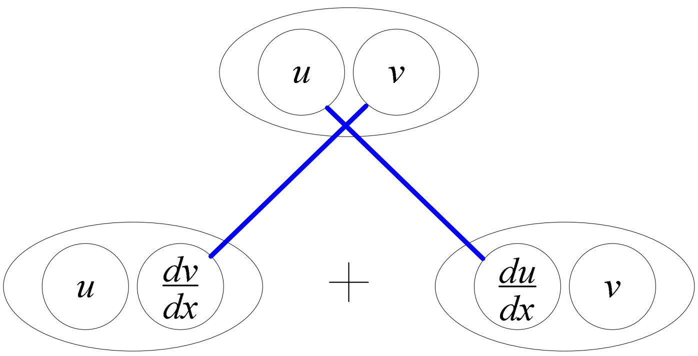


diagram 3

(I adapted this diag from http://www.ballooncalculus.org/examples/reference.html#tanDiff)

Examples

1) If $y = \sin x \cos x$ then

$$\frac{dy}{dx} = \sin x \, . \frac{d}{dx}(\cos x) + \cos x \, . \frac{d}{dx}(\sin x) \ ,$$

$$= -\sin^2 x + \cos^2 x \ ,$$

$$= \cos 2x \, .$$

and

$$\frac{d^2y}{dx^2} = -\sin x \, . \frac{d}{dx}(\sin x) - \sin x \, . \frac{d}{dx}(\sin x)$$

$$+ \cos x \, . \frac{d}{dx}(\cos x) + \cos x \, . \frac{d}{dx}(\cos x) \, .$$

$$= -2 \sin x \cos x - 2 \sin x \cos x \ ,$$

$$-2 \sin 2x \, .$$

2) If $f(x) = x . \ln x$ then

$$\frac{df}{dx} = x . \frac{d}{dx}(\ln x) + \ln x \, . \frac{d}{dx}(x) \ ,$$

$$= 1 + \ln x \, .$$

3) If $y = x^2 e^x + 1 + 1/x^2$ then

$$\frac{dy}{dx} = x^2 . \frac{d}{dx}(e^x) + e^x . \frac{d}{dx}(x^2) + \frac{d}{dx}(1) + \frac{d}{dx}(x^{-2}) \ ,$$

$$= x^2 e^x + 2x . e^x - 2x^{-3} \, .$$

4) If $f(t) = \left(1 + \sqrt{t^3}\right)\left(x^2 - 2 . \sqrt[3]{t}\right)$ then

$$\frac{df}{dt} = \left(1 + \sqrt{t^2}\right) . \frac{d}{dt}\left(x^2 - 2 . \sqrt[3]{t}\right) + \left(x^2 - 2 . \sqrt[3]{t}\right) . \frac{d}{dt}\left(1 + \sqrt{t^3}\right) \ ,$$

$$= \left(1 + \sqrt{t^2}\right)\left(2x - \frac{2}{3} . \frac{1}{\sqrt[3]{t^2}}\right) + \left(x^2 - 2 . \sqrt[3]{t}\right)\left(\frac{3}{2}\sqrt{t}\right) .$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.

1.4.4 A proof of the power rule for positive integers using the product rule

In *Studies on Differentiation I* we saw the proof that $d(x^n)/dx = n.x^{n-1}$. Here we will see how the product rule can be used as part of the steps in the proof of this, the proof method being that of mathematical induction. Note that this proof is only valid when n is a positive integer since induction is only valid in these cases.

So, given $f(x) = x^n$, where n is a positive integer,

- let $n = 1$. Therefore $f(x) = x^1 = x$. Then

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{(x + \delta x) - x}{\delta x} = \lim_{\delta x \to 0} \frac{\delta x}{\delta x} = 1$$

  which is the same as $df/dx = 1.x^0 = 1$. Therefore we have shown that $df/dx = n.x^{n-1}$ for $n = 1$;
- let $n = k$. Therefore $f(x) = x^k$. Then, assume that $df/dx = d(x^k)/dx = k.x^{k-1}$;
- when $n = k + 1$ we have $f(x) = x^{k+1}$ and therefore

$$\frac{d(x^{k+1})}{dx} = \frac{d(x.x^n)}{dx}.$$

  By the product rule we get

$$\frac{d(x^{k+1})}{dx} = x.\frac{d(x^k)}{dx} + x^k\frac{d(x)}{dx}.$$

  Using our assumption we get

$$\frac{d(x^{k+1})}{dx} = x(k.x^{k-1}) + x^n(1) ,$$

$$= (k+1).x^k$$

  which is what we wanted to show. Therefore

$$\frac{df}{dx} = n.x^{n-1}$$

  is true for all positive integer n.

1.4.5 The product rule for more that two functions

Section 1.4.3 showed us how to differentiate the product of two functions. How would we differentiate the product of three functions? How would we differentiate $f(x) = xe^x \sin x$?

What we effectively have here is a situation where $f(x) = u(x).v(x).w(x)$. In this case we could treat two of these function as one combined function, i.e. $f(x) = U(x).w(x)$ where $U(x) = u(x)v(x)$, and do the product rule on $U(x).w(x)$, then on $u(x).v(x)$.

Hence we can do

$$\begin{aligned}\frac{d}{dx}(xe^x \sin x) &= x\frac{d}{dx}(e^x \sin x) + e^x \sin x \frac{d}{dx}(x)\,, \\ &= x\left(e^x \frac{d}{dx}(\sin x) + \sin x \frac{d}{dx}(e^x)\right) + e^x \sin x \frac{d}{dx}(x)\,, \\ &= xe^x \cos\, x + xe^x \sin x + e^x \sin x\end{aligned}$$

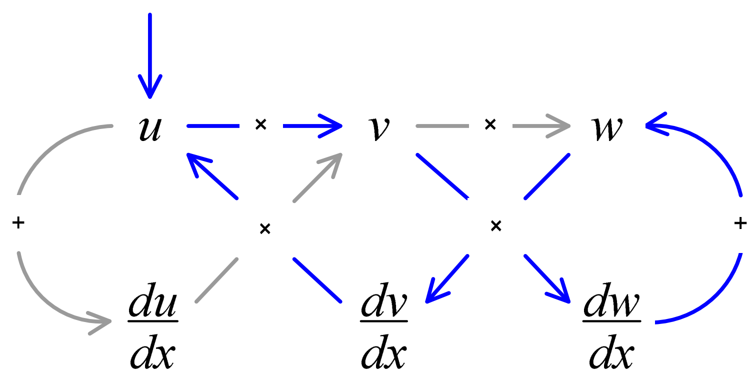
In general we can therefore say that if $f(x) = u.v.w$ then

$$\begin{aligned}\frac{d}{dx}(uvw) &= u\frac{d}{dx}(vw) + vw\frac{du}{dx}, \\ &= u\left(v\frac{dw}{dx} + w\frac{dv}{dx}\right) + vw\frac{du}{dx}, \\ &= u.v\frac{dw}{dx} + u\frac{dv}{dx}w \\ &\qquad + \frac{du}{dx}v.w\,,\end{aligned}$$

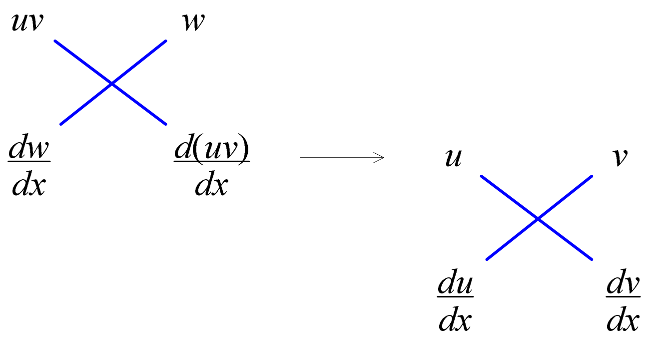
or, in simpler notation,

$$(u.v.w)' = u.v.w' + u.v'.w + u'.v.w \tag{12}$$

The process of this can be represented visually as

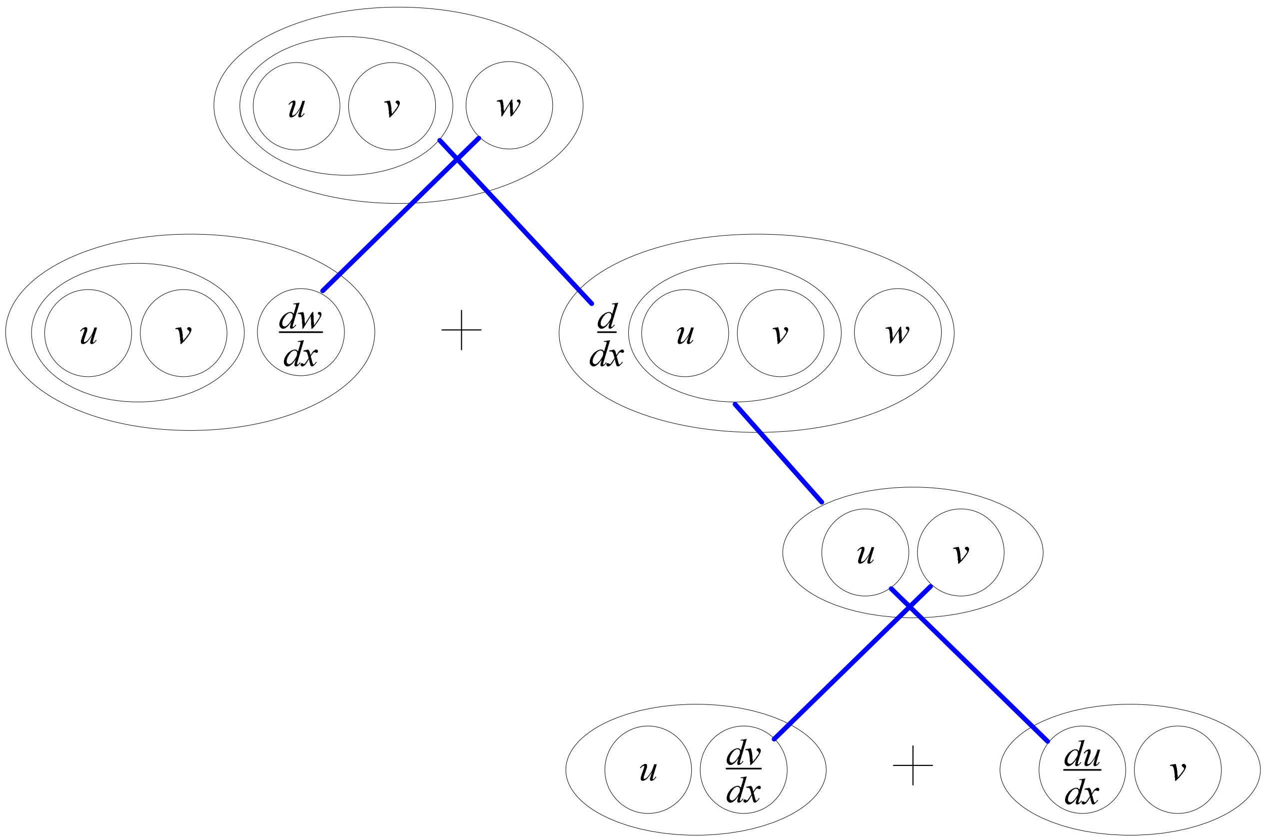


(where you follow the blue lines first (starting at the vertical arrow above u) then continue onto the grey lines), or as



then $u\,v\,\dfrac{dw}{dx} + u\,\dfrac{dv}{dx}\,w + \dfrac{du}{dx}\,v\,w$

or as



Examples

1) If $y = x^2 e^x \sin x$ then

$$\begin{aligned} \frac{dy}{dx} &= x^2 \frac{d}{dx}(e^x \sin x) + e^x \sin x \frac{d}{dx}(x^2)\,, \\ &= x^2\left(e^x \frac{d}{dx}(\sin x) + \sin x \frac{d}{dx}(e^x)\right) + e^x \sin x \frac{d}{dx}(x^2)\,, \\ &= x^2 e^x \cos x + x^2 e^x \sin x + 2x e^x \sin x\,. \end{aligned}$$

2) If $f(x) = x \sin x \cos x$ then

$$\begin{aligned} \frac{df}{dx} &= x \frac{d}{dx}(\sin x \cos x) + \sin x \cos x \frac{d}{dx}(x)\,, \\ &= x\left(\sin x \frac{d}{dx}(\cos x) + \cos x \frac{d}{dx}(\sin x)\right) + \sin x \cos x \frac{d}{dx}(x)\,, \\ &= -x \sin^2 x + x \cos^2 x + \sin x \cos x\,. \end{aligned}$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.

This idea can be extended to find the derivative of the product of n functions

$$f(x) = u_1(x)u_2(x) \ldots u_n(x),$$

namely

$$\frac{d}{dx}(u_1 u_2 \ldots u_n) = u_1 u_2 \ldots u_{n-1}\frac{du_n}{dx} + u_1 u_2 \ldots u_{n-2} u_n \frac{du_{n-1}}{dx} + \cdots$$

$$\cdots + u_1 u_3 \ldots u_{n-1} u_n \frac{du_2}{dx} + u_2 u_3 \ldots u_{n-1} u_n \frac{du_1}{dx}\,.$$

It would be quite complicated (if at all possible) to show this in diagram form. However, this is the strength of algebra: it allows us to generalise the form of something (in this case, a formula) to n-dimensions when it is not possible (yet) to do diagrammatically.

1.4.6 Repeated differentiation of f(x) = u.v – Leibniz's rule

Returning to the product rule for $f(x) = u(x).v(x)$ one question we can ask is, What happens if we differentiate this function three, four, five, … times? Applying the product rule four times to $f(x)$ and simplifying we get:

$$f' = uv' + vu' \,,$$

$$f'' = uv'' + 2u'v' + u''v \,,$$

$$f''' = uv''' + 3u'v'' + 3u''v' + u'''v \,,$$

$$f'''' = uv'''' + 4u'v''' + 6u''v'' + 4u'''v' + u''''v \,.$$

There are two patterns we can notice from this sequence of derivatives: i) the coefficients run like Pascal's triangle of numbers (in fact they run according to the binomial coefficients), and ii) the derivatives "decrease" on *v* and "increase" on *u*. All of this is not a coincidence.

In general if we differentiate our function *n* times (and using the notation $f^{(n)}(x)$ to mean this) it looks like we will get

$$f^{(n)}(x) = uv^{(n)} + n.u'v^{(n-1)} + \frac{n(n-1)}{2!}u''v^{(n-2)} + \frac{n(n-1)(n-2)}{3!}u'''v^{(n-3)} + \cdots$$
$$\ldots + \frac{n(n-1)}{2!}u^{(n-2)}v'' + n.u^{(n-1)}v' + uv^{(n)} \,. \qquad (13)$$

This is in fact true and is called *Leibniz's rule* for differentiation. We will go through the proof of this formula (using induction) in the next section.

As an example let us go through differentiating $y = e^x \sin x$ four times. We could do this manually by finding the first derivative, then the second derivative, then the third derivative, then the fourth derivative …

$$\frac{dy}{dx} = e^x \sin(x) + e^x \cos(x) \,, \qquad \frac{d^2y}{dx^2} = 2e^x \cos(x) \,,$$

$$\frac{d^3y}{dx^3} = 2e^x \cos(x) - 2e^x \sin(x) \,, \qquad \frac{d^4y}{dx^4} = -4e^x \sin(x) \,.$$

… or we could apply (13) directly:

$$(e^x \sin x)'''' = e^x(\sin x)'''' + 4.(e^x)'(\sin x)''' + \frac{4.3}{2!}(e^x)''(\sin x)''$$

$$+\frac{4.3.2}{3!}(e^x)'''(\sin x)' + \frac{4.3.2}{4!}(e^x)''''\sin x\,,$$

which gives

$$y'''' = e^x\sin(x) - 4e^x\cos(x) - 6e^x\sin(x) + 4e^x\cos(x) + e^x\sin(x)\,,$$

and which simplifies to

$$y'''' = -4e^x\sin(x).$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.

1.4.7 Proof of Leibniz's product rule for differentiation: ***To come***

1.4.8 When does d(u.v)/dx = (du/dx)(dv/dx)?

This section is adapted from *When does* $(fg)' = f'g'$?, L. Maharam and E. P. Shaughnessy, in *The To Year College Mathematics Journal*, Vol 7, No. 1. (Feb 1976), pp. 38 – 39. For this section you will need to know some integration.

Given the definition of the product rule we know that $d(\sin x \,.\cos x)/dx \neq d(\sin x)/dx \times d(\cos x)/dx$ since

$$\frac{d}{dx}(\sin x . \cos x) = \cos^2 x - \sin^2 x$$

and

$$\frac{d}{dx}(\sin x) \times \frac{d}{dx}(\cos x) = -\cos x \times \sin x \,.$$

So, given two function $f(x)$ and $g(x)$ is it ever possible for the derivative of a product $(fg)'$ equals the product of the separate derivatives $f'g'$? In other words, can we ever have

$$\frac{d}{dx}(f.g) = \frac{df}{dx}.\frac{dg}{dx} \text{ ?}$$

Well, we know that

$$\frac{d}{dx}(f.g) = f.\frac{dg}{dx} + g.\frac{df}{dx}$$

so the right hand side of the two equations above would have to be equal. Therefore, if we do this (and simplify the notation to make it easier to read) we get

$$f'g + fg' = f'g'.$$

Our aim now is to use algebra to transform this expression into something we know how to integrate. In this case we can write the above as

$$f'(g' - g) = fg'. \qquad (*)$$

At this point we could try to integrate: $\int f'(g' - g)\,dx = \int f.g'dx$, leading to $\int f'g' - f'gdx = \int f.g'dx$.

We could then use by-parts to try to perform the integral on the left hand side and right hand side, but this is quite laborious and there is no guarantee of coming up with a viable result.

Instead let us bring all the functions of $f(x)$ on one side of the equals (sign by division) and all the functions of $g(x)$ on the other side of the equals sign (again by division).

Doing this we get

$$\frac{f'}{f} = \frac{g'}{(g'-g)}.$$

Let us now integrate:

$$\int \frac{f'}{f}\,dx = \int \frac{g'}{(g'-g)}\,dx\,.$$

If you remember your integration you will see that left hand side can be integrated directly as a log function:

$$\ln f + k = \int \frac{g'}{(g'-g)}\,dx\,.$$

where k is the constant of integration. Since k is a number it is going to be easier in the following steps if we rewrite it as ln c. So now we have

$$\begin{aligned} \ln f + \ln c &= \int \frac{g'}{(g'-g)}\,dx\,, \\ \ln(cf) &= \int \frac{g'}{(g'-g)}\,dx\,, \\ f &= c.\exp\int \frac{g'}{(g'-g)}\,dx\,. \qquad (**) \end{aligned}$$

So the reason it was useful to make k = ln c was because it allowed us to use log rules to combine a sum of logs into the log of a product. We can then take the exponential of both sides. Expression (**) is what we are looking for. What it means is that the only way $(f.g)' = f'g'$ is if $f(x)$ satisfies equation (**).

To see what this means in practice consider $f(x) = 1/(x-1)$ and $g(x) = x$. Then

$$\begin{aligned} (fg)' &= f'g' &&= -\frac{1}{(x-1)^2}. \\ (fg)' &= fg' + f'g &&= \left(-\frac{1}{(x-1)^2}\right)(x) + \left(\frac{1}{x-1}\right)(1)\,, \\ & &&= \frac{-x(x-1)+(x-1)^2}{(x-1)^3}, \\ & &&= \frac{(x-1)(-x+x-1)}{(x-1)^3}, \\ & &&= -\frac{1}{(x-1)^2}. \end{aligned}$$

So $f(x)$ and $g(x)$ satisfy the incorrect product rule.

The way to create the correct $f(x)$ which will work for the incorrect product rule is simply to choose any function as your $g(x)$ and perform the integral (**). The only issue will be in how easy it is to perform (**). For example if we choose $g(x) = \sin x$ then

$$\int \frac{g'}{(g' - g)}\, dx = \int \frac{\cos x}{\cos x - \sin x}\, dx$$

which may or may not be integratable.

Furthermore, there is one problem with equation (**). If we choose $g(x) = e^x$ then the denominator of (**) is zero. So (**) does not work for all functions. See the paper by L. Maharam and E. P. Shaughnessy for more details in this case.

Three pairs of functions for which the incorrect product rule gives the correct answer are:

$$f(x) = \sin x + \cos x\,, \qquad g(x) = \sqrt{e^x \csc x}\,,$$

$$f(x) = x.e^x\,, \qquad g(x) = e^{x^2/(2+x)},$$

$$f(x) = e^{-\cot(x/2)}\,, \qquad g(x) = \sec x + \tan x\,.$$

For how one can generalise the idea above to higher derivatives see *The Naïve Product Rule for Derivatives*, Carter C. Gay, Akalu Tefera and Aklilu Zeleke, *The College Mathematics Journal*, Vol. 39, No. 2 (Mar., 2008), pp. 145-148.

1.5 On the quotient rule for differentiation: $d(u/v)/dx$

1.5.1 An informal derivation of the quotient rule

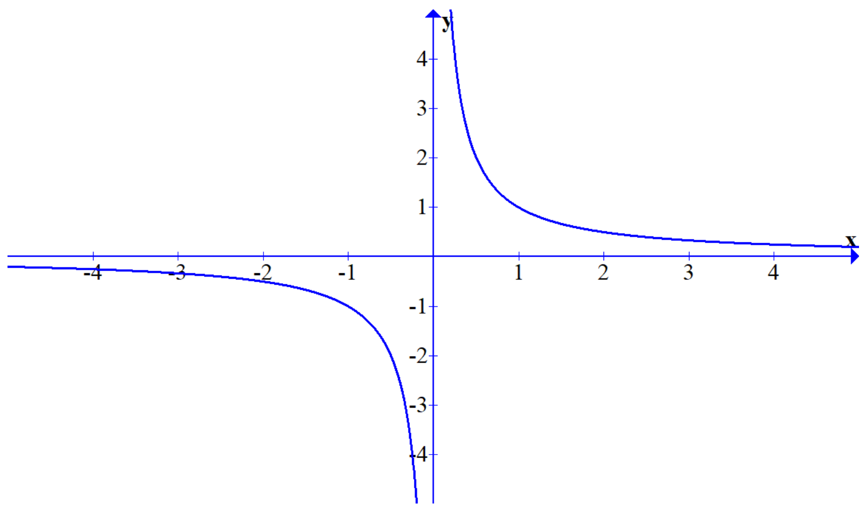
Now that we have seen that differentiating the product of a function made up of two separate "sub-functions" functions isn't as obvious as we might have thought, we need to ask ourselves how we differentiate a functions when one of its sub-fucntions is is divided by another of its subfunctions.

So, if we have $f(x) = u(x)/v(x)$, and we want to differentiate $f(x)$ what will we get? Do we really think that we will get

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{du/dx}{dv/dx} \text{ ?}$$

From what we have done in the previous section it seem very unlikely.

To see what happens consider $f(x) = 1/x$, the graph of which is shown below.



If we assume the above equation is true then we have

$$\frac{d}{dx}\left(\frac{1}{x}\right) = \frac{\frac{d}{dx}(1)}{\frac{d}{dx}(x)} = 0\,.$$

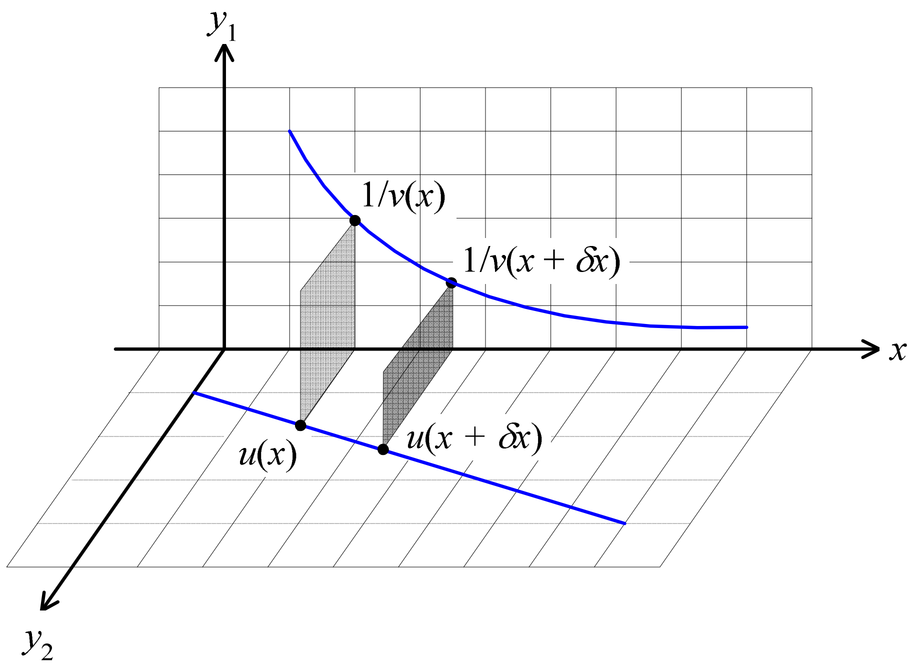
This says that the curve of $1/x$ has no slope of gradient, implying that the curve is a horizontal line (just like y = 1 or y = 2, ... are horizontal lines). But this is clearly not true. So our assumption that the above equation is the derivative of a fraction is incorrect. The same thing happens if we try differentiating $f(x) = x^2/x$.

If we take our two sub-functions to be $u(x) = x^2$ and $v(x) = x$ and apply the derivative equation above we get

$$\frac{d}{dx}\left(\frac{x^2}{x}\right) = \frac{\frac{d}{dx}(x^2)}{\frac{d}{dx}(x)} = \frac{2x}{1} = 2x\ .$$

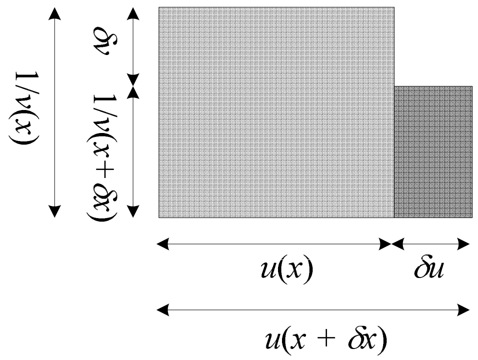
But if we simplify $f(x)$ to $f(x) = x$ we know the derivative to be $df/dx = 1$.

To see how to get the correct quotient rule let us again set up two (x, y) coordinate sytems both sharing the same x-axis, as we did in section 1.4.1. Rewriting $f(x)$ as $f(x) = u(x)\times 1/v(x)$, and drawing $u(x)$ and $1/v(x)$ on our coordinate system, will then enable us to consider the area of rectangles, one of height $1/v(x)$ and length $u(x)$, and another of height $1/v(x + \delta x)$ and length $u(x + \delta x)$, as shown below:



To make things easier to visualise we will superpose the two rectangles and look at them in 2-D. The pink rectangle represents the area $A(x) = u(x)\times 1/v(x)$ and the blue rectangle represents the area $A(x + \delta x) = A(x) + \delta A = u(x + \delta x)\times 1/v(x + \delta x)$ as shown in the diagram on the next page.

We want to find out the rate of change of area as δx approaches 0. This will then tell us the rate of change of $A(x)$ at the point x itself. Since rate of change is given mathematically by taking the limit of $\delta A/\delta x$ we need to find an expression for δA from the diagram above. This can be done by considering the different parts of the diagram, leading to an expression for δA.



So our original area is

$$A(x) = u(x).\frac{1}{v(x)}$$

Our new expanded area is

$$A(x+\delta x) = u(x+\delta x).\frac{1}{v(x+\delta x)}.$$

The change in area, δA, as a result of moving from x to $x + \delta x$ is

$$\delta A = A(x+\delta x) - A(x) = u(x+\delta x)\frac{1}{v(x+\delta x)} - u(x)\frac{1}{v(x)}.$$

Simplifying gives

$$\begin{aligned}
\delta A &= (u+\delta u).\frac{1}{v+\delta v} - u.\frac{1}{v}, \\
&= \frac{(u+\delta u) - u.\frac{1}{v}(v+\delta v)}{v+\delta v}, \\
&= \frac{\delta u - u.\frac{1}{v}\delta v}{v+\delta v}, \\
&= \frac{v.\delta u - u.\delta v}{v.(v+\delta v)}.
\end{aligned}$$

We can now divide δA by δx to give $\delta A/\delta x$, after which we take limits to get dA/dx. Hence,

$$\begin{aligned}
\frac{\delta A}{\delta x} &= \frac{1}{\delta x}.\frac{v.\delta u - u.\delta v}{v.(v+\delta v)}, \\
&= \frac{v.\delta u/\delta x - u.\delta v/\delta x}{v.(v+\delta v)}.
\end{aligned}$$

As $\delta x \to 0$, $\delta u/\delta x = du/dx$, $\delta v/\delta x = dv/dx$, and $\delta v \to 0$. Since A represents $u(x)\times 1/v(x)$ we can replace it by our original notation $f(x)$ so we get

$$\frac{d}{dx}\big(f(x)\big) = \frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v.\frac{du}{dx} - u.\frac{dv}{dx}}{v^2}. \tag{14}$$

1.5.2 The formal proof of the quotient rule

Going back to first principles we can formally prove the quotient rule as follows. If $f(x) = u(x)/v(x)$, then

$$\frac{d}{dx}\left(\frac{u}{u}\right) = \lim_{\delta x \to 0} \frac{\frac{u(x+\delta x)}{v(x+\delta x)} - \frac{u(x)}{v(x)}}{\delta x}.$$

Our aim now is again to convert the above into some combination of $[u(x + \delta x) - u(x)]/\delta x$ and $[v(x + \delta x) - v(x)]/\delta x$ since, when we take limits, we know that the answer to these are derivatives.

So from our definition above we firstly cross multiply to get

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \lim_{\delta x \to 0} \frac{v(x)u(x+\delta x) - u(x)v(x+\delta x)}{\delta x\{v(x)v(x+\delta x)\}}.$$

We now add and subtract the term $u(x).v(x)$ to the numerator. This will allow us to recover the forms $u(x + \delta x) - u(x)$ and $v(x + \delta x) - v(x)$ which we need to form our derivatives. Hence

$$\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) = \lim_{\delta x \to 0} \frac{[v(x)u(x+\delta x) - u(x)v(x)] - [u(x)v(x+\delta x) - u(x)v(x)]}{\delta x\{v(x).v(x+\delta x)\}}.$$

Rearranging terms in the right hand side gives

$$\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) = \lim_{\delta x \to 0} \frac{v(x)[u(x+\delta x) - u(x)] - u(x)[v(x+\delta x) - v(x)]}{\delta x\{v(x).v(x+\delta x)\}}.$$

Now we have recovered the forms $u(x + \delta x) - u(x)$ and $v(x + \delta x) - v(x)$ we can separate and apply the limits:

$$\begin{aligned}\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) &= \lim_{\delta x \to 0} \frac{v(x)[u(x+\delta x) - u(x)]}{\delta x\{v(x).v(x+\delta x)\}} - \lim_{\delta x \to 0} \frac{u(x)[v(x+\delta x) - v(x)]}{\delta x\{v(x).v(x+\delta x)\}}, \\ &= v(x)\lim_{\delta x \to 0} \frac{u(x+\delta x) - u(x)}{\delta x\{v(x).v(x+\delta x)\}} - u(x)\lim_{\delta x \to 0} \frac{v(x+\delta x) - v(x)}{\delta x\{v(x).v(x+\delta x)\}}.\end{aligned}$$

Remember that since $u(x)$ and $v(x)$ are not dependent on δx they can be taking out of the limits. Rearranging further we get

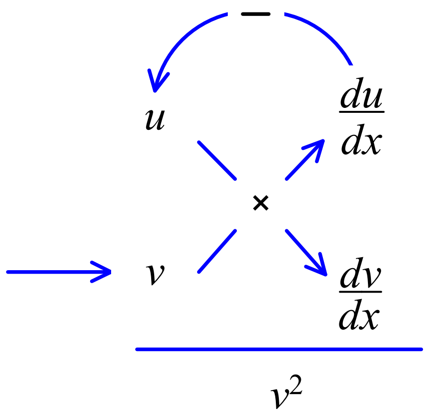
$$\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) = v(x)\lim_{\delta x\to 0}\left\{\frac{u(x+\delta x)-u(x)}{\delta x}\frac{1}{\{v(x).v(x+\delta x)\}}\right\}$$

$$-u(x)\lim_{\delta x\to 0}\left\{\frac{v(x+\delta x)-v(x)}{\delta x}\frac{1}{u(x).v(x+\delta x)}\right\}.$$

Since we now have the forms $[u(x + \delta x) - u(x)]/\delta x$ and $[v(x + \delta x) - v(x)]/\delta x$ we can now apply the limits (note that as $\delta x \to 0$, $v(x + \delta x) \to v(x)$, so the denominator becoems $v^2(x)$). Doing this give us

$$\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) = \frac{v(x)\frac{du(x)}{dx} - u(x)\frac{dv(x)}{dx}}{[v(x)]^2}. \quad (15)$$

The process of the quotient rule can be represented visually as a form of cross multiplication, subtraction, then division as seen in the diagrams below:



Examples

1) If $f(x) = (e^x + 1)/(e^x - 1)$ then

$$\frac{df}{dx} = \frac{(e^x-1).\frac{d}{dx}(e^x+1) - (e^x+1).\frac{d}{dx}(e^x-1)}{(e^x-1)^2},$$

$$= \frac{e^x(e^x-1) - e^x(e^x+1)}{(e^x-1)^2},$$

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

2) If $y = 2\sin x\,/(\sin x - \cos x)$ then

$$\frac{dy}{dx} = \frac{(\sin x - \cos x).\frac{d}{dx}(2\sin x) - 2\sin x\,.\frac{d}{dx}(\sin x - \cos x)}{(\sin x - \cos x)^2},$$

$$= \frac{2\cos x\,(\sin x - \cos x) - 2\sin x\,(\cos x + \sin x)}{(\sin x - \cos x)^2},$$

$$= \frac{-2\cos^2 x - 2\sin^2 x}{(\sin x - \cos x)^2},$$

$$= \frac{-2}{(\sin x - \cos x)^2}.$$

where we get the last step using the trig identity $\cos^2 x + \sin^2 x = 1$.

3) If $x = t^2/\ln t$ then

$$\frac{dx}{dt} = \frac{\ln t\,.\frac{d}{dx}(t^2) - t^2.\frac{d}{dx}(\ln t)}{(\ln t)^2},$$

$$= \frac{2t\ln t - t}{(\ln t)^2}.$$

4) If $y = (x^2 + 2x)/(1 - x^2)$ then

$$\frac{dy}{dx} = \frac{(1 - x^2).\frac{d}{dx}(x^2 + 2x) - (x^2 + 2x).\frac{d}{dx}(1 - x^2)}{(1 - x^2)^2},$$

$$= \frac{(1 - x^2)(2x + 2) + 2x(x^2 + 2x)}{(1 - x^2)^2},$$

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

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.

1.5.3 An alternative proof of the quotient rule

The following is adapted from *Differentiation Using the Product Rule*, Brian J. Philp, *The Mathematical Gazette*, Vol. 83, No. 497 (Jul., 1999), pp. 303-304.

Here we will go derive the quotient rule using only algebra and the product rule. Let us therefore consider again $f(x) = u/v$, where u and v are functions of x, but this time look at $f(x)$ as

$$f(x) = u\frac{1}{v}.$$

The derivative is symbolised as

$$f'(x) = \left(u\frac{1}{v}\right)',$$

and by the product rule we have

$$f'(x) = u\left(\frac{1}{v}\right)' + \frac{1}{v}u' . \qquad [*]$$

All we need to do now is find $(1/v)'$. We know this is not the same as $(1)'/(v)'$, so we do the following: let $h = 1/v$. Then $h.v = 1$. Using the product rule on this get

$$h'v + hv' = 0 .$$

Adding $-2hv'$ to both sides we get

$$h'v - hv' = -2hv' .$$

Since $h = 1/v$ we get

$$h'v - hv' = -\frac{2v'}{v}.$$

Dividing by v gives

$$h' - \frac{h}{v}v' = -\frac{2v'}{v^2},$$

and again since $h = 1/v$ we get

$$h' - \frac{1}{v^2}v' = -\frac{2v'}{v^2},$$

from which we get

$$h' = -\frac{v'}{v^2}.$$

So we have been able to show that

$$\text{if} \quad h = \frac{1}{v} \quad \text{then} \quad h' = -\frac{v'}{v^2}.$$

We now use this in [*] above. Hence

$$
\begin{aligned}
f'(x) &= u\left(\frac{1}{v}\right)' + \frac{1}{v}u' , \\
&= u\frac{-v'}{v^2} + \frac{1}{v}u' , \\
&= \frac{u'v^2 - v'uv}{v^3} , \\
&= \frac{u'v - v'u}{v^2} .
\end{aligned}
$$

1.5.4 An extension of the quotient rule

This section is adapted from *The Formula* $(u/v)' = (vu' - uv')/v^2$, L. H. LeBon, *The Mathematical Gazette*, Vol. 49, No. 369 (Oct., 1965), pp. 296-297.

Here we will look at a slightly modified version of the quotient rule which may, in some instances, be more convenient and may reduce the likelyhood of algebraic errors during the process of differentiation. The purpose of this quotient rule is to be able to handle situations involving v^n.

So, given $f(x) = u/v^n$, where *u* and *v* are functions of *x*, we have

$$
\begin{aligned}
\frac{d}{dx}\left(\frac{u}{v^n}\right) &= \frac{v^n u' - n.uv^{n-1}}{v^{2n}} , \\
&= \frac{vu' - n.uv'}{v^{n+1}} .
\end{aligned}
\qquad (16)
$$

To see the usefulness of this form consider differentiating $f(x) = (1 + 2x^2)/(1 + x)^3$. Using (16) we have $u = 1 + 2x^2$ and $v = 1 + x$, with $n = 3$, hence

$$
\begin{aligned}
\frac{d}{dx}\left(\frac{1 + 2x^2}{(1 + x)^3}\right) &= \frac{(1 + x)(4x) - 3(1 + 2x^2)(1)}{(1 + x)^4} , \\
&= \frac{-2x^2 + 4x - 3}{(1 + x)^4} .
\end{aligned}
$$

Differentiating by the standard quotient rule would give the same result, but would involve much more algebraic manipulation (try it).

1.5.5 Proof of the power rule for negative exponents using the quotient rule

We can use the quotient rule to prove the power rule when the exponent is negative. So, if $y = x^{-n}$, where n is a positive integer, we want to show that $\frac{dy}{dx} = -n.x^{-n-1}$. We do this as follows:

$$y = x^{-n},$$

$$y = \frac{1}{x^n},$$

$$\frac{dy}{dx} = \frac{0 - n.x^{n-1}}{(x^n)^2},$$

$$= -n.(x^{n-1}).(x^{-2n}),$$

$$= -n.x^{-n-1},$$

1.5.6 The quotient rule for more that two functions

In section 1.5.1 we saw how to differentiate two functions when one of them is divided by the other. How would we differentiate the product of three functions? How would we differentiate $f(x) = e^x/x/\sin x$?

What we effectively have here is a situation where *f*(*x*) = $u(x)/v(x)/w(x)$. In this case things are implified by the fact that *f*(*x*) = $u(x)/[v(x).w(x)]$. We can then apply the quotient rule for *f*(*x*) as a whole, and apply the product rule to the denominator.

Hence we can do

$$\frac{d}{dx}\left(\frac{\frac{e^x}{x}}{\sin x}\right) = \frac{d}{dx}\left(\frac{e^x}{x\sin x}\right) = \frac{\frac{d}{dx}(e^x).(x\sin x) - e^x\frac{d}{dx}(x\sin x)}{x^2\sin^2 x},$$

$$= \frac{e^x x\sin x - e^x\sin x - e^x x\cos x}{x^2\sin^2 x}.$$

In general we can therefore say that if $f(x) = u/v/w = u/(v.w)$ then

$$\frac{d}{dx}\left(\frac{u}{v.w}\right) = \frac{(v.w)\frac{du}{dx} - u\frac{d}{dx}(v.w)}{(v.w)^2},$$

$$= \frac{(v.w)\frac{du}{dx} - u(v.\frac{dw}{dx} + w.\frac{dv}{dx})}{(v.w)^2},$$

or, in simpler notation,

$$\left(\frac{u}{v.w}\right)' = \frac{u'.v.w - u.v'.w - u.v.w'}{(v.w)^2}. \qquad (17)$$

Notice the pattern in the numerator of the result: the first function to be differentiated is the numerator u. Then the subtraction occurs on the combination of the derivatives of the denominator functions v and w.

Example

If $f(x) = x^2 \ln x \tan x$ then by (17) we have

$$f'(x) = \frac{(x^2)'.\ln x \tan x - x^2(\ln x)' \tan x - x^2 \ln x\,(\tan x)'}{(\ln x \tan x)^2}$$

which simplifes to

$$f'(x) = \frac{2x.\ln x \tan x - x \tan x - x^2 \ln x \sec^2 x}{(\ln x \tan x)^2}.$$

Similarly the following extensions to the quotient rule for multiple functions in the numerator and denominator can be shown:

If $\quad f(x) = \frac{uv}{wz} \quad$ then $\quad \left(\frac{uv}{wz}\right)' = \frac{u'.v.w.z + uv'w.z - u.v.w'.z - u.v.w.z'}{(w.z)^2}.$

If $\quad f(x) = \frac{uv}{wyz} \quad$ then $\quad \left(\frac{uv}{wyz}\right)' = \frac{u'vwyz + uv'wyz - uvw'yz - uvwy'z - uvwyz'}{(w.y.z)^2},$

etc. Again notice the pattern in the plus and minus sings in the numerator: u and v are functions in the numerator, and their derivative combinations are added. Functions w, y and z are in the denominator and their derivative combinations are subtracted.

1.5.7 When does d(u/v)/dx = (du/dx)/(dv/dx)?

In section 1.4.8 we saw the case of when it was possible for the derivative of a product to be equal to the product of the derivatives. We can ask the same question about the derivative of the division of two functions: when will

$$\frac{d}{dx}\left(\frac{u}{u}\right) = \frac{du/dx}{dv/dx} \ ?$$

We know this is not true in general since

$$\frac{d}{dx}\left(\frac{\sin x}{\cos x}\right) = \sec^2 x$$

and

$$\frac{d(\sin x)/dx}{d(\cos x)/dx} = -\cot x \ .$$

And as for section 1.4.8 we will need to use integration.

So, we know that

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v}$$

so the right hand side of the two equations above would have to be equal. Therefore, if we do this (and simplify the notation to make it easier to read) we get

$$\frac{u'}{v'} = \frac{v.u' - u.v'}{v^2}.$$

Our aim now is to use algebra to transform this expression into something we know how to integrate. In this case we can write the above as

$$u'(v.v' - v^2) = u(v')^2.$$

Doing appropriate algebra in order to get an expression we can integrate we end up with

$$\frac{u'}{u} = \frac{(v')^2}{v.v' - v^2} \ .$$

Let us now integrate:

$$\int \frac{u'}{u} \ dx = \int \frac{(v')^2}{v.v' - v^2} \ dx \ .$$

As in section 1.4.8 the left hand side can be integrated directly as a log function:

$$\ln u + k = \int \frac{(v')^2}{v.v' - v^2} \ dx \ .$$

where k is the constant of integration. Since k is a number it is going to be easier in the following steps if we rewrite it as ln c. So now we have

$$\begin{aligned}\ln u + \ln c &= \int \frac{(v')^2}{v.v' - v^2}\,dx\,,\\ \ln(c.v) &= \int \frac{(v')^2}{v.v' - v^2}\,dx\,,\\ v &= c.\exp\int \frac{(v')^2}{v.v' - v^2}\,dx\,. \qquad [*]\end{aligned}$$

Expression [*] is what we are looking for. What it means is that the only way $(u/v)' = u'/v'$ is if $f(x)$ satisfies equation [*].

To see what this means in practice consider differentiating

$$f(x) = \frac{x}{x(x-1)}\,.$$

Without cancelling the terms in x we have $u(x) = x/(x-1)$ and $v(x) = x$. Then

$$\begin{aligned}f'(x) = \left(\frac{u}{v}\right)' = \frac{u'}{v'} &= \frac{(1-x)^{-1} + x(1-x)^{-2}}{1}\,,\\ &= \frac{1}{1-x} + -\frac{x}{(x-1)^2}\,,\\ &= \frac{(1-x)[(1-x)+x]}{(x-1)^3}\,,\\ &= \frac{1}{(x-1)^2}\,,\end{aligned}$$

and

$$\begin{aligned}f'(x) = \left(\frac{u}{v}\right)' = \frac{vu' - uv'}{v^2} &= \frac{x.\left(\frac{1}{(1-x)^2}\right) + \frac{x}{1-x}.(1)}{x^2}\,,\\ &= \frac{\frac{x}{(1-x)^2} + \frac{x}{1-x}}{x^2}\,,\\ &= \frac{x(1-x) - x(1-x)^2}{x^2(x-1)^3}\,,\\ &= \frac{x(x-1)(1-(1-x))}{(x-1)^3}\,,\\ &= \frac{1}{(x-1)^2}\,.\end{aligned}$$

So $f(x) = u/v$ satisfies the incorrect quotient rule.

The way to create the correct $u(x)$ which will work for the incorrect quotient rule is the same as that for the incorrect product rule: simply to choose any function as your $v(x)$ and perform the integral [*]. The only issue will be in how easy it will be to perform [*].

For example if we choose $v(x) = \sin x$ then

$$\int \frac{(v')^2}{(v.v' - v^2)}\,dx = \int \frac{\cos^2 x}{\sin x \cos x - \sin^2 x}\,dx$$

which may of may or may not be integratable.

As with the incorrect product rule there is one problem with equation [*]. If we choose $v(x) = e^x$ then the denominator of [*] is zero. So [*] does not work for all functions (for example if $v(x) = e^x$).

1.6 On the chain rule for differentiation: $d[f(g(x))]/dx$

Suppose we have a function f containing another function $g(x)$. In other words we have $f(g(x))$. What will be the result of differentiating $f(g(x))$? Will it be

$$\frac{d}{dx}\left(f\left(\frac{dg}{dx}\right)\right)\ ?$$

In other words will we need to substitute the derivative dg/dx into f, after which we can differentiate f?

To answer this consider $f(g(x)) = g^2$ where $g(x) = 2x$. This gives $f(g(x)) = (2x)^2 = 4x^2$. We know the derivative of this is $d[f(g(x))]/dx = 8x$.

Let us now see what happens when we input into f our derivative of g, and then differentiate f : $dg/dx = 2$, which means we have $f(2)$. But $f(2)$ is just the evaluation of f at $x = 2$, i.e. $f(2)$ is a constant, so $df/dg = 0$ which implies that $df/dx = 0$ which we know is not the correct result.

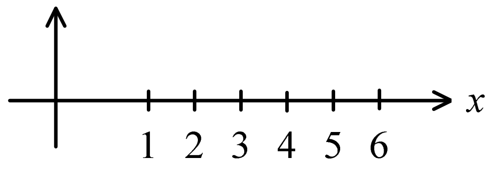
However, if we instead do

$$\frac{d}{dx}\big(f(g(x))\big) = \frac{df}{dg} \times \frac{dg}{dx} = 2g.(2) = 2(2x)(2) = 8x\,,$$

we get the correct result. This is the chain rule for differentiation.

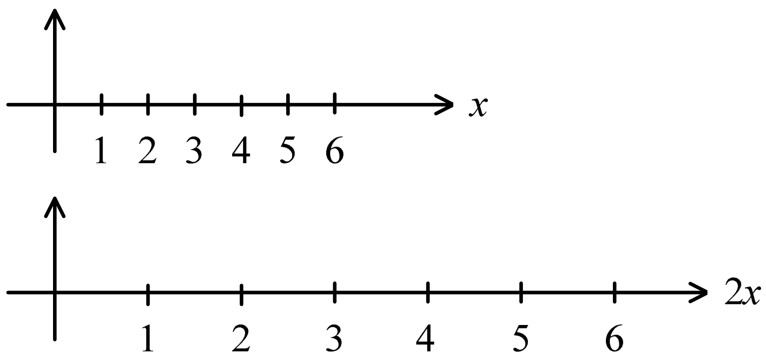
1.6.1 A conceptual description of the chain rule

To understand why the chain does what it does let us first consider the distribution of numbers along the x-axis. To make things easy let us just consider the integer values of x along the x-axis. The distribution of these x values gives rise to the standard x-axis. Any function $f(x)$ is then plotted against this standard, evenly distributed, set of x values which (for the purpose of this example) are all one unit apart:



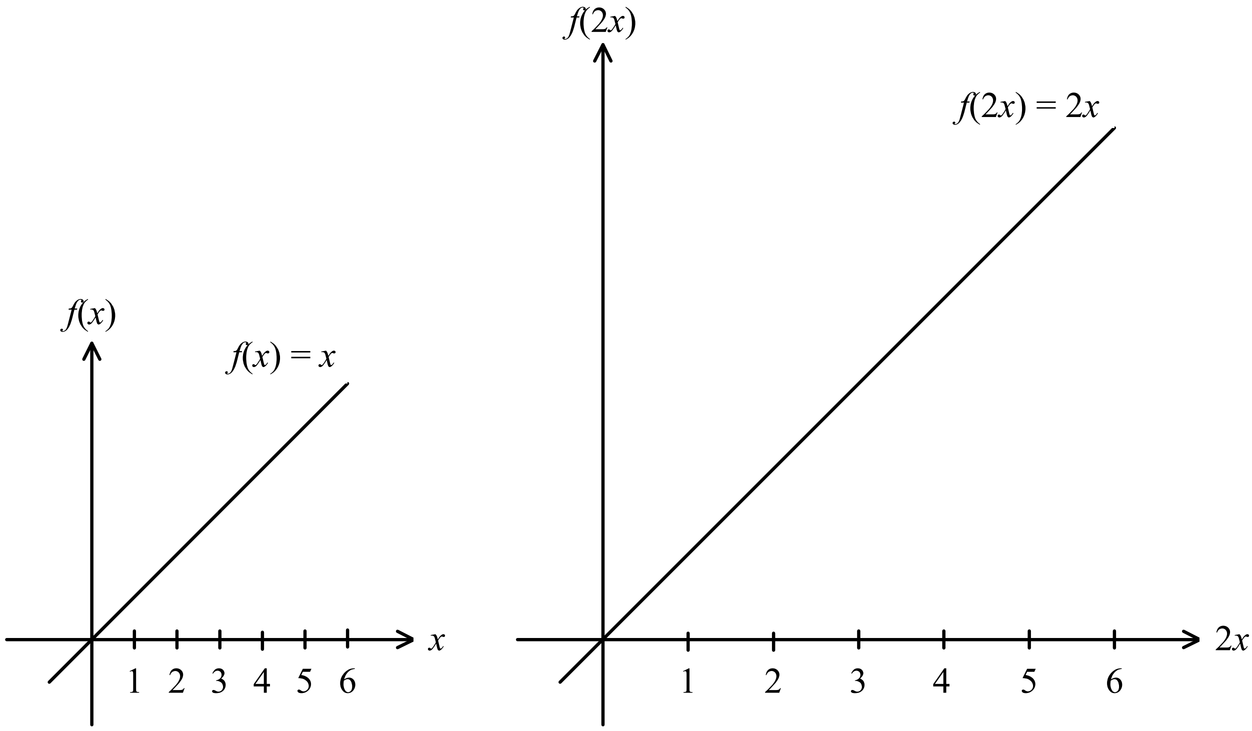
So $f(x)$ simply represents how x changes according to the function specified by $f(x)$, and the derivative of $f(x)$ represents the instantaneous rate of change of $f(x)$ compared to a change in x.

Let say we now change the scale of the x-axis to be $2x$. Every number on the x-axis is now doubled, and the new "$2x$" axis can then be said to be a stretched or expanded version of the standard x-axis. This gives us a new horizontal axis $g(x) = 2x$ where the x values along $g(x)$ are still evenly distributed, but are now two units apart:



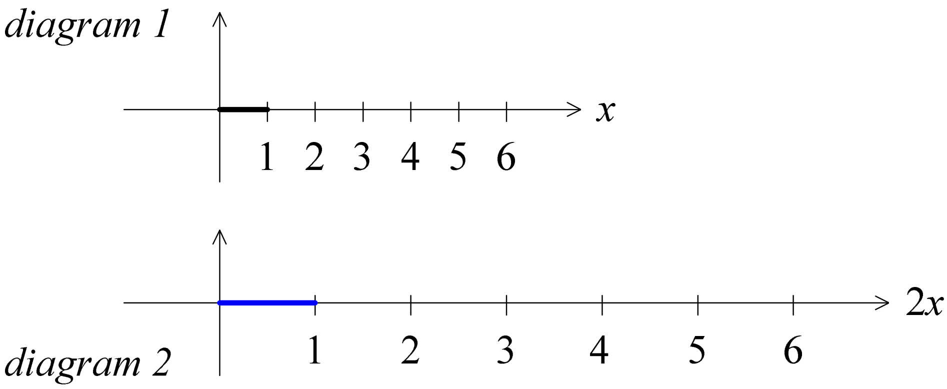
Effectively we are recalibrating the x-axis to be a $g(x)$ axis, and the y-axis to be a $f(g)$ axis. Any function plotted against this new $g(x)$ axis will be a function of $2x$ not simply a function of x. In general any function plotted against this new $g(x)$ axis will be an $f(g(x))$ function. The derivative of $f(g(x))$ will then represent the instantaneous rate of change of $f(g(x))$ compared to a change in $g(x)$.

Just as plotting $f(x) = x$ gives us a staight line of slope 45^0 so plotting the function $f(g(x)) = 2x$ against $g(x) = 2x$ will also gives us a straight line with slope 45^0 since we are plotting the function against an axis calibrated according to its own scale:



The derivative of $f(g((x))$ will therefore be w.r.t. $g(x)$, i.e. $d(f(g))/dg$. But we don't want to plot $f(g(x))$ against the transformed $g(x)$ axis, and we don't want the derivative with respect to $g(x)$. We want to plot $f(g(x))$ against the (untransformed) standard x-axis so that we can get the derivative of $f(g(x))$ w.r.t to x. And the way we can do this is to contract the $g(x)$ axis so that its units line up with, or match, those of the standard x-axis.

To explain the effect of this contraction consider the diagrams below. In diagram 1 the length of one unit of distance along the x-axis is 1 (shown in black). In diagram 2 the length of one unit of distance along the $g(x)$ axis is also 1 (shown in blue). However, compared to the x-axis the length of one unit of distance is 2.

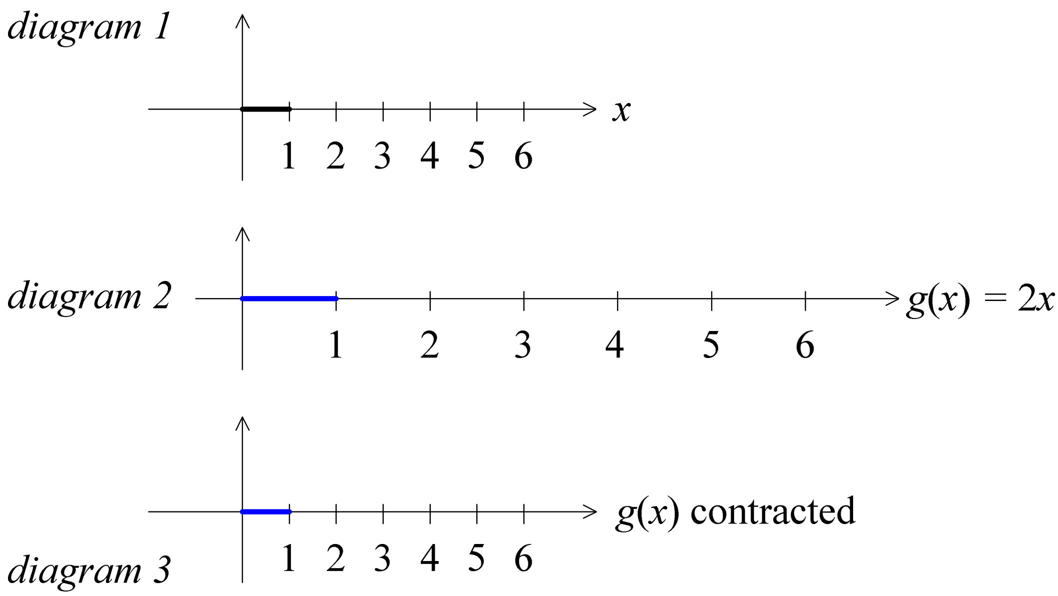


As a comparison of the length of each unit distance of the x-axis to each respective unit length of the $g(x)$ axis we have

Interval	Length of unit distance	
	x-axis	**$g(x)$ axis compared to x-axis**
...	...	...
[0, 1]	1	2
[1, 2]	1	2
[2, 3]	1	2
[3. 4]	1	2
...	...	...

Contracting the $g(x)$ axis by a factor of ½ will then rescale the g(x) axis to match the scale of the x-axis and allow the unit distances of the $g(x)$ axis to match those of the x-axis (we could expand the x-axis to make it the same scale as the $g(x)$ axis but we want to do the former because we are looking for the rate of change of f with respect to x. If we did the latter we would be finding the rate of change of f w.r.t. g which we know straight away if simply df/dg).

This contraction is shown in diagram 3:

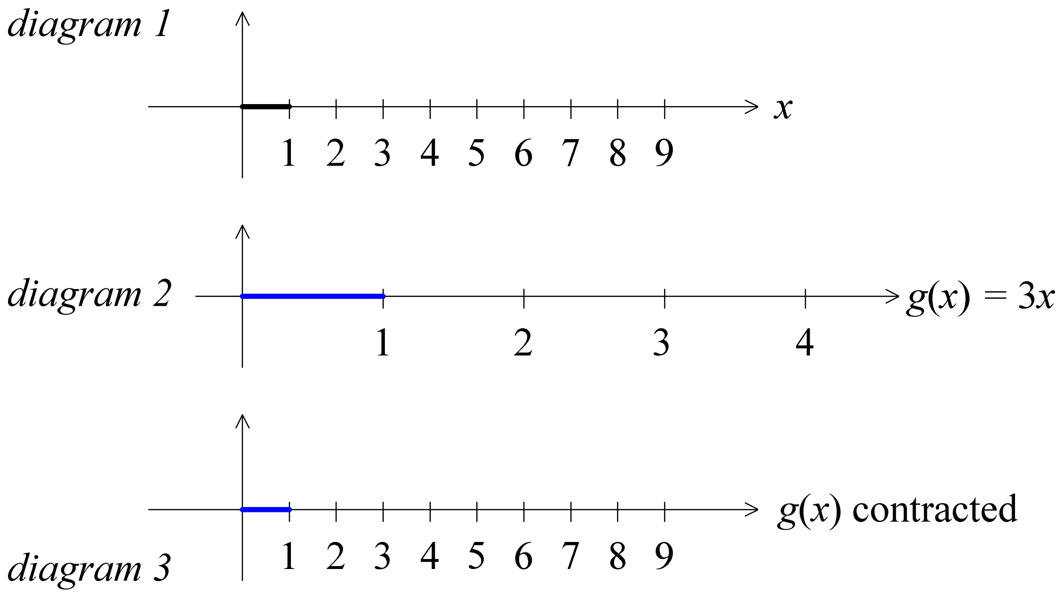


The contraction of the $g(x)$ axis now causes it to run twice as fast as it used to when it was uncontracted (or we can say that the $g(x)$ axis covers 1 unit of distance in half the time it used to).

In contracting the $g(x)$ axis, $f(2x)$ will also be contrated. Specifically, the horizontal movement of $f(2x)$ has also been contracted whilst keeping the height of $f(2x)$ the same. This means that the function $f(2x)$ is also running twice as fast. The function $f(2x)$ will reach its value/height in half the time or distance along the x-axis (i.e. twice as quickly) than $f(x)$. For example, if the value of $f(x)$ is 6 at $x = 2$, then $f(2x)$ will reach a value of 6 at $x = 1$.

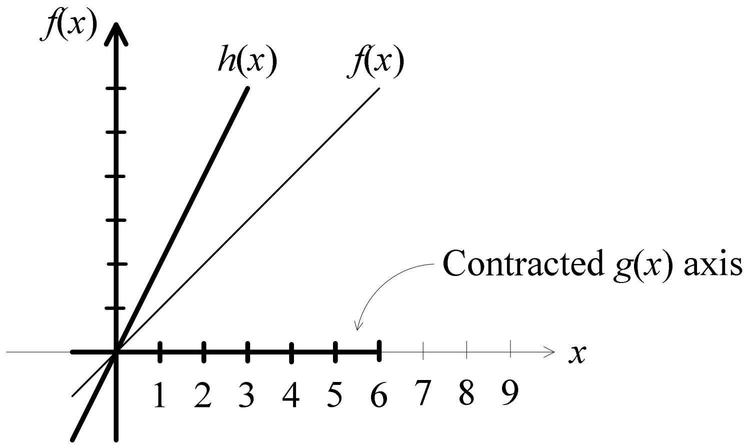
This has the effect of changing the slope of $f(2x)$ for a given height compared to the slope of $f(x)$. And in the example of $g(x) = 2x$ the slope of $f(2x)$ has changed by a factor of 2 compared to $f(x)$. This is because the $g(x)$ axis has changed at a rate of 2, i.e. $dg/dx = 2$. But this is simply the derivative of $g(x) = 2x$.

If instead we had $g(x) = 3x$, we would have the following sequence of axes diagrams:



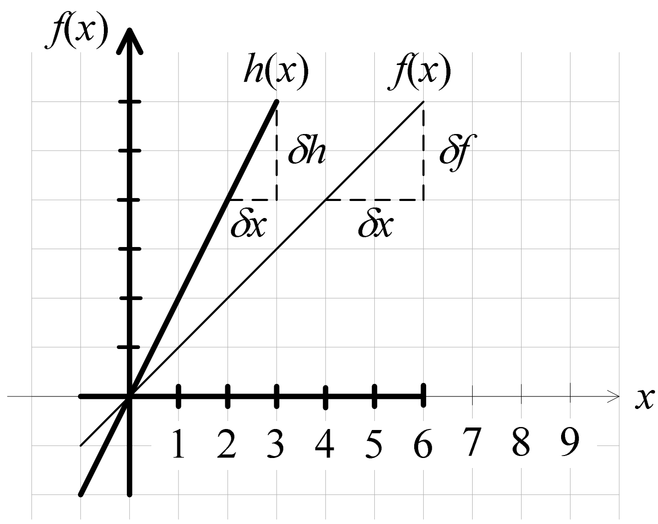
and the $g(x)$ axis is now running three times as fast as it used to compared to its original rate implying that function $f(3x)$ is also running three times as fast, reaching a given height in one-third of the distance/time compared to $f(x)$, implying that the slope of $f(3x)$ will be affected, in this case by a factor of 3 compared to $f(x)$.

Returning to the case where $g(x) = 2x$, the graph of $f(2x) = 2x$ will now also be contracted when plotted against the x-axis:



Now, $f(2x)$ had it particular slope when originally plotted against its won $2x$ axis. But in contracting the $g(x)$ axis we have contracted the horizontal part of $f(2x)$, whilst the vertical part of $f(2x)$ has remained unchanged. This contraction of the $g(x)$ axis has then had the effect of increasing the gradient of $f(2x)$ by a factor of 2. This is because any ratio of δy over δx on $f(2x)$ will be twice that of $f(x)$.

Geometrically speaking this means that, for the gradient of a function $h(x) = f(2x)$, the vertical distances δf and δh remain the same but the horizontal distance δx of $h(x)$ will be halved when compared to $f(x)$. This therefore increases the gradient of h(x) by a factor of 2:



$$\frac{\delta h}{\delta x} = \frac{3}{1/2} \qquad \frac{\delta f}{\delta x} = \frac{3}{1}$$

In contracting $g(x)$ by the amount we have, we have halved the δx values of $h(x)$ compared to those of $f(x)$, whilst keeping the vertical distances δh and δf unchanged. So the slope of $h(x)$ will be increased by a factor of 2. This factor of 2 is simply the derivative of $g(x) = 2x$.

Therefore we can say

$$\frac{dh}{dx} = \frac{d}{dx}f(2x) = \frac{d}{dx}(2x)\,\frac{d}{dx}f(x) = 2\frac{df}{dx}.$$

In other words,

$$\frac{dh}{dx} = \frac{d}{dx}f\big(g(x)\big) = \frac{dg}{dx}\,\frac{df}{dx}.$$

This can be interpreted as dh/dg = (rate of change of $f(g(x))$ w.r.t. transformed axis $g(x)$) * (rate of change of transformed axis $g(x)$ w.r.t. standard, untransformed, axis x).

1.6.2 More on the effect of g(x)

In terms of the contraction (or expansion) effect of $g(x)$, and the effect of this on the slope of $f(g(x))$, the same argument applies whatever the function $g(x)$. If $g(x) = x^2$ we have

diagram 1

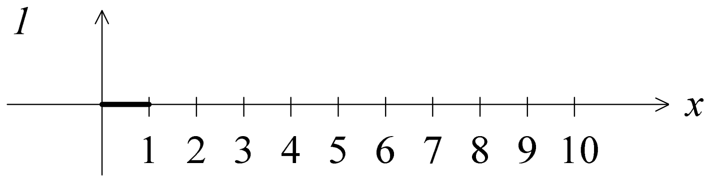
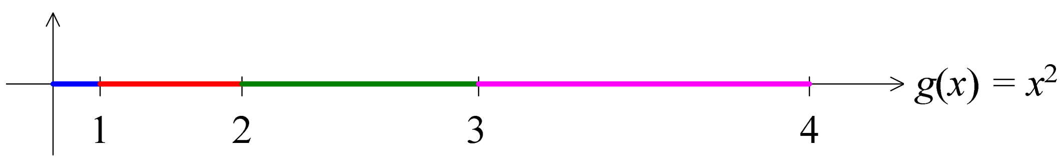


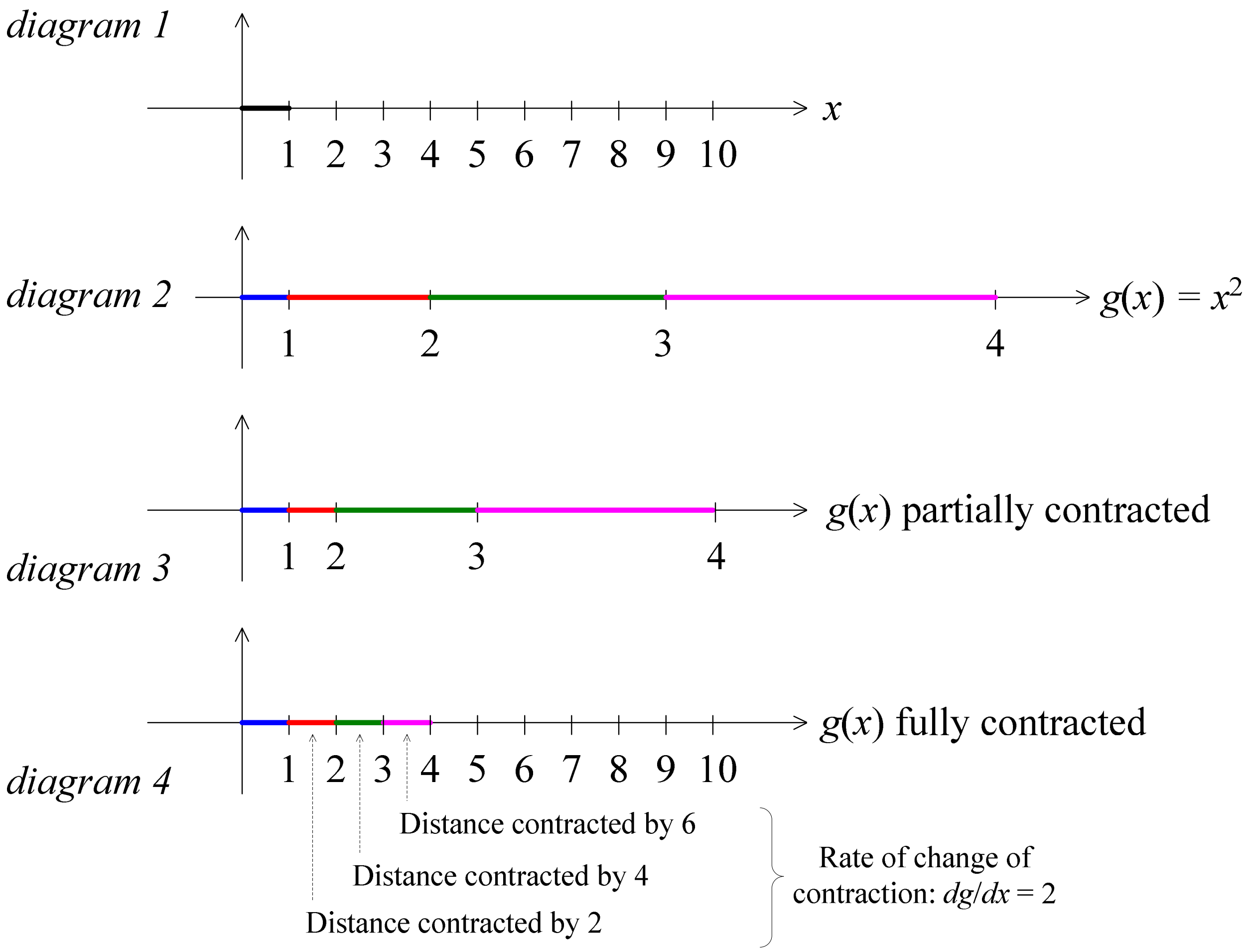
diagram 2



A comparison of the length of each unit distance of the x-axis to each respective unit distance of the $g(x)$ axis can be seen in the table on the next page. From this we see that in order to match up the scale of the $g(x)$ axis with that of the x-axis we have to contract the $g(x)$ axis. If we contract it by a distance of 2 we match up the scale of the x-axis and $g(x)$ axis only for [0, 2] but all other intervals between the axes do not match up (as seen in diagram 3 below).

Interval	**Length of unit distance**	
	***x*-axis**	***g*(*x*) axis compared to *x*-axis**
...	...	...
[0, 1]	1	1
[1, 2]	1	3
[2, 3]	1	5
[3, 4]	1	7
...	...	...

A comparison of the length of each unit distance of the x-axis to each respective unit length of the g(x) axis

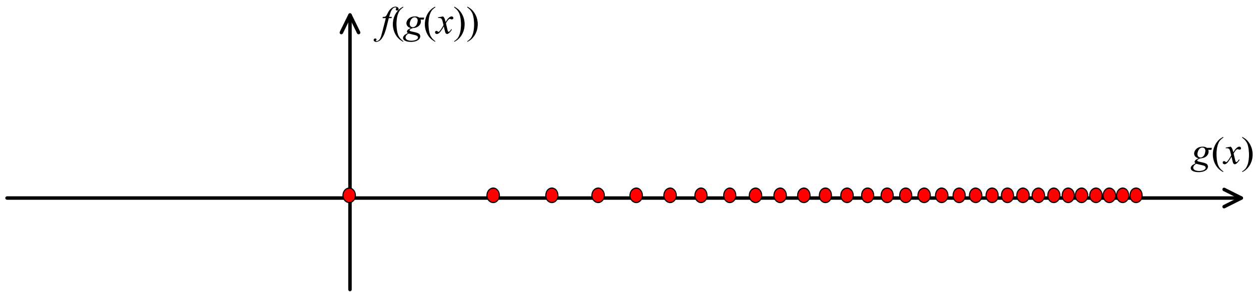


So in order to match up the whole of the scale of the $g(x)$ axis with the whole of the x-axis scale we have to contract each interval by an amount which increases as we move through each interval.

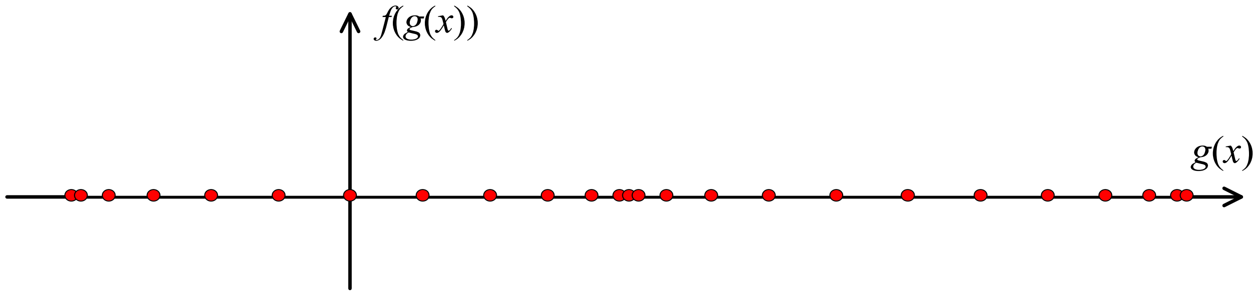
In fact if we reduce the interval [1, 2] on $g(x)$ by 2 we then need to reduce the interval [2, 3] by 4, the interval [3, 4] by 6, etc., as seen in diagram 4. This is because the $g(x)$ axis is changing at a non-constant rate. In fact, the rate of change of the $g(x)$ axis is $2x$, and this just happens to be the derivative of x^2 : for $g(x) = x^2$, $dg/dx = 2x$. This dg/dx is the factor by which the slope of $f(g)$ changes when measured against the x-axis.

In general, function $g(x)$ can then be said to have the effect of contracting (or expanding) the standard x-axis prior to being used as input for function f. Different functions $g(x)$ will therefore contract (or expand) the standard x-axis differently and at different rates.

For example, if $g(x) = \sqrt{x}$, all x values will be distributed along the horizontal axis according to a square root pattern:



If $g(x) = \sin(x)$, all x values are now distributed along the horizontal axis according to a $\sin(x)$ pattern:



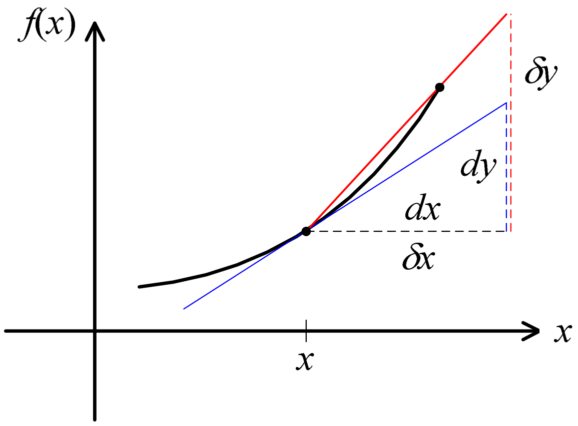
In general this means that there will be a rate of change associate in going from the standard input values, x, to the new input values, $g(x)$. This rate of change will then affect the rate at which f changes.

In summary, the degree of adjustement in the slope of $f(g)$ as measured against the g axis to the slope of $f(g)$ as measured against the x axis is given by the factor dg/dx.

1.6.3 A formal proof of the chain rule

Here we will go through the proof of the chain rule. I will present the same proof in two different forms. One form will be easier on the eye and on one's mathematical thinking load, but will not explicitly show the complete interrelationship between all variables involved. The other form will explicitly show this interrelationship but will, as a consequence, require more mathematical thinking effort.

Before we start we need to consider the difference between the slope of the secant passing through point x and a second point on a function, and the slope of the tangent at x on that function:



The slope of the secant is $\delta y/\delta x$ (the red dashed line divided by the black dashed line). But this is an approximation to the exact slope of the tangent which is dy/dx (the blue dashed line divided by the black dashed line). So we can say

$$\frac{\delta y}{\delta x} \approx \frac{dy}{dx}.$$

So there is an error ε between these two slopes, and we can make the above expression an equality if we include this error:

$$\frac{\delta y}{\delta x} = \frac{dy}{dx} + \epsilon\,. \qquad (*)$$

We will need to bear this in mind through the proof below.

<u>*Proof: Form 1*</u>

Given a function $h(x) = f(g(x))$, we want to find dh/dx. Based on (*) above we have, for $f(g(x))$,

$$\frac{\delta f}{\delta g} = \frac{df}{dg} + \epsilon\,.$$

Cross multiplying we get

$$\delta f = \delta g \frac{df}{dg} + \epsilon . \delta g \; .$$

We can now divide by δx so that we are looking at a rate of change w.r.t. x :

$$\frac{\delta f}{\delta x} = \frac{\delta g}{\delta x} . \frac{df}{dg} + \epsilon . \frac{\delta g}{\delta x} \, .$$

Taking limits we get

$$\begin{aligned}
\lim_{\delta x \to 0} \frac{\delta f}{\delta x} &= \lim_{\delta x \to 0} \left(\frac{\delta g}{\delta x} . \frac{df}{dg} + \epsilon . \frac{\delta g}{\delta x} \right), \\
&= \lim_{\delta x \to 0} \left(\frac{\delta g}{\delta x} . \frac{df}{dg} \right) + \lim_{\delta x \to 0} \left(\epsilon . \frac{\delta g}{\delta x} \right), \\
&= \frac{df}{dg} \lim_{\delta x \to 0} \left(\frac{\delta g}{\delta x} \right) + \lim_{\delta x \to 0} (\epsilon) \lim_{\delta x \to 0} \left(\frac{\delta g}{\delta x} \right) .
\end{aligned}$$

We can take out df/dg from the first limit since it is independent of δx. The first limit then simply becomes dg/dx (remember that we only use limits so that we can do $\delta x \to 0$. If there is no δx in our term then it is not affected by the limit, and can therefore be factored out of the limit).

In the second limit we have that, as $\delta x \to 0$, $\delta g/\delta x \to dg/dx$ so $\epsilon \to 0$. Therefore $\lim_{\delta x \to 0} \delta \epsilon = 0$. Hence we have $0 \times dg/dx = 0$. So we end up with the chain rule to be

$$\frac{df}{dx} = \frac{df}{dg} \frac{dg}{dx} .$$

Proof: Form 2

Given $f(g(x))$ we have from 1st principles

$$\begin{aligned}
\frac{d}{dx}\left[f(g(x)) \right] &= \lim_{\delta x \to 0} \frac{f(g(x + \delta x)) - f(g(x))}{\delta x} , \\
&= \lim_{\delta x \to 0} \frac{f(g + \delta g) - f(g)}{\delta x} . \qquad (**)
\end{aligned}$$

Expression (**) looks like the form $\left(f(x + \delta x) - f(x)\right)/\delta x$ except that the numerator is evaluated at $g(x)$ instead of x.

If the denominator of (**) was δg then we would have the standard expression for the ratio of vertical height and horizontal distance, which would then lead to the derivative df/dg.

However, we can transform (**) so that we do get a denominator δg by multiplying the ratio in (**) by $\delta g/\delta g$:

$$\frac{df}{dx} = \lim_{\delta x \to 0} \left(\frac{f(g+\delta g) - f(g)}{\delta x} \times \frac{\delta g}{\delta g} \right) .$$

Now we swap the denominators:

$$\frac{df}{dx} = \lim_{\delta x \to 0} \left(\frac{f(g+\delta g) - f(g)}{\delta g} \times \frac{\delta g}{\delta x} \right) .$$

We can do this since everything inside the bracket represents numbers divided by other numbers. Remember that before one applies the limit, the fraction still represents a division of numbers. It is only after applying the limits that the "object" (i.e. the derivative) is no longer a division of numbers.

We now have

$$\frac{df}{dx} = \lim_{\delta x \to 0} \left(\frac{f(g+\delta g) - f(g)}{\delta g} . \frac{g(x+\delta x) - g(x)}{\delta x} \right) , \qquad (***)$$

provided $\delta g \neq 0$.

Applying the limits then gives

$$\frac{df}{dx} = \frac{df}{dg}\frac{dg}{dx} .$$

But we now have a problem: δg cannot be 0 in (***). Remember that $\delta g = g(x+\delta x) - g(x)$, so if $g(x+\delta x) = g(x)$ then we have δg = 0. This is the case for functions which have points of inflection at some point x in the domain of $g(x)$. In this situation we have $g(x+\delta x) \to g(x)$ as $\delta x \to 0$ implying that δg = 0. The chain rule still works at points of inflections but (***) says that we cannot differentiate at these points since we would be dividing by 0. So we need to find a work around to make (***) work even for cases where there are points of inflection.

So, going back to expression (*) which refers to the difference between the slope of the secant and the slope of the derivative we have, in this case,

$$\epsilon = \frac{f(g+\delta g) - f(g)}{\delta g} - \frac{df}{dg} ,$$

so that by cross multiplying and rearranging we get

$$f(g+\delta g)-f(g)=\epsilon.\delta g+\frac{df}{dg}.\delta g\,.$$

We now substitute the right hand side of this expression into (***) above to get

$$\frac{df}{dx}=\lim_{\delta x\to 0}\left(\frac{\delta g\left(\epsilon+\frac{df}{dg}\right)}{\delta g}.\frac{g(x+\delta x)-g(x)}{\delta x}\right),$$

From which the δg cancels and we no longer have a problem with division by 0. Therefore we end up with

$$\begin{aligned}\frac{df}{dx} &= \lim_{\delta x\to 0}\left(\left(\epsilon+\frac{df}{dg}\right).\frac{g(x+\delta x)-g(x)}{\delta x}\right),\\ &= \lim_{\delta x\to 0}\left(\left(\epsilon+\frac{df}{dg}\right)\right).\lim_{\delta x\to 0}\left(\frac{g(x+\delta x)-g(x)}{\delta x}\right),\\ &= \left(\lim_{\delta x\to 0}(\epsilon)+\lim_{\delta x\to 0}\left(\frac{df}{dg}\right)\right).\lim_{\delta x\to 0}\left(\frac{g(x+\delta x)-g(x)}{\delta x}\right)\end{aligned}$$

For the first limit we have that, as $\delta x\to 0$, $\delta f/\delta g\to df/dg$ so $\epsilon\to 0$. Therefore $\lim_{\delta x\to 0}\epsilon=0$. The term df/dg in the second limit is independent of δx so the limit is simply df/dg. The third limit is simply the definition of the derivative of $g(x)$.

So for a function $f(g(x))$ we end up with the usual chain rule

$$\frac{df}{dx}=\frac{df}{dg}\frac{dg}{dx}\,. \tag{18}$$

This can be read as

"derivative of $f(g(x))$ = (derivative of outside function f) × (derivative of inside function g)"

Remember that d/dx is an operation, the operation being that of differentiation. In terms of "operation" notation (18) is

$$\frac{d}{dx}=\frac{d}{dg}\frac{dg}{dx}\,.$$

It is an operation we carry out in order to convert the rate of change with respect to one variable (g) into the rate of change with respect to another variable (x), with dg/dx being the necessary conversion factor.

The process of the chain rule can be represented visually as any one of the diagrams below:

diagram 1 | *diagram 2*

(I adapted diagram 2 from
http://www.ballooncalculus.org/examples/reference.html#tanDiff)

Examples
Here we will see how to apply the chain rule in practice.

1) Problem: Find the first, second and third derivative of $y = e^{3x}$.
Solution: Let u = 3x. Then $y = e^{u}$. By the chain rule we have

$$\frac{dy}{dx} = \frac{dy}{du}.\frac{du}{dx},$$

$$= e^{u}.(3) = 3e^{3x}.$$

For the second derivative we simply repeat. We will always have u = 3x, and this will always gives us a multiply of 3 on the "outside". So thinking in the way of "inside/outside" differentiation we have

$$\frac{d^2y}{dx^2} = \frac{d}{dx}(3e^{3x}),$$

$$= 3e^{3x}(3) = 9e^{3x}.$$

and

$$\frac{d^3y}{dx^3} = \frac{d}{dx}(9e^{3x}),$$

$$= 9e^{3x}(3) = 27e^{3x}.$$

2) <u>Problem</u>: Find the first derivative of $f(x) = \ln \sin x$.

<u>Solution</u>: Let $u = \sin(x)$. Then $f(x) = \ln u$. By the chain rule we have

$$\frac{df}{dx} = \frac{df}{du}.\frac{du}{dx},$$

$$= \frac{1}{u}.\cos x = \cot x.$$

3) <u>Problem</u>: If $y = e^{-2x}(\sin 2x + \cos 2x)$ show that $y'' + 4y' + 8y = 0$.

<u>Solution</u>: Here we are going to have to differentiate twice and then subsitute our answers into the equation above to see if it equals zero.

Notice that the major operation is a product, so we we will have to do the product rule. During this process we will then have to do the chain rule three times, once on each of the three functions. To speed things up I will do this by thinking in the way of the "inside-outside" concept of the chain rule. So the "inside" part is $u = 2x$ and the "outside" parts are e^u, $\cos u$, and $\sin u$.

Hence

$$y' = \sin(2x)\frac{d}{dx}(e^{-2x}) + e^{-2x}.\frac{d}{dx}(\sin(2x))$$

$$+\cos(2x)\frac{d}{dx}(e^{-2x}) + e^{-2x}.\frac{d}{dx}(\cos(2x)),$$

$$= -2e^{-2x}\sin 2x + 2e^{-2x}\cos 2x - 2e^{-2x}\cos 2x - 2e^{-2x}\sin 2x,$$

$$= -4e^{-2x}\sin 2x.$$

Differentiating again we have

$$y'' = 8e^{-2x}\sin 2x - 8e^{-2x}\cos 2x\,.$$

Substituting these two derivatives into $y'' + 4y' + 8y$ we get

$$8e^{-2x}\sin 2x - 8e^{-2x}\cos 2x + 4(-4e^{-2x}\sin 2x) + 8\big(e^{-2x}(\sin 2x + \cos 2x)\big)$$

$$= 8e^{-2x}\sin 2x - 16e^{-2x}\sin 2x + 8e^{-2x}\sin 2x - 8e^{-2x}\cos 2x + 8e^{-2x}\cos 2x = 0.$$

4) Problem: Find the first derivative of $y = \tan^2 x^2$.

Solution: Let $u = x^2$. Then $y = \tan^2 u$. By the chain rule we have

$$\frac{dy}{dx} = \frac{dy}{du}.\frac{du}{dx},$$

$$= 2\tan u \sec^2 u\,(2x)\,.$$

Here I did the chain rule on $y = \tan^2 u$ using the "inside/outside" way of thinking about it. Substituing u back into dy/dx we get

$$\frac{dy}{dx} = 4x\tan x^2 \sec^2 x^2\,.$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic. Also, see http://www.ballooncalculus.org/examples/reference.html for many of worked examples using this balloon approach.

1.6.4 The chain rule for higher derivatives

Just as the product rule and quotient rule can be extended to higher derivatives, so can the chain rule be extended. The chain rule may look simpler in structure than both of the former rules but care has to be taken when wanting the find higher derivatives of the chain rule. We will therefore take things steps by step when applying it to find such derivatives.

Let us therefore start with the standard chain rule

$$\frac{df}{dx} = \frac{df}{dg}\frac{dg}{dx}.$$

We now want to find the 2nd derivative version of this equation.

This means we want to do

$$\frac{d}{dx}\left(\frac{df}{dx}\right) = \frac{d}{dx}\left(\frac{df}{dg}\frac{dg}{dx}\right).$$

So we are going to need to use the product rule on the right hand side:

$$\frac{d^2f}{dx^2} = \underset{①}{\frac{d}{dx}\left(\frac{df}{dg}\right)}.\frac{dg}{dx} + \frac{df}{dg}.\underset{②}{\frac{d}{dx}\left(\frac{dg}{dx}\right)}.$$

The derivative of dg/dx in ② is simply going to be d^2g/dx^2. To find the derivative $d(df/dg)/dx$ in ① we need to apply the chain rule. The reason for this is that we can only ever differentiate a function with respect to its own independent variable. So in ① we can only do $d(df/dg)/dg$. But we don't want this. So in order to convert it into a derivative with respect to x (or any other variable) we use the chain rule.

Therefore we have

$$\frac{d^2f}{dx^2} = \frac{d}{dg}\left(\frac{df}{dg}\right).\frac{dg}{dx}.\frac{dg}{dx} + \frac{df}{dg}.\frac{d}{dx}\left(\frac{dg}{dx}\right).$$

Now we can perform all the relevant differentiations to get

$$\frac{d^2f}{dx^2} = \frac{d^2f}{dg^2}.\left(\frac{dg}{dx}\right)^2 + \frac{df}{dg}.\frac{d^2g}{dx^2}. \tag{19}$$

This is the second derivative formula for the chain rule.

<u>Example</u>

To find the second derivative of $f(x) = \sin(x^2 + x)$ we let $u = x^2 + x$ so that $f(x) = \sin u$. Applying (19) we get

$$\begin{aligned} \frac{d}{dx}(\sin(x^2+x)) &= \frac{d^2}{du^2}(\sin u).\left(\frac{d}{dx}(x^2+x)\right)^2 + \frac{d}{du}(\sin u)\frac{d^2}{dx^2}(x^2+x), \\ &= -\sin u \times (2x+1)^2 + \cos u \times (2), \\ &= -(2x+1)^2.\sin(x^2+x) + 2\cos(x^2+x). \end{aligned}$$

Similar formulae can be developed for higher order derivatives of the chain rule.

For example the formula for the 3rd derivative of the chain rule is

$$\frac{d^3f}{dg^3} = \frac{df}{dg}.\frac{d^3g}{dx^3} + 3\frac{d^2f}{dg^2}.\frac{d^2g}{dx^2}.\frac{dg}{dx} + \frac{d^3f}{dg^3}.\left(\frac{dg}{dx}\right)^3.$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.

1.7 Implicit differentiation

All the differentiation we have done so far has been done on functions of the form $y = f(x)$, in other words function in which the variables y and x were completely separate. For example, the function $y + x.y = 1$ can be transformed into a function $y = f(x)$ where the variables are completely separated, i.e. $y = 1/(1 + x)$.

However, there are functions in which the variables y and x cannot be totally separated. For example $y + x.y + y^2 = 1$ cannot be transformed by any algebra into the form $y = f(x)$. This function is then said to be an *implicit equation*.

How do we differentiate such functions? Well, consider again $y + x.y + y^2 = 1$. If we differentiate this, the second term will need to be differentiated by the product rule, and the third term will need to be differentiated by the chain rule (in fact it could also be differentiated by the product rule also).

So we have

$$\frac{d}{dx}(y) + \frac{d}{dx}(xy) + \frac{d}{dx}(y^2) = \frac{d}{dx}(1).$$

We know how to perform the differentiation on the first two terms, but how do we perform the differentiation of the third term? Given that we are looking for the derivative with respect to x, we can't differentiate y^2 as it stands. We can only differentiate it with respect to y, i.e. we can only do

$$\frac{d}{dy}(y^2) = 2y.$$

So how do we transform a derivative with respect to y into a derivative with respect to x?

By using the chain rule:

$$\frac{d}{dx}(y^2) = \frac{d}{dy}(y^2)\frac{dy}{dx} = 2y\frac{dy}{dx}.$$

So when differentiating one variable with respect to another we always use the chain rule. The complete derivative of our function above therefore becomes

$$\frac{dy}{dx} + y + x\frac{dy}{dx} + 2y\frac{dy}{dx} = 0\,.$$

Factorizing and simplifying gives

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

Examples

1) If $y = \sin(xy)$ then to differentiate implicit differentiation we have

$$\frac{d}{dx}(\sin(xy)) = \cos(xy).\frac{d}{dx}(xy)\,,$$

where I used the idea of "derivative of outside × derivative of inside". Now we use the product rule, within which we use implicit differentiation:

$$\frac{d}{dx}(\sin(xy)) = \cos(xy)\left(x\frac{dy}{dx} + y\right).$$

2) If $\sin y + x^2y^3 = 2y$ then

$$\begin{aligned}
\frac{d}{dx}(\sin y + x^2y^3) &= \frac{d}{dx}(2y)\,, \\
\frac{d}{dy}(\sin y).\frac{dy}{dx} + \frac{d}{dx}(x^2y^3) &= \frac{d}{dy}(2y).\frac{dy}{dx}, \\
\frac{d}{dy}(\sin y).\frac{dy}{dx} + y^3\frac{d}{dx}(x^2) + x^2.\frac{d}{dy}(y^3).\frac{dy}{dx} &= \frac{d}{dy}(2y).\frac{dy}{dx}.
\end{aligned}$$

Only now, having formally set up the derivatives with respect to the correct variable, can we actually perform differentiation. As such we get

$$\cos y\,.\frac{dy}{dx} + 2xy^3 + 3x^2y^2\frac{dy}{dx} = 2y\frac{dy}{dx}\,,$$

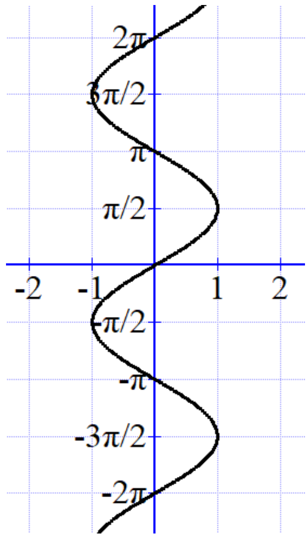
from which

$$\frac{dy}{dx} = -\frac{2x^3}{\cos y + 3x^2y^2 - 2y}.$$

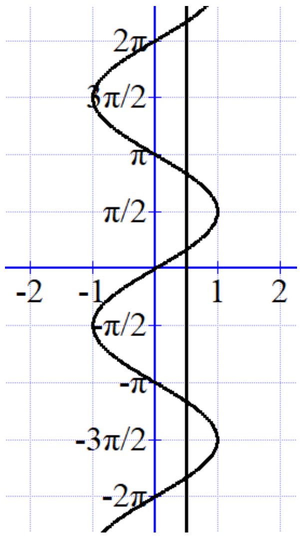
1.7.1 The derivative of inverse trig functions

Now that we know about implicit differentiation we can use this to find the derivative of inverse trig functions.

Firstly let us look at the inverse sine function. You might think that the inverse sine function is simply given as $y = \sin^{-1} x$. But if we draw this equation we get:



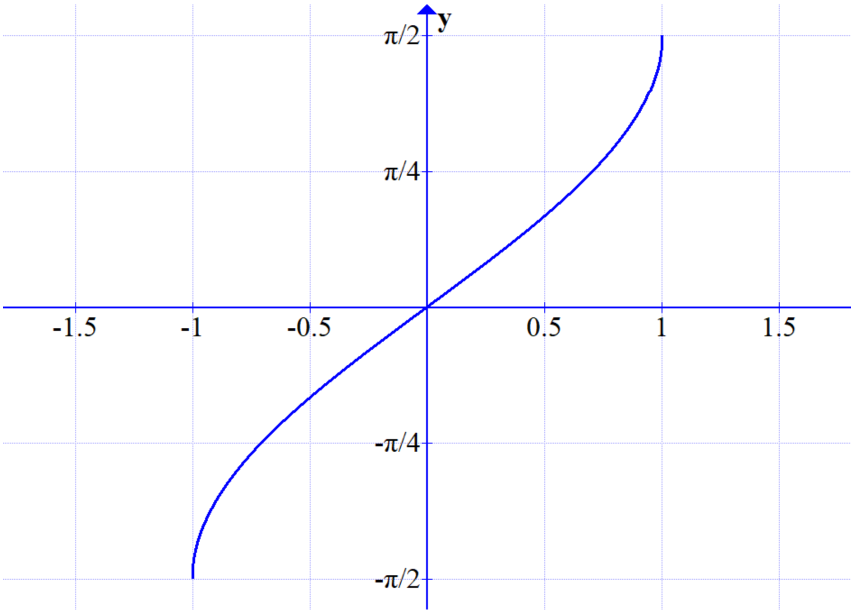
This is not a function. Remember that a function is defined as an equation whereby one or more *x* values gives *only one y* value. Here any single *x* value, (such as *x* = ½) will an infinite number of *y* values, as seen in the graph below by looking at the black vertical line:



In order to stop this from happening we restrict the range of y values for which arcsin(x) is valid. Therefore our inverse sine function is defined as:

$$y = \sin^{-1} x \text{ where } x \in [-1,1] \text{ and } y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],$$

the graph of which is now



It is important to know the range of y in any and all inverse trig functions since we will need to use this range information to help us get the correct derivative (more on this when we get to it).

Returning to our problem of finding the derivative of inverse trig functions, the fact is that we do not know how to differentiate arcsin (or the inverse of any other trig function). But we do know how to differentiate the sin of a function. So we convert the above to be sin y = x and then use implicit differentiation. The whole process goes like this:

$$\begin{aligned} y &= \sin^{-1} x \,, \\ \sin y &= x \,, \\ \frac{d}{dx}(\sin y) &= \frac{d}{dx}(x) \,, \\ \frac{d}{dy}(\sin y).\frac{dy}{dx} &= 1 \,, \\ \cos y.\frac{dy}{dx} &= 1 \,, \end{aligned}$$

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

We could stop here and say that this is our derivative of $\sin^{-1}x$, but we want an answer in terms of x not y. We can find such answer by replacing $\cos y$ by something in terms of x, and we do this by going back to our trig identities, namely $\cos^2 y + \sin^2 y = 1$. Making $\cos y$ the subject of the equation gives $\cos y = \pm\sqrt{1 - \sin^2 y}$.

But since $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, the function $\cos y$ is always positive in this interval, so we take the positive root as our answer, hence

$$\frac{dy}{dx} = \frac{1}{\sqrt{1 - \sin^2 y}}.$$

But $\sin y = x$ so we end up with

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

Exactly the same process can be used to find the derivatives of the other five trig functions: differentiate implicitly, use a trig identity to express the derivative in terms of x, and look at the graph of the inverse trig function to see if its gradient is positive of negative over its domain and (if necessary) choose the correct sign for the square root, or look at the derivative function to see what sign it has over its domain.

In summary we have

for $y = \cos^{-1} x$, $\frac{dy}{dx} = -\frac{1}{\sqrt{1 - x^2}}$,	for $y = \sec^{-1} x$, $\frac{dy}{dx} = \frac{1}{x\sqrt{x^2 - 1}}$,
for $y = \tan^{-1} x$, $\frac{dy}{dx} = \frac{1}{1 + x^2}$,	for $y = \operatorname{cosec}^{-1} x$, $\frac{dy}{dx} = -\frac{1}{x\sqrt{x^2 - 1}}$,
for $y = \cot^{-1} x$, $\frac{dy}{dx} = -\frac{1}{1 + x^2}$.	

Examples

1) If $y = \cos^{-1} x$ then in order to make this a function we need our domain and range to be $x \in [-1,1]$ and $y \in [0,\pi]$. If we want to differentiate $y = \cos^{-1} x^2$, then $x \in [-1,1]$ and y will be in a subrange of $y \in [0,\pi]$.

Differentiating $y = \cos^{-1} x^2$ we have

$$\begin{aligned} \cos y &= x^2 , \\ \frac{d}{dx}(\cos y) &= \frac{d}{dx}(x^2) , \\ \frac{d}{dy}(\cos y).\frac{dy}{dx} &= 2x , \\ -\sin y.\frac{dy}{dx} &= 2x , \\ \frac{dy}{dx} &= \frac{-2x}{\sin y} . \end{aligned}$$

By the trig identity $\cos^2 y + \sin^2 y = 1$ we have $\sin y = \pm\sqrt{1-\cos^2 y}$. By looking at the graph of $y = \cos^{-1} x^2$ we see that its gradient dy/dx is sometimes positive and sometimes negative. But we know that our range of y values is $y \in [0,\pi]$ and therefore that $\sin y$ is always positive in this range. So we take the positive square root, hence

$$\begin{aligned} \frac{dy}{dx} &= \frac{-2x}{\sqrt{1-\cos^2 y}} , \\ &= \frac{-2x}{\sqrt{1-x^4}} . \end{aligned}$$

2) If $y = 10\sin^{-1}(3/x)$ for all $x \geq 3$ then

$$\begin{aligned} \sin\left(\frac{y}{10}\right) &= 3/x , \\ \frac{1}{10}\cos\frac{y}{10}.\frac{dy}{dx} &= -\frac{3}{x^2} , \\ \frac{dy}{dx} &= \frac{-30}{x^2 \cos y/10} . \end{aligned}$$

By the trig identity $\cos^2(y/10) + \sin^2(y/10) = 1$ we have $\cos(y/10) = \pm\sqrt{1 - \sin^2(y/10)}$. Since $\cos(y/10)$ is positive in any part of the interval $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ we take the positive root, hence

$$\begin{aligned} \frac{dy}{dx} &= \frac{-30}{x^2\sqrt{1 - \sin^2(y/10)}}, \\ &= \frac{-30}{x^2\sqrt{1 - 9/x^2}}, \\ &= \frac{-30}{x\sqrt{x^2 - 9}}. \end{aligned}$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.

1.7.2 A proof of the power rule for rational exponents

We can use implicit differentiation to prove that, if $y = x^{p/q}$, where p and q are integers, then $dy/dx = \frac{p}{q}.x^{(p/q)-1}$:

$$\begin{aligned} y &= x^{\frac{p}{q}}, \\ y^q &= x^p, \\ \frac{d}{dx}(y^q) &= \frac{d}{dx}(x^p), \\ \frac{d}{dy}(y^q).\frac{dy}{dx} &= p.x^{p-1}, \\ q.y^{q-1}\frac{dy}{dx} &= p.x^{p-1}, \\ \frac{dy}{dx} &= \frac{p.x^{p-1}}{q.y^{q-1}}. \\ &= \frac{p}{q}.\frac{x^{p-1}}{(x^{p/q})^{q-1}}, \\ &= \frac{p}{q}.\frac{x^{p-1}}{x^{p-p/q}}, \\ &= \frac{p}{q}.x^{(p/q)-1}. \end{aligned}$$

1.7.3 Proof of the power rule for all real exponents

Having gone through implicit differentiation we can now prove the power rule for any real number n. Therefore what we are going to prove is that, if $y = x^n$, where n is a real number, then $dy/dx = n.x^{n-1}$. To do this we will convert our equation into log form.

So,

$$y = x^n ,$$

$$\ln y = \ln x^n ,$$

$$= n.\ln x ,$$

Differentiating implicitly we get

$$\frac{1}{y}.\frac{dy}{dx} = \frac{n}{x},$$

from which

$$\frac{dy}{dx} = n.\frac{y}{x},$$

$$= n.\frac{x^n}{x},$$

$$= n.x^{n-1} .$$

The reason this proof works for all real n (and not just integer or rational n) is because all operations (×, ÷, ln) done above are valid for real n. At no stage was any operation used that restricted n to be integer or rational only.

*1.7.4 The derivative of the general exponential y = a^x^*

We can now use implicit differentiation to find the first derivative of $y = a^x$ where a is any real number. Again we use logs:

$$y = a^x ,$$

$$\ln y = x.\ln a ,$$

$$\frac{1}{y}.\frac{dy}{dx} = \ln a ,$$

$$\frac{dy}{dx} = y.\ln a = x^a \ln a .$$

1.8 A study in differentiation

In this section we will look at some non-standard of derivatives, derivative rules and differentiation examples based on these. I have included this simply for the pleasure of learning more about differentiation. Also, the variation of the material presented below will set into relief all the work above, this helping to crystallise one's understanding of differentiation.

1.8.1 A fictitious definition for dy/dx

The standard defection for the first derivative of a function $f(x)$ is

$$f'(x) = \frac{df}{dx} = \lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x}. \qquad \text{(a)}$$

Consider now inventing a new definition for the derivative given by

$$f^*(x) = \frac{Df}{Dx} = \lim_{\delta x \to 0} \left(\frac{f(x+\delta x)}{f(x)} \right)^{1/\delta x}. \qquad \text{(b)}$$

Let us now go through seeing how $f^*(x)$ transforms $f(x)$ for some simple cases.

1) Let $f(x) = x$. What is Df/Dx? Well, using (b) we have

$$\frac{Df}{Dx} = \lim_{\delta x \to 0} \left(\frac{x+\delta x}{x} \right)^{1/\delta x},$$

$$= \lim_{\delta x \to 0} \left(1 + \frac{\delta x}{x} \right)^{1/\delta x}.$$

This limit is very close to being of the form $\lim_{z \to 0}(1+z)^{1/z}$. If the power/exponent of the above limit was $x/\delta x$ then we would have this exact form, and we know that this limit evaluates to e.

Our aim is therefore going to be to transform the above limit expression in such a way as to get the form which give us the answer e. To do this we will first have to take logs. This will allow us to use the rules of logs to convert the term inside the limit into the form we want.

Hence

$$\ln\left(\frac{Df}{Dx}\right) = \ln\left(\lim_{\delta x \to 0} \left(1 + \frac{\delta x}{x} \right)^{1/\delta x} \right).$$

For what we are doing, the log of the limit is the same as the limit of the log. Hence

$$\ln\left(\frac{Df}{Dx}\right) = \lim_{\delta x\to 0}\left(\ln\left(1+\frac{\delta x}{x}\right)^{1/\delta x}\right).$$

We can now do the trick of multiplying by x/x which will help us later on to get the exponential form we are looking for:

$$\ln\left(\frac{Df}{Dx}\right) = \frac{x}{x}.\lim_{\delta x\to 0}\left(\ln\left(1+\frac{\delta x}{x}\right)^{1/\delta x}\right).$$

Normally when x is independent of dx we can factorise it out of the limit. But this also means that we can multiply it into the limit, which is what we will now do:

$$\ln\left(\frac{Df}{Dx}\right) = \frac{1}{x}.\lim_{\delta x\to 0}\left(x\ln\left(1+\frac{\delta x}{x}\right)^{1/\delta x}\right).$$

Using rules of logs we now have

$$\begin{aligned}\ln\left(\frac{Df}{Dx}\right) &= \frac{1}{x}.\lim_{\delta x\to 0}\left(\ln\left(1+\frac{\delta x}{x}\right)^{x/\delta x}\right),\\ &= \frac{1}{x}.\ln\left(\lim_{\delta x\to 0}\left(1+\frac{\delta x}{x}\right)^{x/\delta x}\right).\end{aligned}$$

The limit is now in a form we can evaluate. Hence we have

$$\ln\left(\frac{Df}{Dx}\right) = \frac{1}{x}.\ln e = 1\,.$$

Taking inverse logs will give us the answer. Hence

$$\text{when } f(x) = x \text{ we have } \quad \frac{Df}{Dx} = e^{1/x}\,.$$

This process can be repeated to find Df/Dx when $f(x) = x^2$. The full derivation is shown below without commentary since exactly the same thinking is used as above.

$$\frac{Df}{Dx} = \lim_{\delta x\to 0}\left\{\frac{(x+\delta x)^2}{x^2}\right\}^{1/\delta x},$$

$$= \lim_{\delta x \to 0} \left\{ \frac{x^2 \left(1 + \frac{\delta x}{x}\right)^2}{x^2} \right\}^{1/\delta x} ,$$

$$= \lim_{\delta x \to 0} \left(1 + \frac{\delta x}{x}\right)^{2/\delta x} .$$

Taking logs and subsequently using log rules we have

$$\ln\left(\frac{Df}{Dx}\right) = \ln\left[\lim_{\delta x \to 0} \left(1 + \frac{\delta x}{x}\right)^{2/\delta x}\right],$$

$$= \lim_{\delta x \to 0} \left[\ln\left(1 + \frac{\delta x}{x}\right)^{2/\delta x}\right],$$

$$= \frac{x}{x}. \lim_{\delta x \to 0} \left[\ln\left(1 + \frac{\delta x}{x}\right)^{2/\delta x}\right],$$

$$= \frac{1}{x}. \lim_{\delta x \to 0} \left[x \ln\left(1 + \frac{\delta x}{x}\right)^{2/\delta x}\right],$$

$$= \frac{1}{x}. \lim_{\delta x \to 0} \left[\ln\left(1 + \frac{\delta x}{x}\right)^{2x/\delta x}\right],$$

$$= \frac{1}{x}. \ln\left[\lim_{\delta x \to 0} \left(1 + \frac{\delta x}{x}\right)^{2x/\delta x}\right],$$

$$= \frac{1}{x}. \ln\left[\lim_{\delta x \to 0} \left(1 + \frac{\delta x}{x}\right)^{\frac{x}{\delta x}}\right]^2 ,$$

$$= \frac{1}{x}. \ln e^2 ,$$

$$= \ln e^{2/x} .$$

Taking inverse logs on both sides will give us the answer. Hence

$$\text{when } f(x) = x^2 \text{ we have } \quad \frac{Df}{Dx} = e^{2/x}\,.$$

In general it can be shown that when $f(x) = x^n$ we have $Df/Dx = e^{n/x}$, where n is a real number.

2) Let $y = k.f(x)$, where k is a constant. What is the derivative of this function according to (b)? Well, applying (b) we have

$$\frac{Dy}{Dx} = \frac{D}{Dx}\big(k.f(x)\big) = \lim_{\delta x \to 0}\left(\frac{k.f(x+\delta x)}{k.f(x)}\right)^{1/\delta x}.$$

Since we can cancel k we end up with

$$\frac{D}{Dx}\big(k.f(x)\big) = \lim_{\delta x \to 0}\left(\frac{f(x+\delta x)}{f(x)}\right)^{1/\delta x} = \frac{Df}{Dx}.$$

So the constant has no effect under differentiation. For example,

$$\text{if } f(x) = x \text{ then } \quad \frac{Df}{Dx} = e^{1/x}\,,$$

$$\text{if } f(x) = 2x \text{ then } \quad \frac{Df}{Dx} = e^{1/x}\,,$$

$$\text{if } f(x) = 3x \text{ then } \quad \frac{Df}{Dx} = e^{1/x}\,,$$

etc.

3) Let $h(x) = f.g$ where f and g are two functions of x. Then the formula for the derivative of the product is given as follows:

$$\begin{aligned}\frac{Dh}{Dx} = \frac{D}{Dx}(f.g) \quad &= \quad \lim_{\delta x \to 0}\left(\frac{f(x+\delta x)}{f(x)}.\frac{g(x+\delta x)}{g(x)}\right)^{1/\delta x},\\ &= \quad \lim_{\delta x \to 0}\left\{\left(\frac{f(x+\delta x)}{f(x)}\right)^{1/\delta x}.\left(\frac{g(x+\delta x)}{g(x)}\right)^{1/\delta x}\right\},\\ &= \quad \lim_{\delta x \to 0}\left(\frac{f(x+\delta x)}{f(x)}\right)^{1/\delta x}.\lim_{\delta x \to 0}\left(\frac{g(x+\delta x)}{g(x)}\right)^{1/\delta x}.\end{aligned}$$

But these last limits are simply the definition of this new derivative. Hence

$$\text{when } h(x) = f.g \text{ we have } \quad \frac{Dh}{Dx} = \frac{Df}{Dx}\frac{Dg}{Dx}.$$

4) Let $h(x) = f/g$ where f and g are two functions of x. Then the formula for the derivative of the quotient is given as follows:

$$\begin{aligned}
\frac{Dh}{Dx} = \frac{D}{Dx}(f.g) \quad &= \quad \lim_{\delta x \to 0}\left(\frac{f(x+\delta x)}{f(x)}.\frac{g(x)}{g(x+\delta x)}\right)^{1/\delta x}, \\
&= \quad \lim_{\delta x \to 0}\left\{\left(\frac{f(x+\delta x)}{f(x)}\right)^{1/\delta x}.\left(\frac{g(x)}{g(x+\delta x)}\right)^{1/\delta x}\right\}, \\
&= \quad \lim_{\delta x \to 0}\left(\frac{f(x+\delta x)}{f(x)}\right)^{1/\delta x}.\lim_{\delta x \to 0}\left(\frac{g(x)}{g(x+\delta x)}\right)^{1/\delta x}.
\end{aligned}$$

The first limit is simply the definition of this new derivative. The second limit is the reciprocal of the derivative (this has to be proved, and is left as an exercise). Hence

$$\text{when } h(x) = f/g \text{ we have } \quad \frac{Dh}{Dx} = \frac{Df}{Dx}\frac{1}{Dg/Dx}.$$

5) Let $f(x) = e^x$. What is Df/Dx? Well, applying (b) we have

$$\frac{Df}{Dx} = \lim_{\delta x \to 0}\left(\frac{e^{(x+\delta x)}}{e^x}\right)^{1/\delta x}.$$

By the rule of exponents we can rewrite this as

$$\begin{aligned}
\frac{Df}{Dx} \quad &= \quad \lim_{\delta x \to 0}\left(e^{\delta x}\right)^{1/\delta x}, \\
&= \quad \lim_{\delta x \to 0} e = e,
\end{aligned}$$

since e is independent of δx (e is simply a fixed number). In other words, according to this new derivative when $f(x) = e^x$, Df/Dx = e = constant.

6) By taking logs of our new derivative we can show that there is a connection between df/dx and Df/Dx. To show this we will start with

$$\frac{\delta f}{\delta x} = \left\{\frac{f(x+\delta x)}{f(x)}\right\}^{1/\delta x} .$$

Taking logs and applying rules of logs we get

$$\begin{aligned}\ln\left(\frac{\delta f}{\delta x}\right) &= \ln\left\{\frac{f(x+\delta x)}{f(x)}\right\}^{1/\delta x} ,\\ &= \frac{1}{\delta x}.\ln\left\{\frac{f(x+\delta x)}{f(x)}\right\},\\ &= \frac{\ln f(x+\delta x) - \ln f(x)}{\delta x}.\end{aligned}$$

We can now apply and perform the limit as $\delta x \to 0$:

$$\lim_{\delta x\to 0} \ln\left(\frac{\delta f}{\delta x}\right) = \lim_{\delta x\to 0}\left\{\frac{\ln f(x+\delta x) - \ln f(x)}{\delta x}\right\}.$$

For certain mathematical reasons which are beyond the scope of these notes we can swap the log and limit operations. As for the right hand side, this is just the definition of the derivative of $\ln x$. So we have

$$\ln\left(\frac{Df}{Dx}\right) = \frac{d}{dx}(\ln f(x)) .$$

We know that $d(\ln f(x))/dx = f'(x)/f(x)$ hence the above becomes

$$\ln\left(\frac{Df}{Dx}\right) = \frac{f'(x)}{f(x)}. \qquad (*)$$

Adopting our notation of f^* for Df/Dx we then have the following three forms for the connection between Df/Dx and df/dx:

$\ln f^*(x) = \frac{f'(x)}{f(x)}$	$f'(x) = f(x)\ln f^*(x)$,	$f^*(x) = e^{f'(x)/f(x)}$,
by leaving (*) as it is;	by cross multiplying in (*);	by taking inverse logs;
(1)	(2)	(3)

We should recognise form (3) since this is the form we got when differentiation the functions $f(x) = x$ and $f(x) = x^2$.

We can now use (3) to find Df/Dx for $f(x) = x^n$. In this case we have

$$\frac{f'(x)}{f(x)} = \frac{n.x^{n-1}}{x^n} = \frac{n}{x},$$

Hence

$$\frac{Df}{Dx} = e^{n/x}.$$

Form (2) can be used to confirm the results we get for Df/Dx. For example, when $f(x) = x$ we have that $Df/Dx = e^{1/x}$. We also know that $df/dx = 1$. So by (2) we have the left hand side to be

$$\frac{df}{dx} = 1$$

and the right hand side to be

$$\frac{Df}{Dx} = x.\ln e^{1/x} = x\frac{1}{x}\ln e = 1,$$

so our Df/Dx computation was correct. The same checks can be done for $f(x) = x^2$, $f(x) = x^n$ and $f(x) = e^x$ (left as an exercise).

1.8.2 Another fictitious definition for dy/dx

Let us created another fictitious definition for the first derivative:

$$f^{\sim}(x) = \frac{\Delta f}{\Delta x} = \lim_{\delta x \to 0} \frac{f^2(x+\delta x) - f^2(x)}{\delta x}. \qquad \text{(c)}$$

For a constant k, and two functions $f(x)$ and $g(x)$ the rules for $\Delta(kf)/\Delta x$, $\Delta(f.g)/\Delta x$, and $\Delta(f/g)/\Delta x$ can be derived in a straightforward manner. I will therefore leave it as an exercise for you to show that

$$\frac{\Delta(kf)}{\Delta x} = k.\frac{\Delta f}{\Delta x}, \qquad \frac{\Delta}{\Delta x}(f.g) = g^2\frac{\Delta f}{\Delta x} + f^2\frac{\Delta g}{\Delta x}, \qquad \frac{\Delta}{\Delta x}\left(\frac{f}{g}\right) = \frac{g^2\frac{\Delta f}{\Delta x} - f^2\frac{\Delta g}{\Delta x}}{g^4}.$$

Finding the expression for $\Delta(f \pm g)/\Delta x$ is more involved, and is not the same as that of $d(f \pm g)/dx$, so I will leave this until the end of this section. What we will do here is to go through the derivative of some specific functions, as well as some properties of $\Delta f/\Delta x$.

Hence if $f(x) = x$ we have

$$\begin{aligned} \frac{\Delta f}{\Delta x} &= \lim_{\delta x \to 0} \frac{(x + \delta x)^2 - x^2}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{(x^2 + 2x.\delta x + (\delta x)^2) - x^2}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{2x.\delta x + (\delta x)^2}{\delta x}, \\ &= \lim_{\delta x \to 0} (2x + \delta x), \\ &= 2x. \end{aligned}$$

So the derivative of $f(x) = x$ is $\Delta f/\Delta x = 2x$. By the same process it can be shown that the derivative $f(x) = x^2$ is $\Delta f/\Delta x = 4x^3$, the derivative $f(x) = x^3$ is $\Delta f/\Delta x = 6x^5$, etc, and that in general, when $n \in \mathbb{N}$ (i.e. n is a positive integer), the derivative $f(x) = x^n$ is $\Delta f/\Delta x = (n + 1)x^n$.

A good question to ask is, Is there a connection between $\Delta f/\Delta x$ and df/dx? To see if there is let us see if we can spot a pattern between the results of the two derivatives for x^n for various values of n:

$\boldsymbol{f(x)}$	$\boldsymbol{\frac{df}{dx}}$	$\boldsymbol{\frac{\Delta f}{\Delta x}}$
x	1	$2x$
x^2	$2x$	$4x^3$
x^3	$3x^2$	$6x^5$
x^4	$4x^3$	$8x^7$
x^5	$5x^4$	$10x^9$
…	…	…

Looking at the coefficients of each derivative we see that

coefficient of $\Delta f/\Delta x$ = 2 × coefficient of df/dx.

Looking at the powers of each derivative we see that

exponent of $f(x)$ + exponent of df/dx = exponent of $\Delta f/\Delta x$.

This implies that we can multiply df/dx by $f(x)$. Hence it looks like the connection between $\Delta f/\Delta x$ and df/dx is $\Delta f/\Delta x = 2f(x) \times \, df/dx$. We will prove this in a moment, and show that this is true for any function $f(x)$.

Let us now find $\Delta f/\Delta x$ for $f(x) = e^x$. In this case we have

$$\begin{aligned}\frac{\Delta f}{\Delta x} &= \lim_{\delta x \to 0} \frac{e^{2(x+\delta x)} - e^2}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{e^{2x+2\delta x} - e^2}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{e^{2x}e^{2\delta x} - e^2}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{e^{2x}\left(e^{2\delta x} - 1\right)}{\delta x}, \\ &= e^{2x} \lim_{\delta x \to 0} \frac{e^{2\delta x} - 1}{\delta x}.\end{aligned}$$

We now need to find a way of evaluating the limit. This is not a standard one so we have to find a way of converting it into a standard form. Notice that $e^{2\delta x} = \left(e^{\delta x}\right)^2$ so the numerator of the limit is really a difference of two squares. Hence we can write

$$\frac{\Delta f}{\Delta x} = e^{2x} \lim_{\delta x \to 0} \frac{\left(e^{\delta x} + 1\right)\left(e^{\delta x} - 1\right)}{\delta x}.$$

Grouping terms appropriately we have

$$\frac{\Delta f}{\Delta x} = e^{2x} \lim_{\delta x \to 0}\left(e^{\delta x} + 1\right) \lim_{\delta x \to 0} \frac{\left(e^{\delta x} - 1\right)}{\delta x}.$$

The first limit becomes 2 ($e^{\delta x} = 1$ as $\delta x \to 0$), and our previous work we know the second limit to be equal to 1. Hence we have

$$\frac{\Delta f}{\Delta x} = 2e^{2x}.$$

Let us now return to the formula $\Delta f/\Delta x = 2f(x) \times df/dx$ which we guessed to be the general connection between $\Delta f/\Delta x$ and df/dx in the case of $f(x) = x, x^2, x^3$, etc. We can show that this connection is true for any function $f(x)$ as follows: given that

$$\frac{\Delta f}{\Delta x} = \lim_{\delta x \to 0} \frac{f^2(x+\delta x) - f^2(x)}{\delta x}$$

we want to use algebra on $f^2(x+\delta x) - f^2(x)$ in order to change this into combinations of $f(x)$ and $f(x+\delta x)$ (these terms forming part of the defintion of *df*/*dx*). To do this note that $f^2(x+\delta x) - f^2(x)$ is the difference of two squares: $a^2 - b^2 = (a-b)(a+b)$.

Hence $\Delta f/\Delta x$ can be written as

$$\frac{\Delta f}{\Delta x} = \lim_{\delta x \to 0} \frac{[f(x+\delta x) + f(x)][f(x+\delta x) - f(x)]}{\delta x}.$$

Now we group terms in such as way as to recover the derivative *df*/*dx*, viz:

$$\begin{aligned}\frac{\Delta f}{\Delta x} &= \lim_{\delta x \to 0} \left\{[f(x+\delta x) + f(x)] \times \frac{[f(x+\delta x) - f(x)]}{\delta x}\right\}, \\ &= \lim_{\delta x \to 0}[f(x+\delta x) + f(x)] \times \lim_{\delta x \to 0} \frac{[f(x+\delta x) - f(x)]}{\delta x},\end{aligned}$$

The second limit is *df*/*dx*. The first limit becomes $2f(x)$ as $\delta x \to 0$. So we end up with

$$\frac{\Delta f}{\Delta x} = 2f(x)\frac{df}{dx}. \qquad (**)$$

Adopting our notation of $f^{\sim}$ for $\Delta f/\Delta x$ we then have the following forms for the connection between $\Delta f/\Delta x$ and *df*/*dx*:

- $f^{\sim}(x) = 2f(x)f'(x)$ by leaving (**) as it is; (4)
- $f'(x) = \frac{1}{2f(x)} f^{\sim}(x)$ by cross multiplying in (**). (5)

Expression (4) allows us to derive $f^{\sim}(x)$ very easily from $f'(x)$. Expression (5) allows us to derive $f'(x)$ from $f^{\sim}(x)$.

We can now use (4) to formally find $\Delta f/\Delta x$ for $f(x) = x^n$. In this case we have

$$f^{\sim}(x) = 2(x^n)n(x^{n-1}) = 2nx^{2n-1}. \qquad (***)$$

(we would have got the same result if we had used definition (c) to prove this from 1st principles).

Our previous examples of finding $\Delta f/\Delta x$ for x, x^2, x^3, etc. were only valid when n was a positive integer. This is because we were using the binomial theorem to expand the term $f^2(x+\delta x)$. This was also the case (in the Differentiation I notes) when we were finding *df*/*dx* for $f(x) = x^n$ when n was only a positive integer. Only later (in these notes) did we go on to show that the formula for *df*/*dx* for $f(x) = x^n$ applied for all real values of n. Similary, the formula (***) now applies for all real n since, by this last fact, we now know that equation (**) applies for all real n.

We previously showed that the derivative $\Delta f/\Delta x$ for $f(x) = e^x$ was $\Delta f/\Delta x = 2e^{2x}$. From (4) we have

$$f^{\sim}(x) = 2e^x(e^x) = 2e^{2x}$$

which is the result we expect.

What about the derivative a sum or difference? What is $\Delta h/\Delta x$ when $h(x) = f(x) \pm g(x)$? I have left this until now because the proof of this is more involved than what we have done up to now. We will derive the sum and difference formulae separately since these are different

Hence, given $h(x) = f(x) + g(x)$ we have from definition (c)

$$\frac{\Delta h}{\Delta x} = \lim_{\delta x \to 0}\left\{\frac{[f(x+\delta x) + g(x+\delta x)]^2 - [f(x) + g(x)]^2}{\delta x}\right\},$$

For practical purposes let us simplify the notation: let $F = f(x+\delta x), G = g(x+\delta x), f = f(x)$ and $g = g(x)$. Then we have

$$\begin{aligned}\frac{\Delta h}{\Delta x} &= \lim_{\delta x \to 0}\left\{\frac{(F+G)^2 - (f+g)^2}{\delta x}\right\},\\ &= \lim_{\delta x \to 0}\left\{\frac{F^2 + 2FG + G^2 - f^2 - 2fg - g^2}{\delta x}\right\}.\end{aligned}$$

Grouping terms accordingly we have

$$\begin{aligned}\frac{\Delta h}{\Delta x} &= \lim_{\delta x \to 0} \left\{ \frac{F^2 - f^2}{\delta x} + \frac{G^2 - g^2}{\delta x} + \frac{2FG - 2fg}{\delta x} \right\}, \\ &= \lim_{\delta x \to 0} \frac{F^2 - f^2}{\delta x} + \lim_{\delta x \to 0} \frac{G^2 - g}{\delta x} + \lim_{\delta x \to 0} \frac{2FG - 2fg}{\delta x} \ . \qquad \{*\}\end{aligned}$$

Recalling our notation for F and G the first two limits are the definitions of the derivatives for $f(x)$ and $g(x)$. We now have to find a way of tranforming the third limit into something we can evaluate.

What we are looking for is to recover the structure of the derivative, this being a difference of two squares. We therefore need to be judicious in our choice of algebraic trick in order to recover this difference of squares structure.

So to start with we will add and subtract the term $2fG$ to the numerator in order to be able to factorise and obtain the difference aspect we are looking for:

$$\begin{aligned}\lim_{\delta x \to 0} \frac{2FG - 2fg}{\delta x} &= \lim_{\delta x \to 0} \frac{2FG - 2fG + 2fG - 2fg}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{2G(F - f) + 2f(G - g)}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{2G(F - f)}{\delta x} + \lim_{\delta x \to 0} \frac{2f(G - g)}{\delta x} .\end{aligned}$$

If each term in the brackets had been squared terms then the limits would equal the derivatives of $f(x)$ and $g(x)$. But they are not. However, we can transform the terms in the brackets into squared terms by doing the following transformation:

$$\lim_{\delta x \to 0} \frac{2FG - 2fg}{\delta x} = \lim_{\delta x \to 0} \frac{2G(F - f)}{\delta x} . \frac{F + f}{F + f} + \lim_{\delta x \to 0} \frac{2f(G - g)}{\delta x} . \frac{G + g}{G + g} \ .$$

We are effectively multiplying each limit by 1. Now all we need to do is multiply out the numerators in order to get the difference of squares we need.

$$\lim_{\delta x\to 0}\frac{2FG-2fg}{\delta x}=\lim_{\delta x\to 0}\frac{2G(F^2-f^2)}{(F+f)\delta x}+\lim_{\delta x\to 0}\frac{2f(G^2-g^2)}{(G+g)\delta x}.$$

Regrouping terms as necessary we have

$$\lim_{\delta x\to 0}\frac{2FG-2fg}{\delta x}=\lim_{\delta x\to 0}\frac{2G}{F+f}.\lim_{\delta x\to 0}\frac{F^2-f^2}{\delta x}+\lim_{\delta x\to 0}\frac{2f}{G+g}.\lim_{\delta x\to 0}\frac{G^2-g^2}{\delta x}.$$

Recalling our notation $F=f(x+\delta x), G=g(x+\delta x), f=f(x)$ and $g=g(x)$ we have

$$\lim_{\delta x\to 0}\frac{2FG-2fg}{\delta x}=\frac{\Delta f}{\Delta x}.\lim_{\delta x\to 0}\frac{2G}{F+f}+\frac{\Delta g}{\Delta x}.\lim_{\delta x\to 0}\frac{2f}{G+g}.$$

Now we have to evaluate the remaining limits. For the first limits we have

$$\lim_{\delta x\to 0}\frac{2G}{F+f}=\lim_{\delta x\to 0}\frac{2g(x+\delta x)}{f(x+\delta x)+f}=\frac{g}{f}\quad\text{and}\quad\lim_{\delta x\to 0}\frac{2f}{G+g}=\lim_{\delta x\to 0}\frac{2f(x)}{g(x+\delta x)+g(x)}=\frac{f}{g}$$

($g(x+\delta x)\to g(x)$ and $f(x+\delta x)\to f(x)$ as $x\to 0$). These results are so because $f(x)$ and $g(x)$ are constinuous functions and neither the numerator nor the denomenator approaches 0 as $x\to 0$ (assuming $f(x)$ and $g(x)$ are not zero functions).

(see http://www.stumblingrobot.com/2015/09/19/explore-properties-of-an-alternative-definition-of-the-derivative/ for an alternative proof to this where instead of doing two separate algebraic tricks of adding and subtracting $2fG$ and then multiplying by $(F+f)/(F+f)$ and $(G+g)/(G+g)$ the author uses one trick of multiplying by $(FG+fg)/(FG+fg)$ in the thirs limit of expression {*} above).

Hence the derivative of $h(x)=f(x)+g(x)$ is given by

$$h^{\sim}(x)=f^{\sim}(x)+g^{\sim}(x)+\frac{g}{f}f^{\sim}(x)+\frac{f}{g}g^{\sim}(x).$$

A similar approach can be used to show that, if $h(x)=f(x)-g(x)$ then

$$h^{\sim}(x)=f^{\sim}(x)+g^{\sim}(x)-\frac{g}{f}f^{\sim}(x)+\frac{f}{g}g^{\sim}(x).$$

1.8.3 A problem

Let $f(x)$ and $g(x)$ be two functions which can be differentiated as often as we like. Given that $f'g' = 1$ and that $h(x) = f(x).g(x)$ show that

$$\frac{h'''}{h} = \frac{f'''}{f} + \frac{g'''}{g}.$$

Derive the equation

$$\frac{h^{(iv)}}{h} = \frac{f^{(iv)}}{f} + \frac{g^{(iv)}}{g}$$

and state the necessary conditions for this to be true.

<u>Solution</u>

We know that $h(x) = f(x).g(x)$. By the product rule we have

$$h' = f'g + g'f\ ,$$

$$h'' = f''g + 2f'g' + g''f\ .$$

Since $f'g' = 1$ we have

$$h'' = f''g + g''f + 2\ .$$

Differentiating again we get

$$h''' = f'''g + f''g' + g''f' + g'''f\ . \qquad (*)$$

But our condition $f'g' = 1$ implies that $f''g' + f'g'' = 0$ (by the product rule) so (*) becomes

$$h''' = f'''g + g'''f\ .$$

Dividing by h we get

$$\frac{h'''}{h} = \frac{f'''}{f} + \frac{g'''}{g}.$$

To prove the fourth derivative equation we differentiate the third derivative equation:

$$h^{iv} = f^{iv}g + f'''g' + f'g''' + g^{iv}f\ . \qquad (**)$$

In order to get the required fourth derivative equation we need the sum of central two terms to equal zero. We already know that $f'g' = 1$ implies $f''g' + f'g'' = 0$, but does it also imply $f'''g' + f'g''' = 0$? To find out, we differentiate. Therefore

$$f''g' + f'g'' = 0$$

which implies $$f'''g' + f''g'' + g''f'' + g'''f' = 0.$$

So in order for $f'''g' + f'g''' = 0$ we need $f''g'' + f''g'' = 2f''g'' = 0$. If this is true the central two terms in (**) become zero and (**) becomes

$$h^{iv} = f^{iv}g + g^{iv}f\ .$$

Therefore

$$\frac{h^{iv}}{h} = \frac{f^{iv}}{f} + \frac{g^{iv}}{g}$$

only if $f'g' = 1$ and $f''g'' = 0$. Similar results could be proved for higher derivatives.

1.8.4 An alternative proof of the product rule and quotient rule

This section is adapted from *The Product and Quotient Rules Revisited*, Roger Eggleton and Vladimir Kustov, *The College Mathematics Journal*, Vol. 42, No. 4 (September 2011), pp. 323-326

Having previously shown that the derivative of ln x is $1/x$, and having derived the process of the chain rule above we can now go through a way of deriving the product rule and quotient rule for differentiation by using logs. Doing it this way will give a more elegant set of rules, and which share many structural features in common.

Let us start with $f(x) = u.v$, where u and v are functions of x, and let us take the log of bith sides:

$$\begin{aligned} f &= u.v\ , \\ \ln f &= \ln u + \ln v\ , \\ \frac{f'}{f} &= \frac{u'}{u} + \frac{v'}{v}\ . \end{aligned}$$

Cross multiplying by f, and remembering that $f = u.v$, we end up with

$$f' = (u.v)' = u.v\left(\frac{u'}{u} + \frac{v'}{v}\right). \qquad (20)$$

Examples

1) When $f(x) = xe^x$ we have by (20)

$$\frac{df}{dx} = x.e^x\left(\frac{1}{x}+\frac{e^x}{e^x}\right),$$

$$= e^x + xe^x\,.$$

2) When $y = (\ln x)/x$ we have by (20)

$$\frac{dy}{dx} = x^{-1}\ln x\left(\frac{-1/x^2}{1/x}+\frac{1/x}{\ln x}\right),$$

$$= \frac{(1-\ln x)}{x^2}\,.$$

Using the same approach, when $f(x) = u/v$ we get

$$f = \frac{u}{v},$$

$$\ln f = \ln u - \ln v\,,$$

$$\frac{f'}{f} = \frac{u'}{u}-\frac{v'}{v}\,.$$

Cross multiplying by f, and remembering that $f = u/v$, we end up with

$$f' = \left(\frac{u}{v}\right)' = \frac{u}{v}\left(\frac{u'}{u}-\frac{v'}{v}\right). \qquad (21)$$

Examples

1) When $y = (\ln x)/x$ we have by (21)

$$\frac{dy}{dx} = \frac{\ln x}{x}\left(\frac{1/x}{\ln x}-\frac{1}{x}\right),$$

$$= \frac{(1-\ln x)}{x^2},$$

which is the same answer as in the previous example.

2) When $f(x) = \sin x / \cos x$ we have by (21)

$$\begin{aligned} \frac{dy}{dx} &= \frac{\sin x}{\cos x}\left(\frac{\cos x}{\sin x} - \frac{-\sin x}{\cos x}\right), \\ &= 1 + \frac{\sin^2 x}{\cos^2 x}, \\ &= 1 + \tan^2 x \ , \\ &= \sec^2 x \, . \end{aligned}$$

We know this is the right answer since $\sin x / \cos x \ = \tan x$, and the derivative of $\tan x$ is $\sec^2 x$.

Note that we could also have derived (20) and (21) from the original versions of the product rule and quotient rule by relevant factorisations:

$$\begin{aligned} u.v' + v.u' &= u\left(v' + v\frac{u'}{u}\right), \\ &= u.v\left(\frac{v'}{v} + \frac{u'}{u}\right). \end{aligned}$$

$$\begin{aligned} \frac{v.u' - u.v'}{v^2} &= \left(\frac{1}{v^2}\right)u\left(v\frac{u'}{u} - v'\right), \\ &= \left(\frac{1}{v^2}\right)u.v\left(\frac{u'}{u} - \frac{v'}{v}\right), \\ &= \frac{u}{v}\left(\frac{u'}{u} - \frac{v'}{v}\right). \end{aligned}$$

Formula (21) has the advantage of simplifying the structure of the derivative of *f*(*x*) as compared to the standard quotient rule. It has a simple and symmetric-like pattern between the parts of *f*(*x*) and their derivatives. Its disadvantage is that it sometimes requires more algebra in order to get the simplified final result, particularly if our function is a rational polynomial.

This pattern extends to a function contains more that two terms. So for three functions *f*(*x*) containing three terms $u(x)$, $v(x)$, and $w(x)$ we have

$$f' = (u.v.w)' = u.v.w\left(\frac{u'}{u} + \frac{v'}{v} + \frac{w'}{w}\right), \qquad (22)$$

and

$$f' = \left(\frac{\frac{u}{v}}{w}\right)' = \left(\frac{u}{vw}\right)' = \frac{u}{vw}\left(\frac{u'}{u} - \frac{v'}{v} - \frac{w'}{w}\right). \qquad (23)$$

Second derivatives for the product and quotient rules can be found via (20) and (21). For the alternative product rule we could differentiate (20) as is, or we could rewrite it as

$$\frac{(u.v)'}{uv} = \frac{u'}{u} + \frac{v'}{v} \tag{a}$$

and differentiate:

$$\left(\frac{(u.v)'}{uv}\right)' = \left(\frac{u'}{u} + \frac{v'}{v}\right)' . \tag{b}$$

Here we will differentiate the left hand side and the right hand side separately, and then equate the results.

For the right hand side of (b) we have

$$\begin{aligned}
\left(\frac{u'}{u} + \frac{v'}{v}\right)' &= \left(\frac{u'}{u}\right)' + \left(\frac{v'}{v}\right)' , \\
&= \frac{u'}{u}\left(\frac{u''}{u'} - \frac{u'}{u}\right) + \frac{v'}{v}\left(\frac{v''}{v'} - \frac{v'}{v}\right), \\
&= \frac{u'}{u}\frac{u''}{u'} - \frac{u'}{u}\frac{u'}{u} + \frac{v'}{v}\frac{v''}{v'} - \frac{v'}{v}\frac{v'}{v}, \\
&= \frac{u''}{u} + \frac{v''}{v} - \left(\frac{u'}{u}\right)^2 - \left(\frac{v'}{v}\right)^2 . \qquad \text{(c)}
\end{aligned}$$

For the left hand side of (b) we have

$$\left(\frac{(u.v)'}{uv}\right)' = \frac{(uv)'}{uv}\left(\frac{(u.v)''}{(uv)'} - \frac{(uv)'}{uv}\right).$$

Simplfying the right hand side we get

$$\left(\frac{(u.v)'}{uv}\right)' = \frac{(uv)''}{uv} - \left[\frac{(uv)'}{uv}\right]^2 .$$

The term inside the square brackets is simply the alternative product rule in the form of (a) above. So we now have

$$\left(\frac{(u.v)'}{uv}\right)' = \frac{(uv)''}{uv} - \left(\frac{u'}{u} + \frac{v'}{v}\right)^2 .$$

Expanding the square we have

$$\frac{(u.v)'}{uv} = \frac{(uv)''}{uv} - \left(\frac{u'}{u}\right)^2 - 2\left(\frac{u'}{u}\right)\left(\frac{v'}{v}\right) - \left(\frac{v'}{v}\right)^2 . \qquad \text{(d)}$$

Equating (c) and (d) gives

$$\frac{(uv)''}{uv} - \left(\frac{u'}{u}\right)^2 - 2\left(\frac{u'}{u}\right)\left(\frac{v'}{v}\right) - \left(\frac{v'}{v}\right)^2 = \frac{u''}{u} + \frac{v''}{v} - \left(\frac{u'}{u}\right)^2 - \left(\frac{v'}{v}\right)^2 ,$$

which simplifes to

$$\frac{(uv)''}{uv} = \frac{u''}{u} + \frac{v''}{v} + 2\left(\frac{u'}{u}\right)\left(\frac{v'}{v}\right).$$

The same approach can be used to find the second derivative formula for the quotient rule, this being

$$\frac{(u/v)''}{u/v} = \frac{u''}{u} - \frac{v''}{v} - 2\frac{(u/v)'}{u/v}.\frac{v'}{v} .$$

Examples

Returing to the quotient rules for functions composed only of two other functions we will now go through some examples of all three versions of the quotient rule, namely those of section 1.5.1, 1.5.4 and this section.

1) When $f(x) = \sin x \,/\, \cos^2 x$ we have

a) by the standard quotient rule of section 1.5.1

$$\left(\frac{\sin x}{\cos^2 x}\right)' = \frac{(\cos^2 x)(\cos x) - (\sin x)(-2\cos x \sin x)}{\cos^4 x},$$

b) by the extended quotient rule of section 1.5.4

$$\left(\frac{\sin x}{\cos^2 x}\right)' = \frac{(\cos x)(\cos x) - 2(\sin x)(-\sin x)}{\cos^3 x},$$

c) by the alternative quotient rule of this section

$$\left(\frac{\sin x}{\cos^2 x}\right)' = \frac{\sin x}{\cos^2 x}\left(\frac{\cos x}{\sin x} - \frac{2\cos x(-\sin x)}{\cos^2 x}\right).$$

all of which simplify to $f'(x) = (1 + \sin^2 x)/\cos^3 x$.

2) When $f(x) = (3-x)/\sqrt[3]{1-x^2}$ we have

a) by the usual quotient rule of section 1.5.1

$$\left(\frac{3-x}{\sqrt[3]{1-x^2}}\right)' = \frac{-(1-x^2)^{\frac{1}{3}} + \frac{2x}{3}(1-x^2)^{-\frac{2}{3}}}{(1-x^2)^3},$$

b) by the extended quotient rule of section 1.5.4

$$\left(\frac{3-x}{\sqrt[3]{1-x^2}}\right)' = \frac{-(1-x^2) + \frac{2x}{3}(3-x)}{(1-x^2)^{4/3}},$$

c) by the alternative quotient rule of this section

$$\left(\frac{3-x}{\sqrt[3]{1-x^2}}\right)' = \frac{3-x}{\sqrt[3]{1-x^2}}\left(\frac{-1}{3-x} + \frac{\frac{2x}{3}(1-x^2)^{-\frac{2}{3}}}{\sqrt[3]{1-x^2}}\right).$$

all of which simplify to $f'(x) = \frac{1}{3}(x^2+6x-3)/(1-x^2)^{4/3}$.

We can now note the following

- by far the simplest method in terms of differentiation and algebra is b);
- version c) is also very simple in terms of the process of differentiation but requires a lot of algebra in order to simplify to the final result;
- version a) is relatively the least simple in terms of differentiation and requires more algebra than version b) but less than version c) when it comes to simplifying to the final result;

So, for rational polynomials where the denominator is not a linear power, the extended quotient rule of section 1.5.4 is by far the easiset and simplest. For rational polynomials where the denominator is a linear power extended quotient rule and the standard quotient rule require the same amount of effort.

For any function the alternative quotient rule of this section will also be the simplest in terms of differentiation but will almost always require the most effort in terms of algebra needed to simplify to the final result.

1.9 More difficult examples

Here we will go through some examples of differentiating functions which require the use of a combination of the basic power, product, and quotient rules along with implicit differentiation.

1.9.1 Example 1

Problem: Find the first derivative of $y = \ln(x^2+1)(x^3-1)$

Solution: We could use the product rule as part of the process of finding dy/dx, but we can make our lives easier if we use the property of logs, i.e.

$$y = \ln(x^2+1) + \ln(x^3-1)\ .$$

Now we can use the chain rule spearately on each log term. Thinking as "(derivative of inside) ×(derivative of outside)" for each log term we get

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

The same idea applies if we want to find the first derivative of $y = \ln x^4/(3x-4)^2$. Here we would use log rules to rewrite this function as

$$y = \ln x^4 - \ln(3x-4)^2\ .$$

Differentiation then becomes easier, both requiring only the chain rule. For the first log term let $z = \ln x^4$. Then let $u = x^4$ so that $z = \ln u$. For the second log term let $w = \ln(3x-4)^2$. Then let $v = (3x-4)^2$ which is itself a function of a function for which we would have to use the chain rule.

However, things can be evern easier if we continue using log rules since y can be reduced further to

$$y = 4\ln x - 2\ln(3x-4)\ ,$$

and the chain rule only needs to be used once with u = 3x – 4. Ultimately we then get

$$\frac{dy}{dx} = \frac{4}{x} - 2\frac{3}{3x-4} = \frac{6x-16}{x(3x-4)}\,.$$

1.9.2 Example 2

Problem: Find the first derivative of $y = x^2 3^x$.

Solution: We could again use the product rule on this function, but given we have an exponential terms we would have to use logs at some point. So we might as well use logs from the start. So, from

$$y = x^2 3^x$$

we use the rules of logs to get

$$\ln y = \ln x^2 + \ln 3^x \,.$$

More rules of logs gives

$$\ln y = 2 \ln x + x. \ln 3 \,.$$

We can now differentiate as usual:

$$\begin{aligned} \frac{1}{y} \cdot \frac{dy}{dx} &= \frac{2}{x} + \ln 3 \,, \\ \frac{dy}{dx} &= y.\frac{2}{x} + y. \ln 3 \,, \\ &= 2x. 3^x \\ &\quad + x^2 3^x \ln 3 \,. \end{aligned}$$

1.9.3 Example 3

Problem: Find the first derivative of $f(x) = (e^{2x} - e^{-2x})/(e^{2x} + e^{-2x})$.

Solution: We could again use the logs to make this easier, but for practice will will use the quotient rule, during which we will need to use the chain rule (for which I will again use the idea of "derivative of inside × derivative of outside"):

$$\begin{aligned} f(x) &= \frac{e^{2x} - e^{-2x}}{e^{2x} + e^{-2x}} \,, \\ \frac{df}{dx} &= \frac{(e^{2x} + e^{-2x})(e^{2x} - e^{-2x})' - (e^{2x} - e^{-2x})(e^{2x} + e^{-2x})'}{(e^{2x} + e^{-2x})^2} \,, \\ &= \frac{(e^{2x} + e^{-2x})(2e^{2x} + 2e^{-2x}) - (e^{2x} - e^{-2x})(2e^{2x} - 2e^{-2x})}{(e^{2x} + e^{-2x})^2} \,, \\ &= \frac{2e^{4x} + 4 + 2e^{-4x} - (2e^{4x} - 4 + 2e^{-4x})}{(e^{2x} + e^{-2x})^2} \,, \end{aligned}$$

$$= \frac{8}{(e^{2x} + e^{-2x})^2}.$$

1.9.4 Example 4

Find the second derivative of $f(x) = e^{-2x}\sin 3x$. Again logs could be used to break this into an addition, therefore obviating the need to use the product rule. But as practice we will go through using the product rule.

So

$$\begin{aligned} f(x) &= e^{-2x}\sin 3x\,, \\ \frac{df}{dx} &= \sin 3x.(e^{-2x})' + e^{-2x}(\sin 3x)'\,, \\ &= -2\sin 3x\,.\,e^{-2x} + 3e^{-2x}\cos 3x\,. \end{aligned}$$

We now repeat the product rule on df/dx:

$$\begin{aligned} \frac{d^2f}{dx^2} &= -2\sin 3x\,(e^{-2x})' + e^{-2x}(-2\sin 3x)' \\ &\qquad + \cos 3x\,(3e^{-2x})' + 3e^{-2x}(\cos 3x)'\,, \\ \frac{d^2f}{dx^2} &= 4\sin 3x\,.\,e^{-2x} - 6e^{-2x}\cos 3x \\ &\qquad -6\cos 3x\,.\,e^{-2x} - 9e^{-2x}\sin 3x\,, \\ &= 4\sin 3x\,.\,e^{-2x} - 12e^{-2x}\cos 3x - 9e^{-2x}\sin 3x\,. \end{aligned}$$

1.9.5 Example 5

Problem: Find the first derivative of $y = \ln(\sec x + \tan x)$.

Solution: Let $u = \sec x + \tan x$. Then $y = \ln u$, and we use the chain rule:

$$\begin{aligned} \frac{d}{dx}(\ln(\sec x + \tan x)) &= \frac{dy}{du}.\frac{du}{dx}\,, \\ &= \frac{1}{u}.\frac{d}{dx}(\sec x + \tan x)\,, \\ &= \frac{1}{u}.\left[\frac{d}{dx}\left(\frac{1}{\cos x}\right) + \frac{d}{dx}(\tan x)\right]. \end{aligned}$$

By the quotient rule the derivative of $1/\cos(x)$ is $\sec(x).\tan(x)$ and that of $\tan(x)$ is $\sec^2(x)$. Hence

$$\begin{aligned}\frac{d}{dx}(\ln(\sec x + \tan x)) &= \frac{1}{u}.(\sec x \tan x + \sec^2 x)\,, \\ &= \frac{\sec x \tan x + \sec^2 x}{\sec x + \tan x}\,, \\ &= \sec x\,.\end{aligned}$$

See https://chris-fenwick.jimdofree.com/maths-teaching/my-maths-examples/ for more examples on this topic.