# On Differentiation I

## On the definition of the derivative and the proof of some basic derivatives

**By Chris Fenwick**

**Uploaded on 06/02/2024**

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

# Table of Content

## 1.1 Introduction

In these notes we will go through the concept and algebra of the derivative. These notes develop the concept and mathematics of differentiation from scratch, and assume no prior knowledge to, or understanding of, differentiation. We will do this as follows:

1. We will look at how the idea of finding the slope of a straight line can be used as a starting point for finding the slope of a curve. Using this idea we will see that we will have to continually adjust the way we find the slope of a curve, and this will be done by the process of travelling along the curve;
2. We will see that, in applying this process, the true slope we end up calculating is valid only at one point on the curve, not for the whole of the curve. Such a situation will then give rise to the new concepts of tangents and derivatives;
3. We will then see that the direction in which we travel along the curve is important if we are to accept the slope of the curve to be what is;
4. We will then discuss more deeply the idea of "travelling along the curve", an important and fundamental concept to calculus which gives rise to the concept of "limits", and which we will then make more rigorous via algebra;
5. We will look at two ways in which the derivative of a function can be understood and represented as a function. Such an understanding will allow us more easily to find the slope of the curve at any point, anywhere along the curve, and not just for one point on the function. Another way of saying this is that this function will represent the derivative of the $f(x)$ as a whole, and not tjust the derivative at one point of $f(x)$.
6. We will then formally define the first derivative;
7. We will also look at how the derivative, apart from being a measure of slope or a rate of change, can also be seen as a transformation or as a measure of sensitivity to change;
8. We will see a few example of equations where the first derivative is used to represent and describe natural phenomena;
9. We will go through proving the first derivative of basic functions, sometimes in two different ways. Along the way we will go though detailed explanation (with numerical work and diagrams) about the basic effect of dividing by numbers which forever get closer to 0;

10. We will discuss the idea that not all functions have derivatives. In other words there are functions for which we cannot find the derivative/slope at certain points, and there are functions which do not have a derivative at all;

11. Finally we will go through some studies on derivative and tangents. These studies aim to analyse the idea of derivatives and tangents in a non-standard way. The point of including these studies is to learn more broadly about differentiation and tangents;

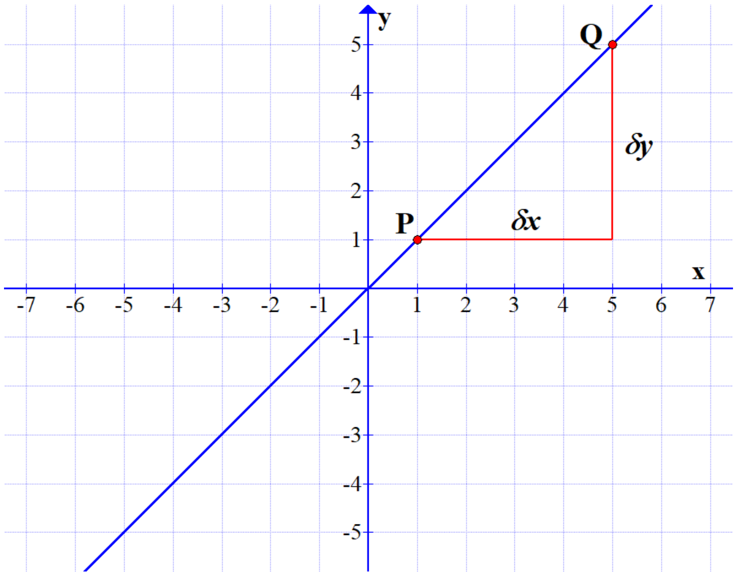
## 1.2 The derivative as the slope of a curve at a point

In this section we will start the process of understanding how to find the slope of a curve. We will first start by reviewing how we find the slope of a straight line, and then we will see how this idea can be modified so that we can apply it to curves.

### *1.2.1 Approaching the derivative from one direction*

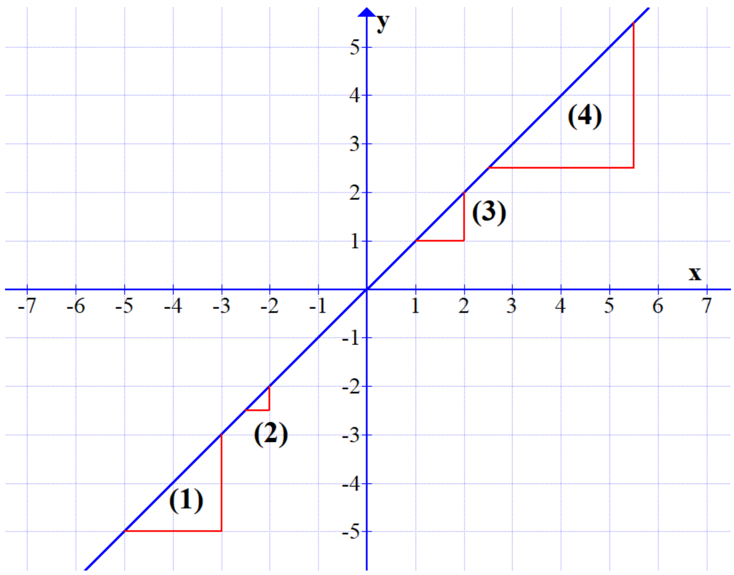
Suppose we want to know how steep a straight line is. To do this we simply calculate how high up we have gone in the vertical direction from a starting point P to an end point Q (sometmes called the rise) and divide this distance by how far across we have gone in the horizontal direction from P to Q (sometimes called the run). The distances travelled in both cases is the difference between our end point and starting point, and the steepness calculation is sometime called "rise over run".

As an example consider the function $y = x$. The "rise" and "run" are represented in the diagram below by the vertical distance $\delta y$ and the horizontal distance $\delta x$. We see that if we travel across from $x = 1$ to $x = 5$ then $\delta x = 5 - 1 = 4$. The corresponding vertical distance travelled $y = 1$ to $y = 5$ gives $\delta y = 5 - 1 = 4$. Hence the line has a steepness of $\delta y/\delta x = 4/4 = 1$ unit.




| Slope |
|---|
| $\frac{\delta y}{\delta x} = \frac{5-1}{5-1} = 1$ |

Note that it doesn't matter where P or Q are located, or how big and small our resulting triangle is. The slope of the line *y* = *x* will always be 1 unit. This is illustrated in the graph and table below.

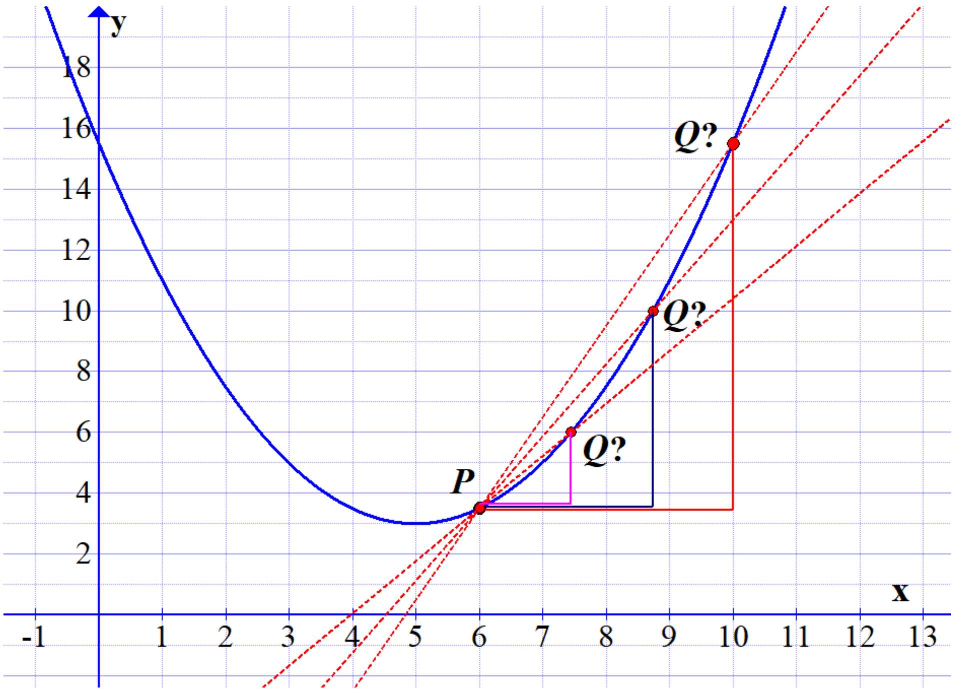



| Triangle | Slope |
|---|---|
| (1) | $\frac{\delta y}{\delta x} = \frac{-3-(-5)}{-3-(-5)} = 1$ |
| (2) | $\frac{\delta y}{\delta x} = \frac{-2-(-2.5)}{-2-(-2.5)} = 1$ |
| (3) | $\frac{\delta y}{\delta x} = \frac{2-1}{2-1} = 1$ |
| (4) | $\frac{\delta y}{\delta x} = \frac{5.5-2.5}{5.5-2.5} = 1$ |

Another way of interpreting the concept of rise-over-run is that, in calculating $\delta y \div \delta x$, we are finding the rate at which the position of *y* changes given a change in the position of *x*. Terms like slope, gradient, and steepness are all used to mean the same thing: rate of change of *y* with respect to *x*.

However, the calculation of $\delta y \div \delta x$ above only applies if the path we travel from P to Q is a straight line. But what if this is not the case? What if our path is curved? How then do we find the slope of a function?

For example, consider the curve below with various different locations for Q. The true path from P to Q is curved so our first problem is that we cannot use our standard calculation of vertical distance divided by horizontal distance to find the slope of the curve PQ. If we did so we would be finding the slope of the straight line path PQ.

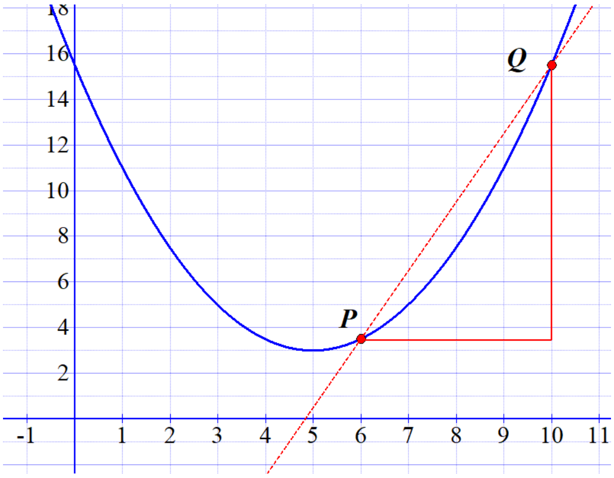



Our second problem is about which line PQ we are finding the slope of. In the diagram above there are three lines PQ, and none of their slopes are either equal to each other or equal to the actual slope of the curve. Worse still there are an infinite number of places we can put point Q, so we end up with an infinite number of slopes PQ all of them different and all of them wrong in terms of measuring the slope of the curve.

So the question is, If we want to use $\delta y \div \delta x$ as our calculation for finding the slope of the curve where should Q be placed? Is it even possible to find a place for Q on the curve such that PQ is a straight line but that we are still measuring the slope of the curve and not the line PQ?

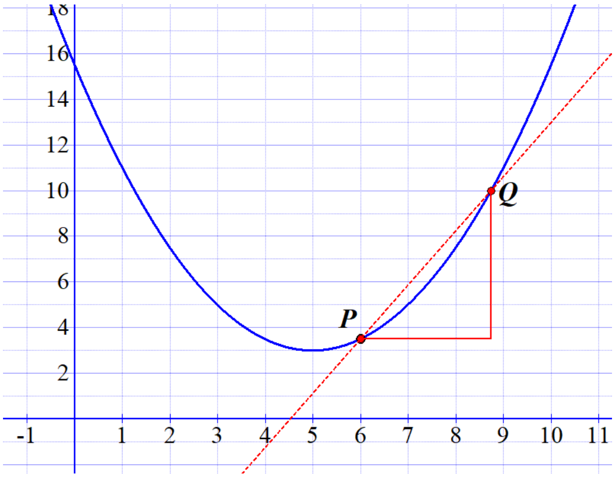
Well, although the second question seems like a contradiction the answer to it is yes, and we will see how this is so as we go through this section. The answer to the first question can be addressed as follows: The function plotted is $y = 0.5(x-5)^2 + 3$.

Let us look at the values of the three slopes PQ in the graph above when P is the point (6, 3.5) and compare these values with the real value of the slope of the curve (which, for the moment, we will assume we know).



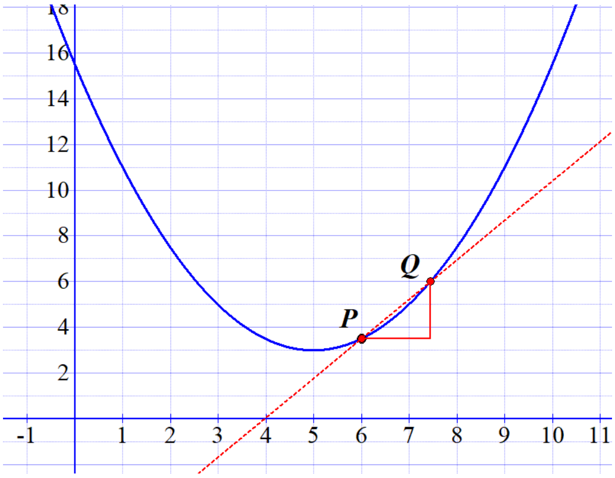

| *P* | *Q* | *δx* | *δy* | *δy ÷ δx* |
|---|---|---|---|---|
| (6, 3.5) | (10, 15.5) | 4 | 12 | 3 |

Real value of slope at P is 1




| *P* | *Q* | *δx* | *δy* | *δy ÷ δx* |
|---|---|---|---|---|
| (6, 3.5) | (8.74, 10) | 2.74 | 7.5 | 2.737 |

Real value of slope at P is 1




| *P* | *Q* | *δx* | *δy* | *δy ÷ δx* |
|---|---|---|---|---|
| (6, 3.5) | (7.75, 6) | 1.25 | 2.5 | 2 |

Real value of slope at P is 1

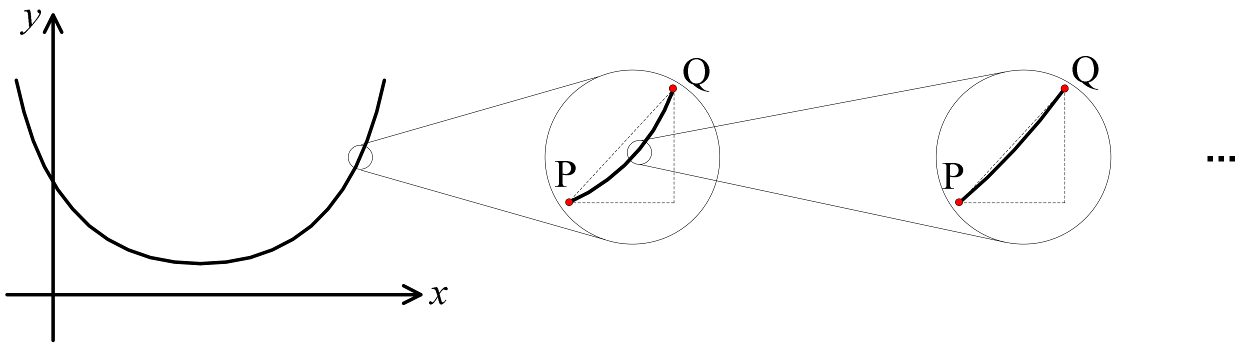
So there seems to be a pattern here: the closer Q gets to P the closer $\delta y \div \delta x$ gets to the real value of the slope at P. This can be seen even more to be the case if we take values of Q eve more close to P:

| **P** | | **Q** | | **$\delta x$** | **$\delta y$** | **$\delta y \div \delta x$** |
|---|---|---|---|---|---|---|
| x | y | x | y | | | |
| 6 | 3.5 | 6.8 | 4.62 | 0.8 | 1.12 | 1.40000 |
| 6 | 3.5 | 6.7 | 4.445 | 0.7 | 0.945 | 1.35000 |
| 6 | 3.5 | 6.6 | 4.28 | 0.6 | 0.78 | 1.30000 |
| 6 | 3.5 | 6.5 | 4.125 | 0.5 | 0.625 | 1.25000 |
| 6 | 3.5 | 6.4 | 3.98 | 0.4 | 0.48 | 1.20000 |
| 6 | 3.5 | 6.3 | 3.845 | 0.3 | 0.345 | 1.15000 |
| 6 | 3.5 | 6.2 | 3.72 | 0.2 | 0.22 | 1.10000 |
| 6 | 3.5 | 6.1 | 3.6050 | 0.1 | 0.105 | 1.05000 |
| 6 | 3.5 | 6.01 | 3.51005 | 0.01 | 0.01005 | 1.00500 |
| 6 | 3.5 | 6.001 | 3.5010005 | 0.001 | 0.001001 | 1.00050 |
| 6 | 3.5 | 6.00010 | 3.500100005 | 0.0001 | 0.000100005 | 1.00005 |
| … | … | … | … | … | … | … |

*Table of slopes for the secant PQ when Q gets closer to P.*

As we go to 9, 10, 11, 12,… decimal places so $\delta y \div \delta x$ approaches even more closely the value 1. Since the line joining P and Q is a straight line, and since $\delta y$ and $\delta x$ are still actual numbers (however small these may be), we are still justified in doing the calculation $\delta y \div \delta x$ as a way of measuring the slope of the curve. However, whereas in measuring the slope of the straight line $y = x$ (or any straight line $y = mx + c$) our slope could be measured over an interval of any length along the $x$-axis, it so happens that our slope is now being measured over smaller and smaller intervals. This means that the slope is being measured closer and closer to the point P. We will see later that we will end up with the slope being measured at point P itself.

Returing to the fact that we can still use $\delta y \div \delta x$ as a way of measuring the slope of the curve we see that when Q contnually approaches P we are zooming into a very local, micro, even nano, part of the curve. This zooming-in effect can generally be illustrated visually as shown below:




In this situation we see that PQ is still a straight line and the zoomed-in part of the curve from P to Q now looks more and more like a straight line. There is, however, one caveat which applies in this situation: however much we zoom in to a part of the curve, Q will never land on top of P, and the curved part of PQ will never theoretically become a straight line. All that the zooming-in process will show us is that Q gets ever closer to P, and in the process the curve between Q and P becomes more and more straight.

The above comment is such an important concept that I will repeat it: Although the function is still curved the fact that it is straightening out at the micro or nano level means that the straight line secant from P to Q is a very close fit to this part of the curve. In theoretical terms the function will always remain curved however much we zoom in, but the secant line will represent the path of the curve better and better. Despite the fact that the secant PQ will never exactly line up with the curve PQ the fact that Q is forever approaching P allows us to more validly use the calculation $\delta y \div \delta x$ as a way of measuring the slope of the curve. This "forever zooming-in" concept is **the** fundamental concept of calculus. Without it there is no calculus. See if you can imagine this zooming in process continuing for ever without Q ever landing on top of P.

As we continue this process of zoom-in, although P and Q remain theoretically separate we are actually no longer finding the slope of the secant PQ. Instead we are finding *the slope of the curve at the point P itself*. What this means is that we are finding the rate at which the curve is changing at one single point P, not over an interval PQ. Or, to put it another way, we are finding the rate of change of the curve at the instantaneous moment we arrive at P.

We now come to the theoretical part of this topic, the part where the concept of $\delta y \div \delta x$ as simply the ratio of vertical distance to horizontal distance changes to the concept $dy/dx$, a brand new object called the "derivative of $y(x)$ with respect to $x$ at the point P".

The theoretical "thing" which changes $\delta y \div \delta x$ to $dy/dx$ is called "limit". So instead of saying "as Q gets closer and closer and closer and … to P" we say

"In the limit as Q approaches P".

Since we are using $\delta x$ as our notation for horizontal distance, and since $\delta x$ approaches 0 as Q approaches P, we more precisely say

"In the limit as $\delta x$ approaches 0",

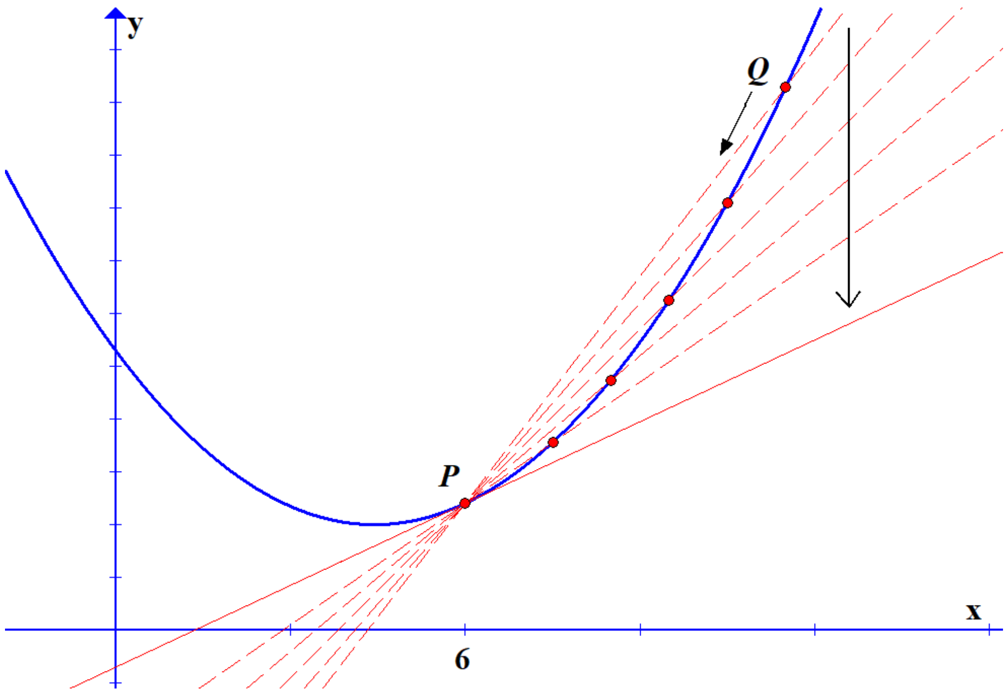
This statement is represented symbolically as

$$\lim_{\delta x \to 0}$$

where we then specify what it is we are taking the limit of. In this case we have

$$\lim_{\delta x \to 0} \frac{\delta y}{\delta x} .$$

So the secant that went through the two points P and Q on the curve now becomes something called a *tangent*, this tangent touching the curve at one point only (namely point P). This is illustrated in the graph below:




Visually we see that, as $\delta x$ approaches 0, the sequence of secants (the dashed lines in the graph above) converges to the tangent at P (seen as the solid line in the graph above), this tangent being the slope or rate of change of the curve of $y = f(x)$ at point P.

Furthermore, there is nothing special about having chosen the function $y = 0.5(x-5)^2 + 3$ or the point $x$ = 6. The concept of a secant sliding up or down along the curve of a function towards any fixed point P($x$, $y$) applies to any function and any point P($x$, $y$) on the curve of that function. In general such an approach will give a tangent at P($x$, $y$) (there are exceptions to this which we shall get to later in section 1.12).

<u>*Example*</u>

Let us say that we want to find the value of the slope of the function $f(x) = x^2$ at the point (1, 1). Let us now choose a second point on the right of (1, 1), say (2, 4). What value will we get for the slope of $f(x)$ when this second point approaches (1, 1)? To answer this we can set up a table of slopes as we did previously:

| $P$ | $Q$ | $\delta x$ | $f(1)$ | $f(1+\delta x)$ | $f(1+\delta x) - f(1)$ | $\frac{f(1+\delta x)-f(1)}{\delta x}$ |
|---|---|---|---|---|---|---|
| (1, 1) | (2, 4) | 1 | 1 | 4 | 3 | 3 |
| (1, 1) | (1.5, 2.25) | 0.5 | 1 | 2.25 | 1.25 | 2.5 |
| (1, 1) | (1.1, 1.21) | 0.1 | 1 | 1.21 | 0.21 | 2.1 |
| (1, 1) | (1.01, 1.0201) | 0.01 | 1 | 1.0201 | 0.0201 | 2.01 |
| (1, 1) | (1.001, 1.002001) | 0.001 | 1 | 1.002001 | 0.002001 | 2.001 |
| … | … | … | … | … | … | … |
| (1, 1) | Infinitely close to (1, 1) | Infinitely close to 0 | 1 | Infinitely close to 1 | Infinitely close to 0 | 2 (limit as $\delta x \to 0$) |

So as we approach infinitely close to the point (1, 1) the slope at point P approaches 2. The technical way of saying this is as follows:

in the limit as $\delta x \to 0$, the gradient at P = 2.

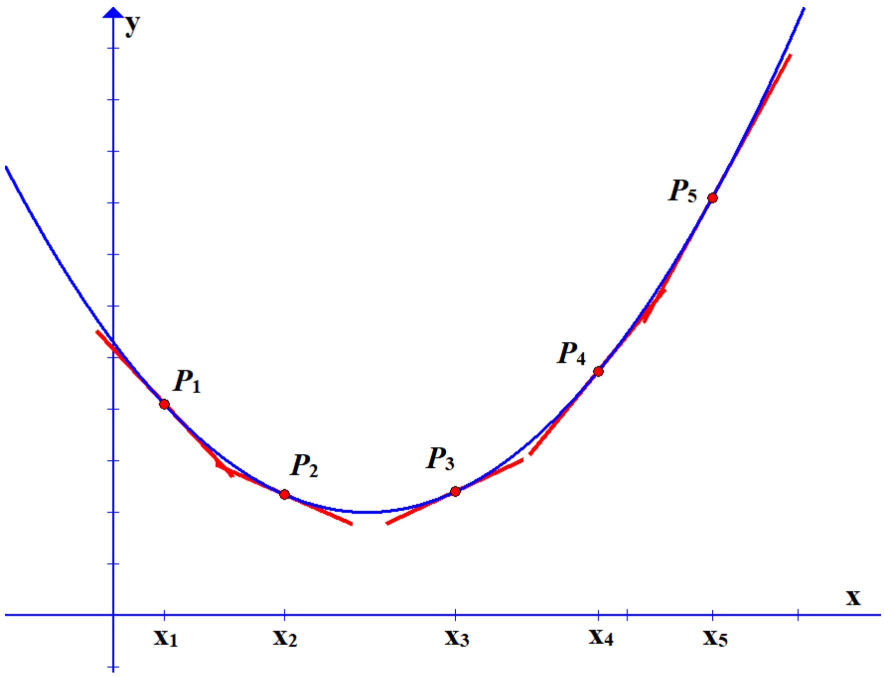
*1.2.2 A comment about tangents*

I have seen a number of animations which demonstrate the effect of Q approaching P to form a tangent at a given point P. These animations then go on to show the tangent sliding along the curve. This sliding of the tangetn is at best totally misleading and at worst false in what it is illustrating about the tangent.

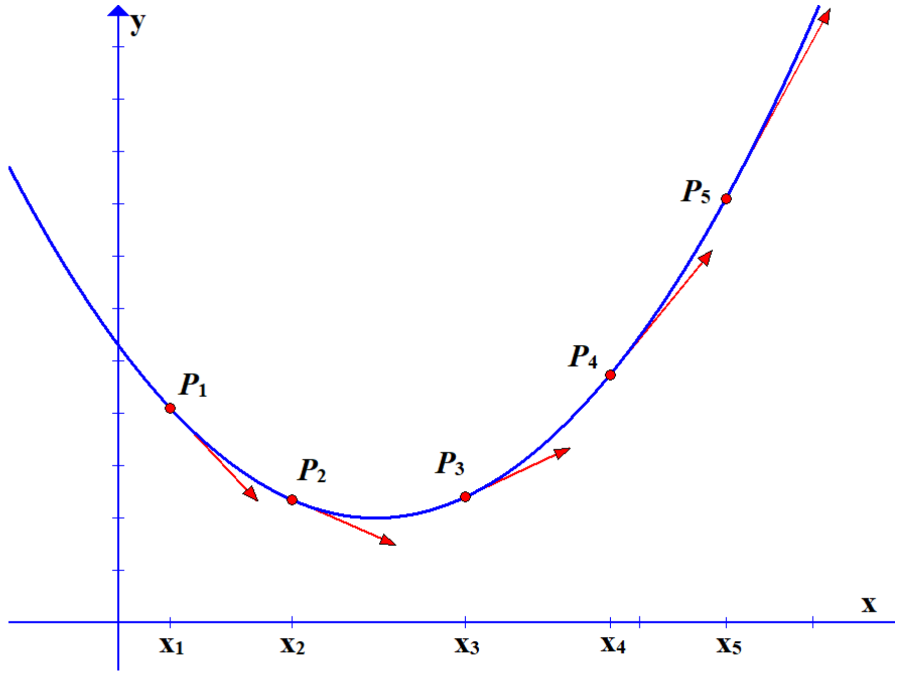
The fact is that the tangent at a point is static. The tangent located at $P_1(x_1, y_1)$ does not move. There is no way we can make a tangent $P_1$ located at $(x_1, y_1)$ move to another point $P_2$ located at $(x_2, y_2)$ since any tangent is unique to one and only one point on the curve. The tangents at $P_1$ and $P_2$ are two separate and distinct tangents. So there is no such thing as a tangent sliding along the curve from $P_1$ so that it arrives at $P_2$ or any other location on the curve.

The idea of sliding one of the points of a secant towards the other (fixed) point is simply a heuristic we used to explore the effects and consequences of this movement on $\delta y \div \delta x$. One might say that the tangent is the final resting place of the sliding secant.

This implies that every point $(x, y)$ on the curve of a function $y(x)$ will have its own unique tangent at a given point, as illustrated below. Each point $P_1$ to $P_5$ has its own tangent represented by a line touching the curve at their respective red dots.




Furthermore, just as the tangent represents the slope or rate of change of the curve of $y = f(x)$ at a specific point $x$ it also represents the direction in which the curve $f(x)$ is going at that specific point, as illustrated below by the arrowed lines:




For example, if the curve represented the path of a planet orbiting the Sun, then if gravity were to be immediately turned off when the planet was at position $P_1$ the planet would continue travelling in the straight line shown by the arrow eminating from $P_1$. The same would be true if gravity were turned off at points $P_2$, $P_3$, $P_4$, $P_5$.

So at any given point on the curve, the curve is travelling in a direction different to that at any other point on the curve. In fact, the curve is continually traveling in a different direction; it is continually changing its direction of travel. So when the curve is at $P_1$ its straight line direction is as indicated by the red arrowed tangent at $P_1$; when the curve is at $P_2$ its straight line direction is as indicated by the red arrowed tangent at $P_2$, etc.

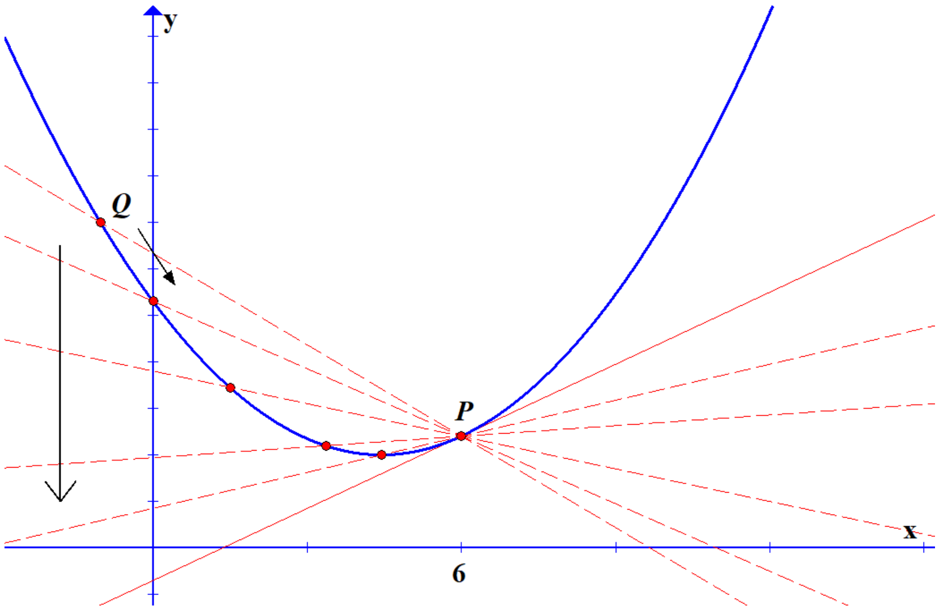
Also, since $\delta y$ and $\delta x$ shrink to an infinitely small distance the tangent itself isn't really a line. Or if it is, it is a line of infinitely small length. It is only for visual reasons that we draw tangents as lines.

*1.2.3 Approaching the derivative from the other direction*

In our previous analysis of measuring the slope of a function at point P, point Q approached point P from the right hand side. However, one thing we will need if this idea of slope-at-a-point is to be valid is to get the same slope if Q approaches P from the left hand side. Then we will have a consistent technique which will work whatever direction of travel we take along the curve $f(x)$.

If we get a different result to the slope at P when Q approaches P from the left hand side then our answer will be inconsistent with our previous answer to the slope at P, and this will make the "slope-at-a-point" idea useless (it may not be obvious but there are functions whereby the slope we get when travelling to P from the left hand side is different to the slope we get when we get if we travel to P from the right hand side).

So, given the situation depicted below let us see what happens to the sequence of secants PQ (dashed lines) when point Q approaches point P from the left hand side:




Visually we see that, as Q gets closer and closer, getting forever closer, to P from the left hand side we end up with the same tangent we got when Q approached P from the right hand side. Phew! What this means is that we will get one and the same result for the slope of the curve at P irrespective of which direction we approach P. We will confirm this numerically when we go through an actual example in the next section. Ultimately, this has to be proved mathematically, and this is done at university level if-and-when we study a topic called Real Analysis.

*Example*

Continuing the example at the end of section 1.2.1 in our study of finding the slope of the function $f(x) = x^2$ at the point (1, 1) the question now is, Will we get the same slope if we approach the point (1, 1) from the left hand side as we did when we approached that point from the right hand side? To test this let us choose our left hand side point to be (0, 0). The answer is in the table below:

| $Q$ | $P$ | $\delta x$ | $f(1-\ \delta x)$ | $f(1)$ | $f(1) - f(1-\delta x)$ | $\dfrac{f(1)-f(1-\delta x)}{\delta x}$ |
|---|---|---|---|---|---|---|
| (0, 0) | (1, 1) | 1 | 0 | 1 | 1 | 1 |
| (0.5, 0.25) | (1, 1) | 0.5 | 0.25 | 1 | 0.75 | 1.5 |
| (0.9, 0.81) | (1, 1) | 0.1 | 0.81 | 1 | 0.19 | 1.9 |
| (0.99, 0.9801) | (1, 1) | 0.01 | 0.9801 | 1 | 0.0199 | 1.99 |
| (0.999, 0.998001) | (1, 1) | 0.001 | 0.998001 | 1 | 0.001999 | 1.999 |
| … | … | … | … | … | … | … |
| Infinitely close to (1, 1) | (1, 1) | Infinitely close to 0 | Infinitely close to 1 | 1 | Infinitely close to 0 | 2 (limit as $\delta x \to 0$) |

So we see that as Q approaches P from the left hand side the slope of the curve is also 2. Since the answer to the slope of the curve at (1, 1), in the limit as $\delta x \to 0$, is the same from both the left hand side and the right hand side of (1, 1) we can now confirm that the slope of the curve $f(x) = x^2$ at the point (1, 1) is indeed 2.

*Examples*

1) Let us find the slope of $f(x) = 4x^2 + 1$ at $x = 2.5$. to do this we will set up a table of values as $\delta x$ approaches 2.5 from the left hand side and from the right hand side. Again, we want the result we get when we approach from the right hand side to be the same as the result we get when we approach from the left hand side.

   So, approaching $x = 2.5$ from the left hand side we have

| $Q$ | $P$ | $\delta x$ | $f(2.5-\delta x)$ | $f(2.5)$ | $f(2.5) - f(2.5-\delta x)$ | $\dfrac{f(2.5)-f(2.5-\delta x)}{\delta x}$ |
|---|---|---|---|---|---|---|
| (1.5, 10) | (2.5, 26) | 1 | 10 | 26 | 16 | 16 |
| (2, 17) | (2.5, 26) | 0.5 | 17 | 26 | 9 | 18 |

| | | | | | | |
|---|---|---|---|---|---|---|
| (2.4, 24.04) | (2.5, 26) | 0.1 | 24.04 | 26 | 1.96 | 19.6 |
| (2.49, 25.8004) | (2.5, 26) | 0.01 | 25.8004 | 26 | 0.1996 | 19.96 |
| (2.499, 25.980004) | (2.5, 26) | 0.001 | 25.980004 | 26 | 0.019996 | 19.996 |
| … | … | … | … | … | … | … |
| Infinitely close to (2.5, 26) | (2.5, 26) | Infinitely close to 0 | Infinitely close to 26 | 26 | Infinitely close to 0 | 20 (limit as $\delta x \to 0$ from the left hand side of x = 2.5) |

Approaching $x$ = 2.5 from the right hand side we have

| $P$ | $Q$ | $\delta x$ | $f(2.5)$ | $f(2.5+\delta x)$ | $f(2.5+\delta x) - f(2.5)$ | $\frac{f(2.5+\delta x)-f(2.5)}{\delta x}$ |
|---|---|---|---|---|---|---|
| (2.5, 26) | (3.5, 50) | 1 | 26 | 50 | 24 | 24 |
| (2.5, 26) | (3, 37) | 0.5 | 26 | 37 | 11 | 22 |
| (2.5, 26) | (2.6, 28.04) | 0.1 | 26 | 28.04 | 2.04 | 20.4 |
| (2.5, 26) | (2.51, 26.2004) | 0.01 | 26 | 26.2004 | 0.2004 | 20.04 |
| (2.5, 26) | (2.501, 26.020004) | 0.001 | 26 | 26.020004 | 0.020004 | 20.004 |
| … | … | … | … | … | … | … |
| (2.5, 26) | Infinitely close to (1, 1) | Infinitely close to 0 | 26 | Infinitely close to 26 | Infinitely close to 0 | 20 (limit as $\delta x \to 0$ from the right hand side) |

Since the slope of the secant on the left hand side and right hand side of x = 2.5 both approach a value of 20 (the last column of both tables above) we can now confirm that the slope of the curve $f(x) = 4x^2 + 1$ at the point (2.5, 26) is indeed 20.

There are many other limits for which we get a numerical result to a ratio when $\delta x$ approaches some specific value (not always 0) or when $x$ approaches infinity, some of which are shown in the appendix to these notes. However, it is illustrative to see two very important examples of thji limiting preocess. The first examples relates to the function $(\sin \delta x)/\delta x$ where the limit of this function (as $\delta x$ approaches 0) happens to be 1, and the second example relates to $(1 + 1/x)^x$ where the limit of this function as $x$ approaches infinity is 2.7818281…

$$\lim_{\delta x \to 0} \frac{\sin \delta x}{\delta x} = 1$$

| $\delta$x | sin($\delta$x) | sin($\delta$x) / $\delta$x |
|---|---|---|
| -0.1 | -0.099833416646828 | 0.998334166468282 |
| -0.01 | -0.009999833334167 | 0.999983333416666 |
| -0.001 | -0.000999999833333 | 0.999999833333342 |
| -0.0001 | -0.000099999999833 | 0.999999998333333 |
| … | … | … |
| 0 | 0 | 1 |
| … | … | … |
| 0.0001 | 0.000099999999833 | 0.999999998333333 |
| 0.001 | 0.000999999833333 | 0.999999833333342 |
| 0.01 | 0.009999833334167 | 0.999983333416666 |
| 0.1 | 0.099833416646828 | 0.998334166468282 |

*Table 1: The limit of $(\sin(\delta x))/\delta x$ as $\delta x$ approaches 0 from the left and right hand side*

$$\lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x = e$$

| x | 1/x | $(1 + 1/x)^x$ |
|---|---|---|
| 1 | 1.000000000000000 | 2.000000000000000 |
| 10 | 0.100000000000000 | 2.593742460100000 |
| 100 | 0.010000000000000 | 2.704813829421530 |
| 1000 | 0.001000000000000 | 2.716923932235520 |
| 10000 | 0.000100000000000 | 2.718145926824360 |
| 100000 | 0.000010000000000 | 2.718268237197530 |
| 1000000 | 0.000001000000000 | 2.718280469156430 |
| 10000000 | 0.000000100000000 | 2.718281693980370 |
| 100000000 | 0.000000010000000 | 2.718281786395800 |
| … | … | … |
| Approaching infinity | Approaching 0 | 2.718281828459045 |

*Table 2: The limit of $(1 + 1/x)^x$ as x approaches infinity*

In the case of the results in table  It should be clear from the above examples that, however small $\delta x$ becomes, both $\delta x$ and $\sin(\delta x)$ are still finite values not equal to 0. This means that we can divide them. It is then because of the relative sizes of these values that the ratio converges to a specific value. However, not all ratios exhibit this property. For example, the function $f(x) = \sin(1/x)$ will not converge to any value as $x \to 0$. All it will do is bounce between +1 and −1. So there are cases for which we get an answer as “Q forever approaches P” and some case for which we cannot get an answer. The mathematical study this forms part of an undergraduate maths degree course.

### *1.2.4 A comment about limits*

Let us return to the concept of a point Q continually approaching point P. So important is this concept that it is worth mentioning again: Q never lands exactly on top of P. In other words, Q is never equal to P. If this were the case then $\delta y \div \delta x$ would give us $0 \div 0$ which is infinity (or, more precisely, not a number). It is simply that Q gets closer and closer to P. In fact, Q can get as close as we like, or arbitrarily close, to P: Q → P, but Q ≠ P.

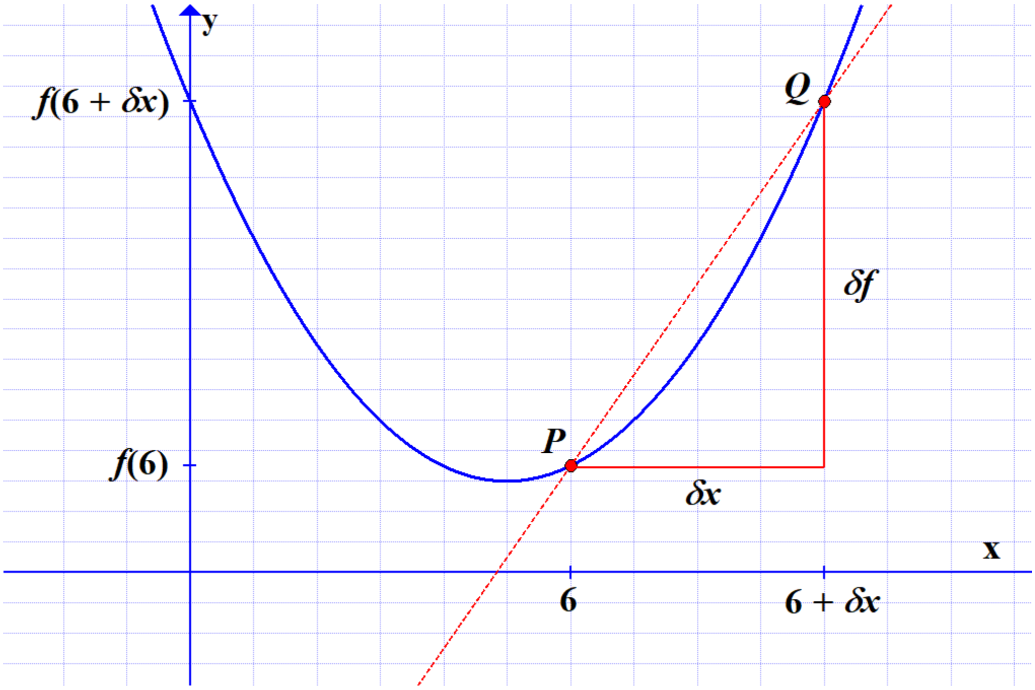
But despite the seeming impossibility of Q forever approaching P but never landing on top of P we still get an answer to the slope of the curve. As we have seen with the numerical work previously done it is possible for $\delta y \div \delta x$ to give an actual numerical answer even when both $\delta x$ and $\delta y$ are extremely (infinitely?) close to zero *but not actually equal to zero*. This is because, even though small, $\delta x$ and $\delta y$ are still finite numbers, and as such they can be divided by each other. Therefore, the result is that the sequence of ratios of $\delta y$ to $\delta x$ we obtain as $\delta x \to 0$ is such that this sequence converges to a fixed value which happens to represent the slope of the curve at a particular point.

### *1.2.5 Becoming more rigorous*

Let us now develop the algebra of $\delta y \div \delta x$ more precisely. This is the beginnings of getting to a formal mathematical definition of the derivative. Therefore, given a function $y = f(x)$ we have the following:

- The point $(6, f(6))$ is our current position on the function (point P in previous graphs);
- The point $(6 + \delta x, f(6 + \delta x)$ is a neighbouring point to the right hand side of $(6, f(6))$ on the function (point Q in previous graphs);

- $f(6+\delta x) - f(6)$ is the vertical distance between $f(6)$ and $f(6+\delta x)$, also called $\delta f$;
- $(6 + \delta x) - 6$ is the horizontal distance between $6 + \delta x$ and 6, and which simplifes to $\delta x$;
- $[f(6+\delta x) - f(6)]/\delta x = \delta f/\delta x$ is the ratio of vertical to horizontal distances. It represents how far up (or down) we go given how far across we have gone;

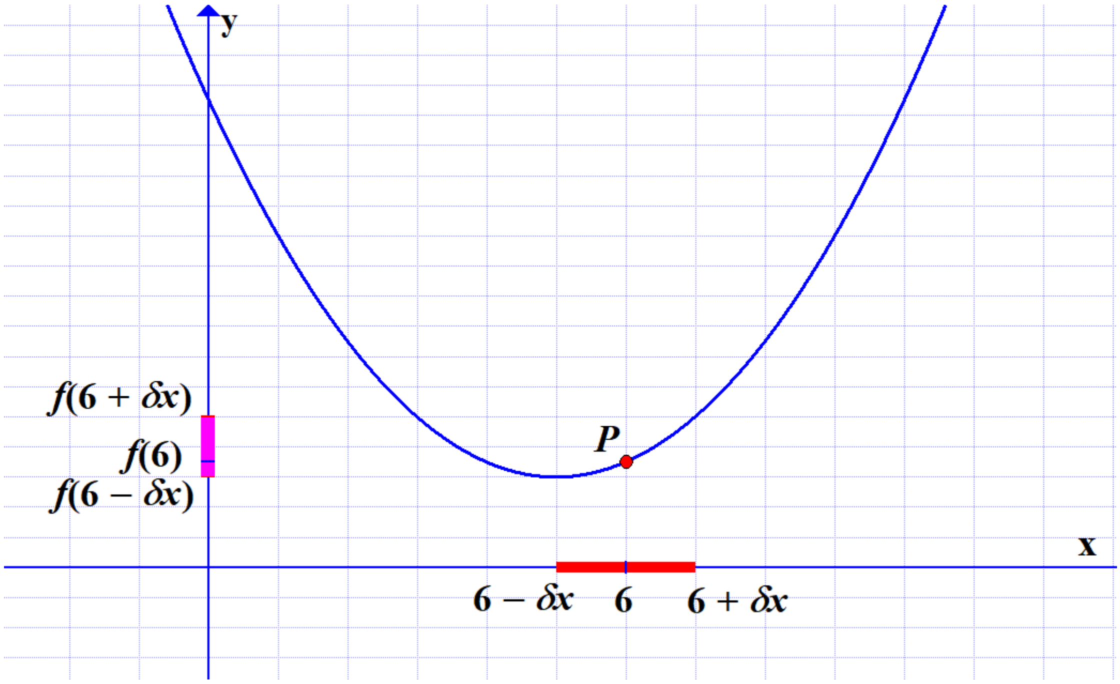



From this we can set up an expression for $\delta f/\delta x$ to be

$$\frac{\delta f}{\delta x} = \frac{f(6+\delta x) - f(6)}{(x+\delta x) - x}$$

The expression on the right hand side is known as the *difference-quotient* and represents the degree to which $f(x)$ changes around the point $f(6)$ due to a small change $\delta x$ around the point $x = 6$.

This is illustrated in the graph below. A wobble or perturbation or nudge by an amount $\delta x$ either side of $x = 6$ (shown by the red bar) will cause a respective wobble, perturbation or nudge around $f(6)$ (shown by the pink bar):




Even though $\delta x$ and $\delta f$ are small distances they are still measurable distances. The slope $\delta f/\delta x$ is therefore still just a division of two separate numbers. But as $\delta x$ approaches 0 this division now becomes a brand new thing, called the *derivative*, something which represents the slope of $f(6)$ at $x$ = 6. What $\delta f/\delta x$ now tell us is the degree to which $f(x)$ changes at the exact moment it arrives at $x$ = 6. This is a new type of slope and is symbolised as $df/dx$:

$$\lim_{\delta x \to 0} \frac{f(6+\delta x) - f(6)}{\delta x} = \frac{df}{dx} \text{ at } x = 6\,.$$

More accurately speaking $df/dx$ represents the *instantaneous rate of change* of $f(6)$ at $x$ = 6.

Conceptually we can break down the above expression as follows:

| Zooming in | the slope of the secant line at the point $x$ = 6 | is | the derivative at $f(6)$ |
|---|---|---|---|
| $\lim_{\delta x \to 0}$ | $\frac{f(6+\delta x) - f(6)}{\delta x}$ | $=$ | $\frac{df}{dx}$ at $x = 6$ |

In general, for any point $x = a$ the expression

$$\lim_{\delta x \to 0} \frac{f(a+\delta x) - f(a)}{\delta x}$$

is for curves what $\delta y/\delta x$ is for straight lines, namely that the former allows us to calculate the slope of a curve at any point $x = a$. All we need to do is to compute the limit expression every time we want to find the slope of a function at a given point (notice that the value of the slope now only applies to a specific point $x = a$, not to the function as a whole as is the case when we are using $\delta y/\delta x$ for straight lines).

*Examples*

1) Let us find the slope of $f(x) = 4x^2 + 1$ at $x = 2.5$. The general expression for the slope of a function at $x = 2.5$ is

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{f(2.5+\delta x) - f(2.5)}{\delta x}.$$

Using the given function we have

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{[4(2.5+\delta x)^2 + 1] - [4(2.5)^2 + 1]}{\delta x}.$$

All we do now is to use algebra to expand the numerator in the expression above:

$$\begin{aligned} \frac{df}{dx} &= \lim_{\delta x \to 0} \frac{[4(6.25 + 5\delta x + (\delta x)^2) + 1] - [4(6.25) + 1]}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{20\delta x + 4(\delta x)^2}{\delta x}, \\ &= \lim_{\delta x \to 0} (20 + 4\delta x). \end{aligned}$$

Evaluating the last limit gives us $df/dx = 20$. This is how steep the curve is at $x = 2.5$. Or we can say that this is the rate at which $f(x)$ is changing the moment we arrive at $x = 2.5$.

2) How steep is the curve of $g(x) = (3+x)/(3-x)$ (where $x \neq 3$) when it reaches $x = 2$? As before the general expression for the slope of a function at $x = 2$ is

$$\frac{dg}{dx} = \lim_{\delta x \to 0} \frac{g(2+\delta x) - g(2)}{\delta x}.$$

Using the given function we have

$$\frac{dg}{dx} = \lim_{\delta x \to 0} \frac{\dfrac{3 + (2 + \delta x)}{3 - (2 + \delta x)} - \dfrac{3 + 2}{3 - 2}}{\delta x}.$$

Doing the algebra and simplifying we get

$$\begin{aligned} \frac{dg}{dx} &= \lim_{\delta x \to 0} \frac{\dfrac{5 + \delta x}{1 - \delta x} - 5}{\delta x}, \\ &= \lim_{\delta x \to 0} \frac{5 + \delta x - 5(1 - \delta x)}{\delta x(1 - \delta x)}, \\ &= \lim_{\delta x \to 0} \frac{6}{1 - \delta x}. \end{aligned}$$

Evaluating the last limit gives us $dg/dx = 6$. This is how steep the curve of $g(x)$ is at $x$ = 2. Or we can say that this is the rate at which $g(x)$ is changing the moment we arrive at $x$ = 5.

As can be seen from the above examples computing this limit a lot of effort, specially if we have to find the slope of the curve of $f(x)$ at many different points on the curve. However, there is a way around this: the limit expression can be used to find a general equation which represents the slope of $f(x)$ as whole. No need to use the limit expression every time we want to find the slope at a single point. Just use the general equation instead. Each function will have its own general equation created from the limit expression. We will see how to create these generalised expressions in section 1.5 and section 1.11 as well as in the notes Differentiation II.

## 1.3 The derivative as a functions representing the slope of the curve as a whole

### *1.3.1 One way of understanding df/dx as a function*

We saw in the previous section that, in taking the slope of $y = x$, our calculation of $\delta y \div \delta x$ could be based on any triangle formed from any starting point P to any end point Q. It didn't matter where we located the triangle (as illustrated in diagram (1) below), or how big or small the triangle was, we would always get the same result for the slope. Even making different triangles of different sizes would not affect our measurement of the value of the slope (as illustrated in diagram (2) below).

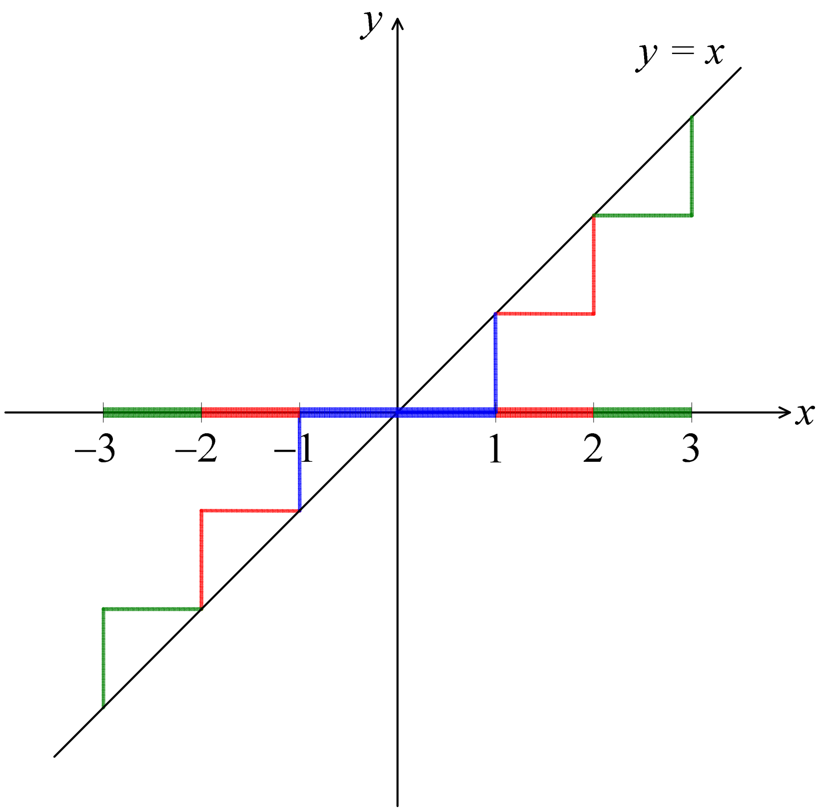



*Diagram (1):*

*Measuring the slope of y = x over intervals of the same lengths*

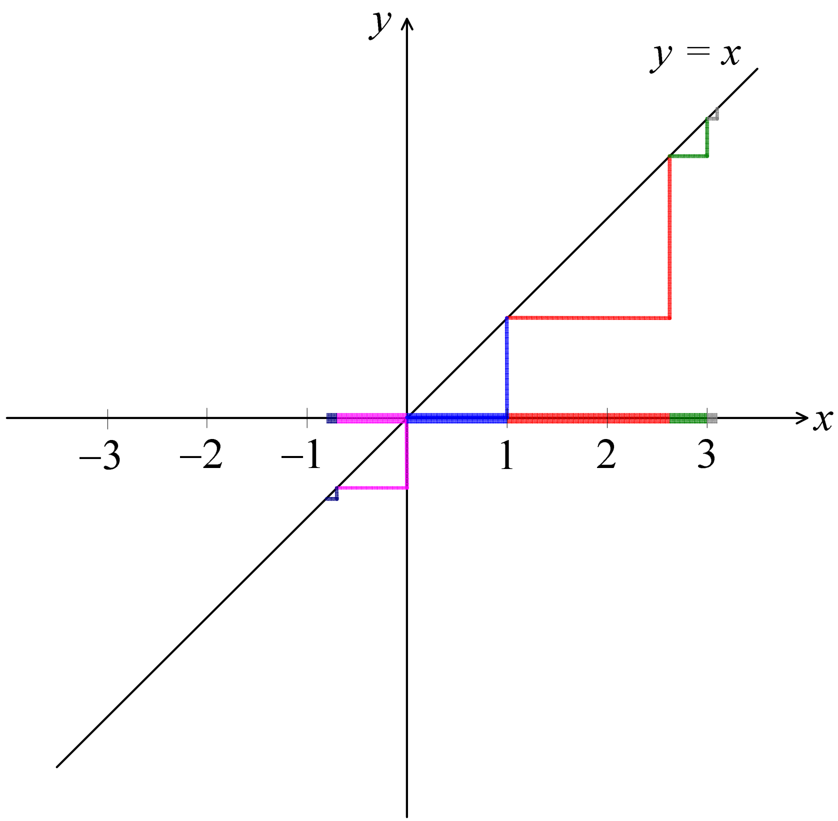



*Diagram (2):*

*Measuring the slope of y = x over intervals of different lengths*

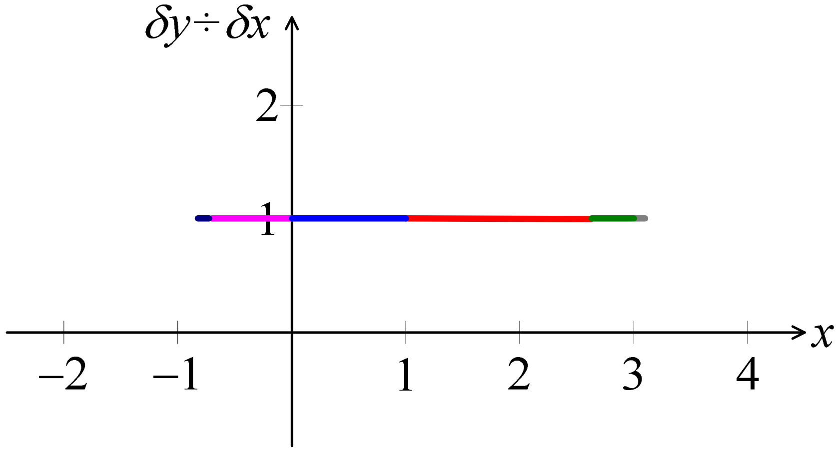
Because of this we only ever needed to form one triangle, and this could be of any size located anywhere along the line $y = x$. More specifically, we could take our $\delta x$ measurement to be of any length anywhere along the $x$-axis.

Now suppose that we want to see the distribution of slopes of $y$ for different intervals along the $x$-axis. In this case we will need to take several slope measurements from different location and over different lengths $\delta x$. The table below shows the interval length $\delta x$ used in diagram (2) and the associated slope over that interval.

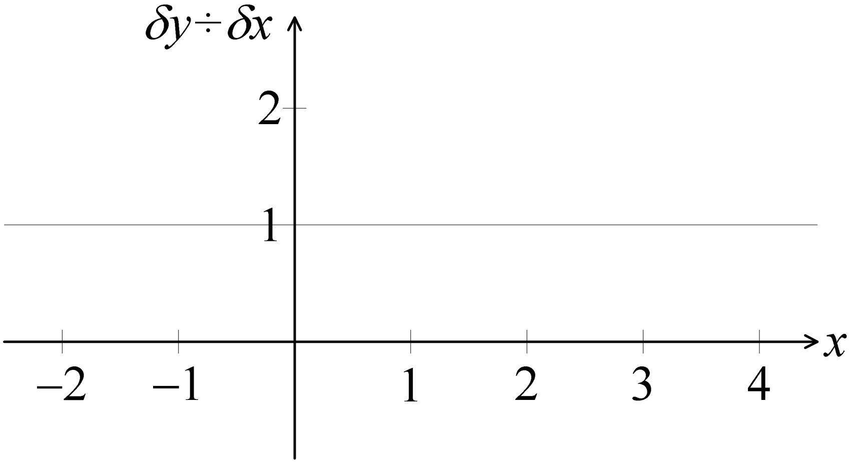
| **Interval (colour)** | … | [– ¾, 0] (pink) | [0, 1] (blue) | [1, 2½] (red) | [2½, 3] (green) | … |
|---|---|---|---|---|---|---|
| $\delta y \div \delta x$ | 1 | 1 | 1 | 1 | 1 | 1 |

*Table of values of the slope of y = x over intervals of different lengths*

We can plot the distribution of these slopes, across the interval over which they apply, as illustrated below:




So, wherever we take our interval, and whatever length our interval is, we see that the slope of $y = x$ over the whole interval [ –¾, 3] is $\delta y \div \delta x = 1$. This slope actually applies for all intervals of any length over any range of $x$ values from $-\infty$ to $+\infty$, so the line segments shown in the graph above actually form part of one continuous line which extends forever left and right as shown below:




But doesn't this line represent a function? Yes. If we call this function $S(x)$ where "$S$" stands for "slope", where this function is the derivative of $y$, then our derivative function is $S(x) = 1$.

We therefore have the situation where the slope $\delta y \div \delta x = 1$, which was only ever measured using one triangle taken over one interval on the $x$-axis, is true when measured over *all intervals of any length located anywhere* along the $x$-axis. We also have the situation where the slope $\delta y \div \delta x = 1$, which was only ever a number, happens now to be a function.

All of this extends the idea of slope from being merely a number to being a function. It extends the idea of slope from (seemingly) being relevant only a one location (or over one interval on the x-axis) to being relevant to $f(x)$ as a whole. These are important extensions to the idea of slope and will be needed for the understanding of the general definition of $df/dx$ to come.

*1.3.2 Another way of understanding df/dx as a function*

Another way of understanding the fact that the derivative is a function is by noticing that, in the difference-quotient, $\delta x$ acts like an independent variable and therefore the whole limit expression acts like a function. Let us call this difference-quotient function $S(\delta x)$. Hence

$$S(\delta x) = \frac{f(6+\delta x) - f(6)}{\delta x}.$$

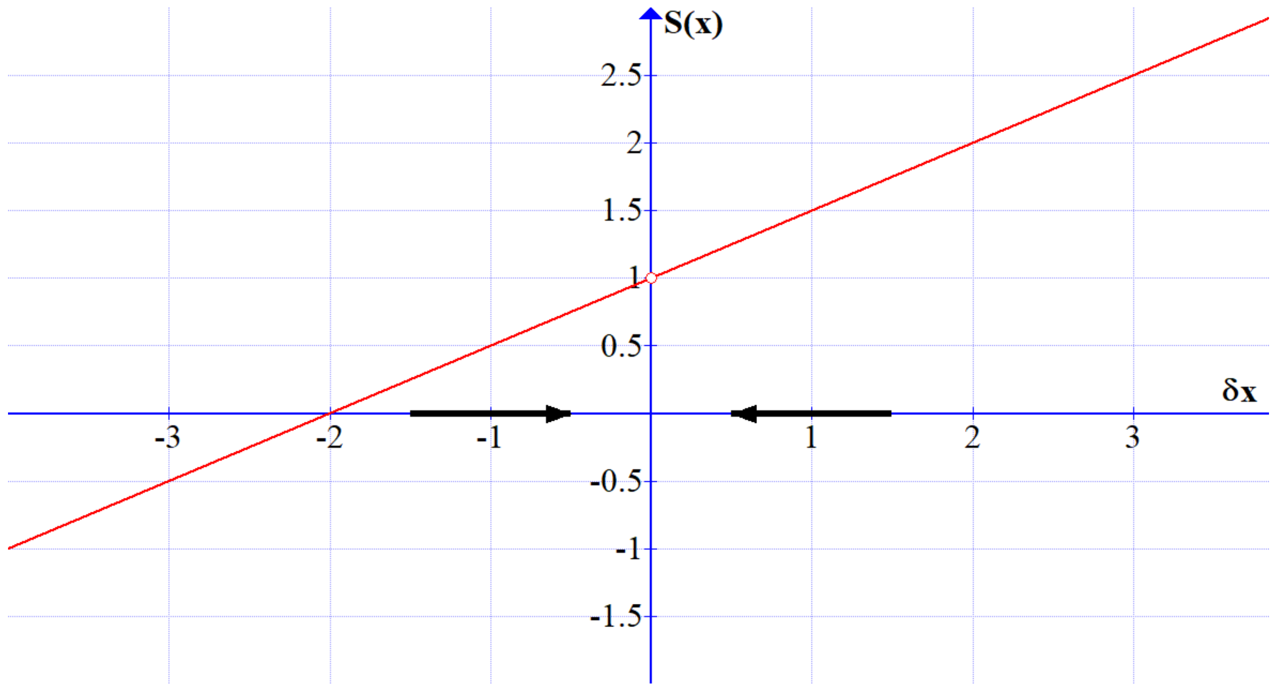
For $f(x) = 0.5(x-5)^2 + 3$ we have

$$S(\delta x) \quad = \quad \frac{[0.5(6-5+\delta x)^2 + 3] - [0.5(6-5)^2 + 3]}{\delta x},$$

$$= \frac{0.5\delta x^2 + \delta x + 3.5 - 3.5}{\delta x},$$

$$= 0.5\delta x + 1 .$$

This is the equation of a straight line. What we can now do is plot this equation. Noting that $S(\delta x)$ represents a slope, readings from the graph of $S(\delta x)$ will represent the slope of the secant PQ of $f(x)$ for various distances $\delta x$. This means that we can use the graph of $S(\delta x)$ to see what happens to the slope of $f(x)$ as $\delta x$ approaches 0:




*Graph of the slope function S(x)*

Note that we are not taking the slope of this line, but that this line already represents the slope of the function $f(x)$. Looking at the graph we can then see that as $\delta x$ approaches 0 (from both the left hand side of 0 and the right hand of 0) $S(\delta x)$ approaches 1. The white/empty circle at (0, 1) represents the fact that, although $\delta x$ can equal 0 in the function $S(\delta x)$, it cannot equal 0 in $\lim_{\delta x \to 0}[f(6 + \delta x) - f(6)]/\delta x$. Nonetheless the derivative of $f(x)$ does exist at $x$ = 6 and is given by

$$\lim_{\delta x \to 0} S(x) = \lim_{\delta x \to 0} \frac{[f(6 + \delta x) - f(6)]}{\delta x} = 1 .$$

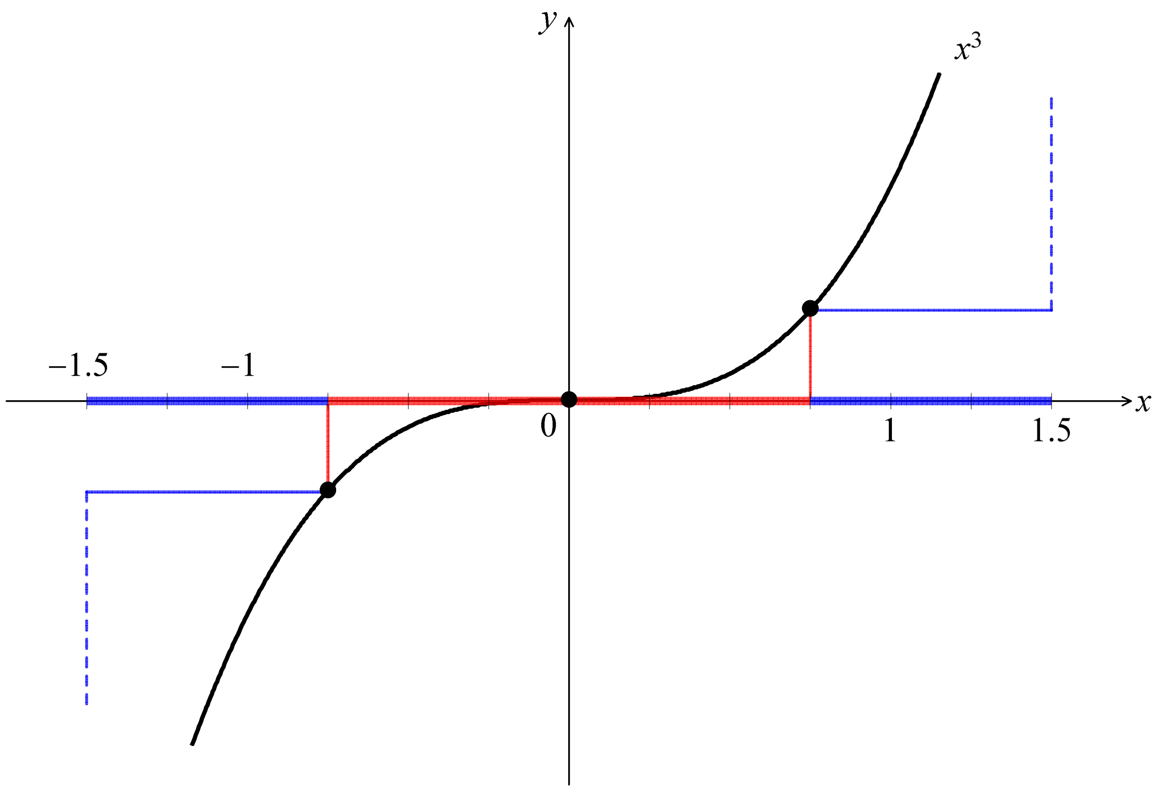
Note that this graph is only valid for finding the slope at $x$ = 6. We would have to draw separate graphs of $S(\delta x)$ if we wanted to find the slope of the curve at other values of $x$.

As mentioned earlier, the graph above can also be used to find the slope of any secant PQ on $f(x) = 0.5(x - 5)^2 + 3$. For example, if $\delta x$ = 2 then the slope of the secant PQ is 2. If $\delta x$ = –2 then the slope of the secant PQ is 0 (i.e. PQ forms a horizontal line. Question: where would Q have to be on the graph for it to be level with P?)

*Another example*

Consider wanting to find the function which represents the derivative of $f(x) = x^3$. We want to see the distribution of slopes of $f(x)$ for different intervals along the $x$-axis. In this case we will need to take several slope measurements over different interval lengths $\delta x$.

Let us start by taking slope measurements at intervals of $\delta x$ = 0.75. Doing so we obtain the following set of triangles from which we can calculate the slope of the secants:

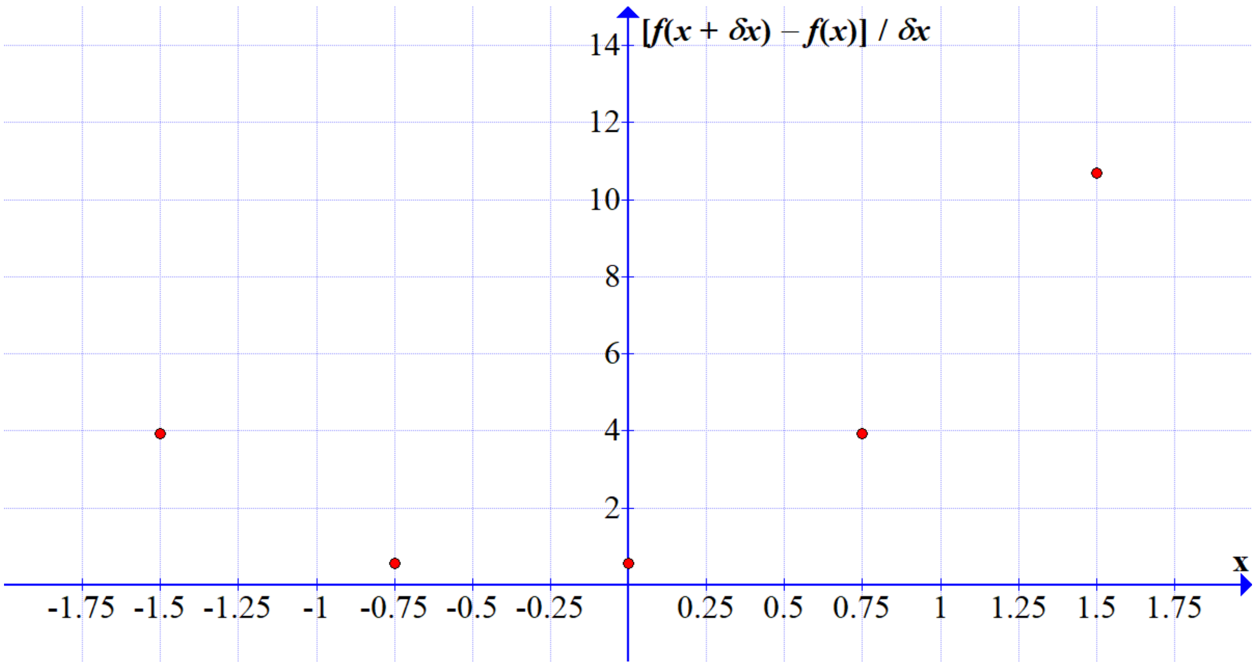



The calculation of the slopes according to $[f(x + \delta x) - f(x)]/\delta x$ at different values of $x$, with $\delta x$ = 0.75, can be seen in the table below:

| x = | -1.5 | -0.75 | 0 | 0.75 | 1.5 |
|---|---|---|---|---|---|
| [f(x + δx) - f(x)] / δx = | 3.9375 | 0.5625 | 0.5625 | 3.9375 | 10.6875 |

*Slope data for $f(x) = x^3$ when $\delta x = 0.75$*

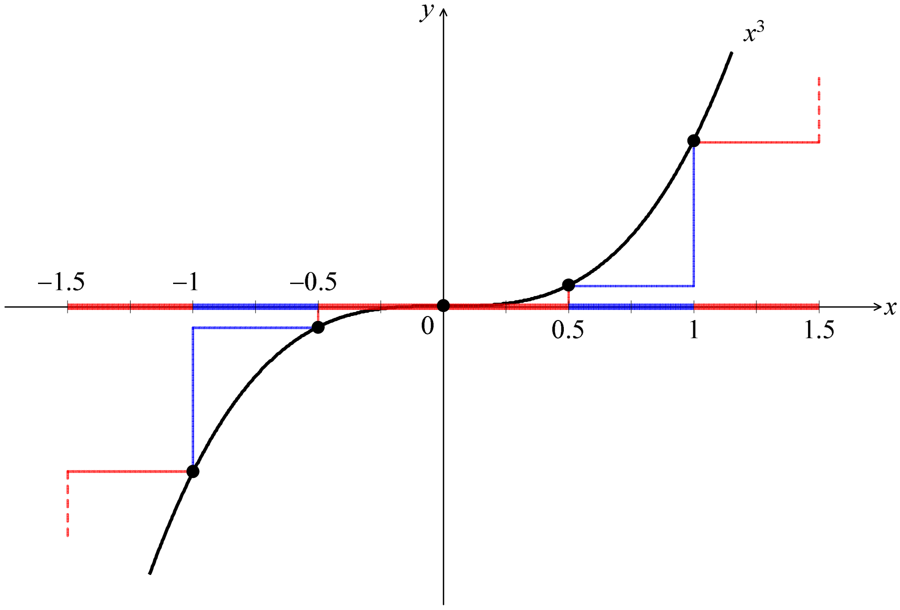
What this data represents is the distribution of the slopes of $f(x) = x^3$. The shape of this distribution should form some kind of coherent and recognisable pattern. To see whether this is so we can plot the above data, and this is shown as the red points in the graph below:




*Distribution of slopes of $f(x) = x^3$ when $\delta x = 0.75$*

This sequence of data points does not look like any recognisable pattern which could be represented by a function. This is because $\delta x = 0.75$ is a far too large an interval to take for an accurate representation of the derivative of $f(x)$, and the sequence data points above do not suggest what function this derivative could be. We will therefore have to take a smaller interval.

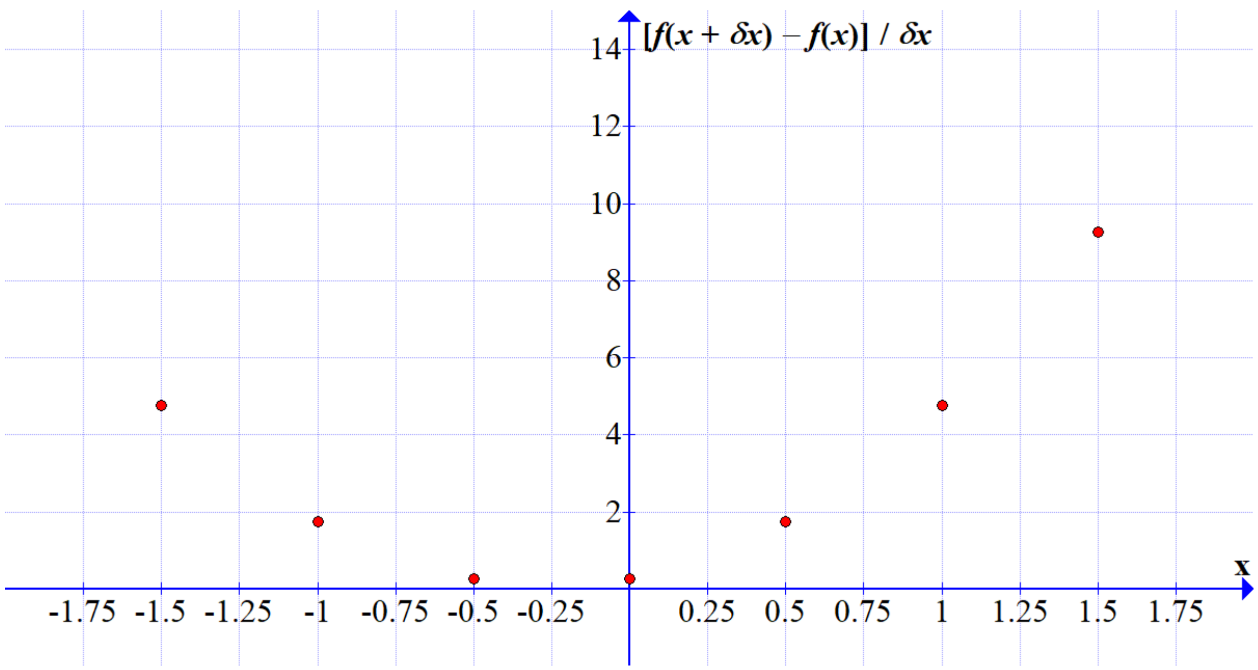
So, taking $\delta x = 0.5$ (seen by the red and blue bars on the $x$-axis) we obtain the following set of triangles from which we can calculate the slope of the secants:




This gives the following slope data and graphs:

| x = | -1.5 | -1 | -0.5 | 0 | 0.5 | 1 | 1.5 |
|---|---|---|---|---|---|---|---|
| $\frac{f(x + \delta x) - f(x)}{\delta x}$ | 4.75 | 1.75 | 0.25 | 0.25 | 1.75 | 4.75 | 9.25 |

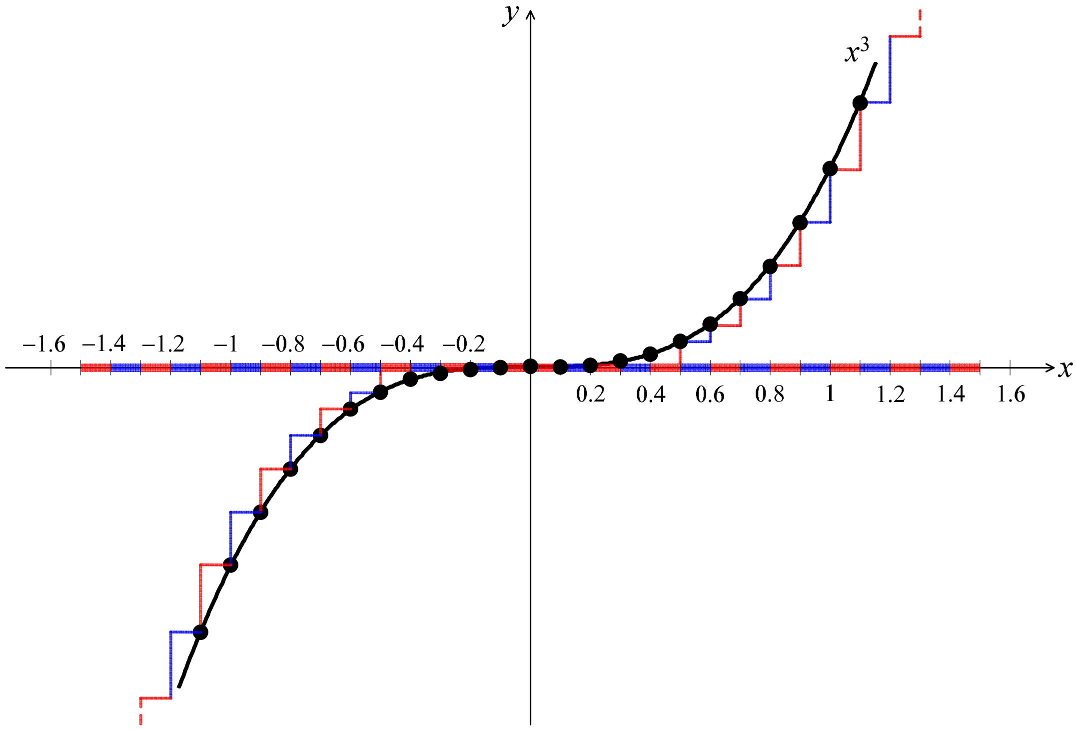
*Slope data for f(x) = $x^3$ when δx = 0.5*




*Distribution of slopes of f(x) = $x^3$ when δx = 0.5*

It looks like we might be getting somewhere. The sequence of calculations of $[f(x + \delta x) - f(x)]/\delta x$, when $\delta x = 0.5$, has produced a set of data that now seem to follow a coherent pattern. Reducing $\delta x$ further will allow us to see if this pattern continues.

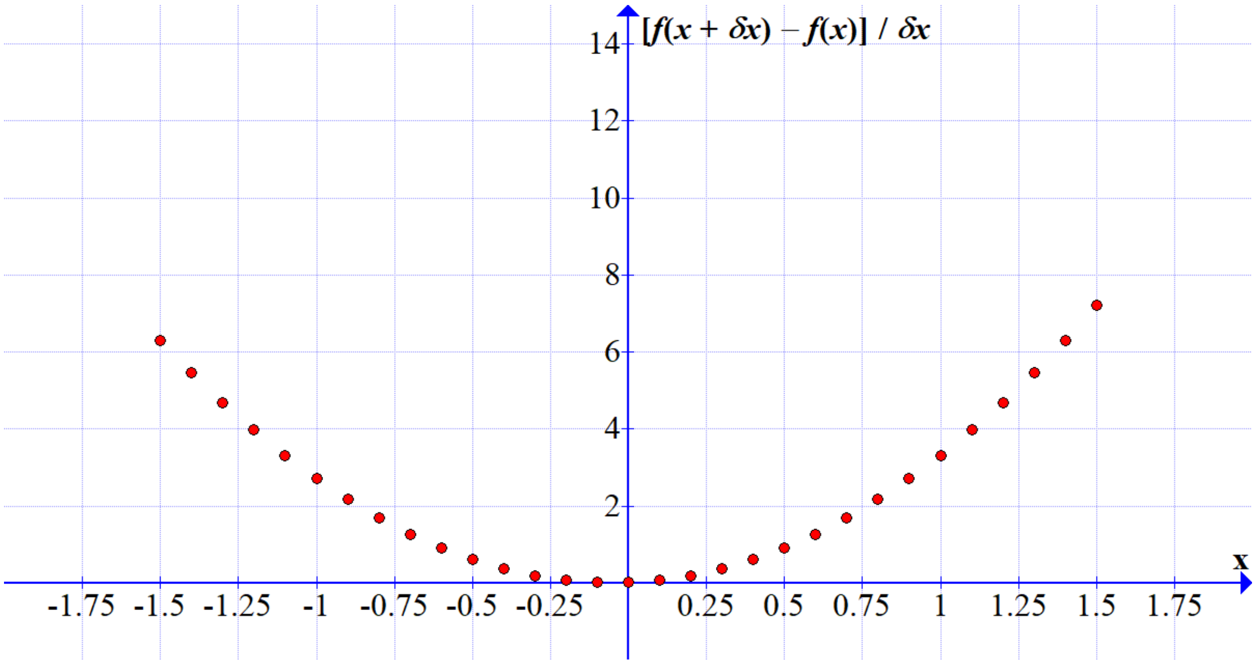
So, taking $\delta x = 0.5$ (seen by the red and blue bars on the *x*-axis) we obtain the following set of triangles from which we can calculate the slope of the secants:




This gives the following slope data and graphs:

| x = | -1.5 | -1.4 | ... | 0 | ... | 1.4 | 1.5 |
|---|---|---|---|---|---|---|---|
| $\frac{f(x + \delta x) - f(x)}{\delta x}$ | 6.31 | 5.47 | ... | 0.01 | ... | 6.31 | 7.21 |

*Slope data for $f(x) = x^3$ when $\delta x = 0.1$*

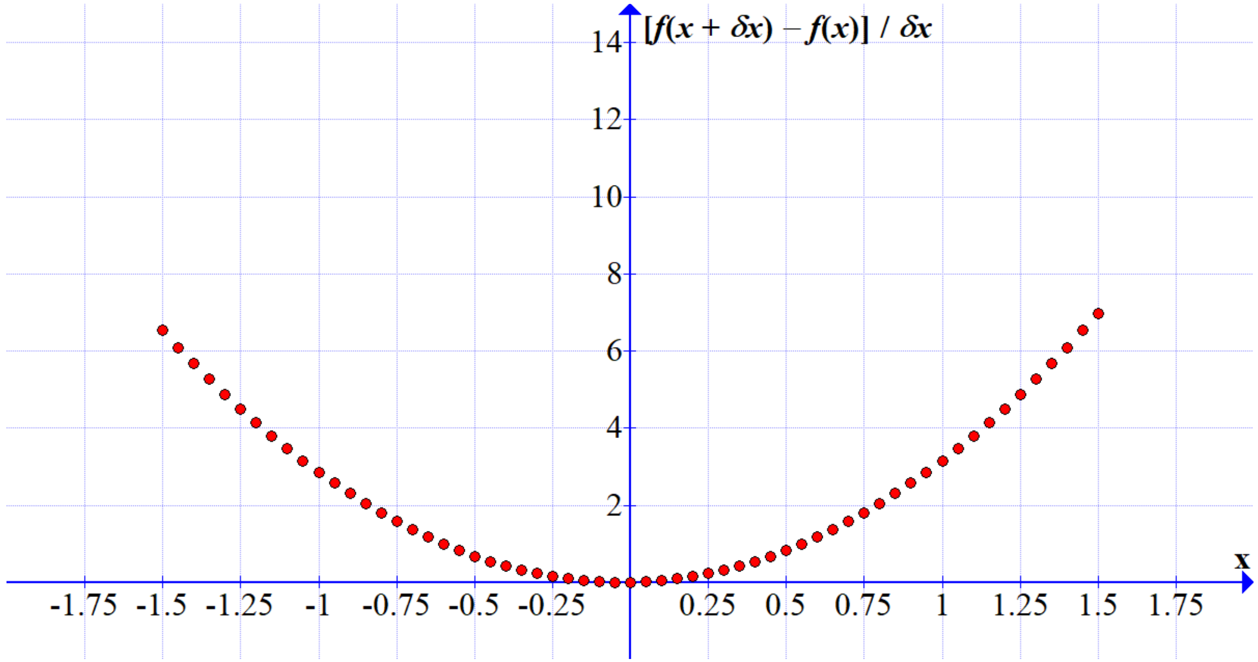



*Distribution of slopes of $f(x) = x^3$ when $\delta x = 0.1$*

Now this looks more promising. The shorter interval of $\delta x = 0.1$ shows that the set of data seems to give an $x^2$ type of pattern. This could be a coincidence so let reduce the interval $\delta x$ to $\delta x = 0.05$. For practical reasons I will not show the districution of triangles along the curve of $f(x) = x^3$ only the table of data and the distribution of slopes, the result of this are shown on the next page.

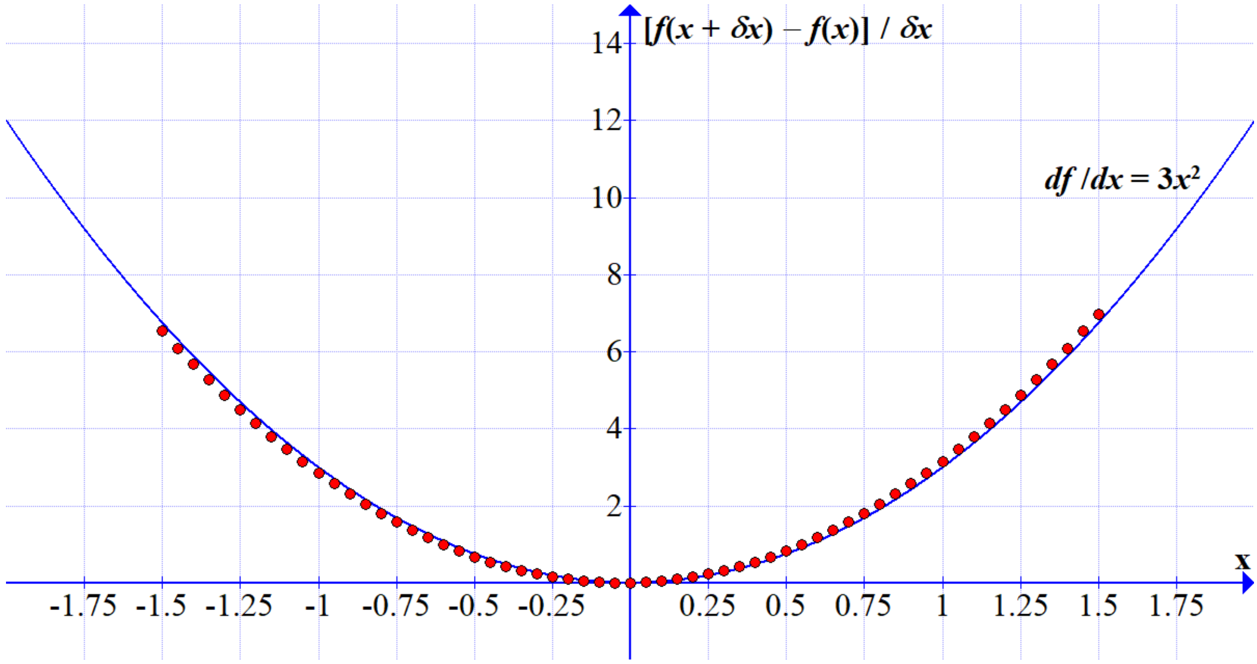
| δx = | 0.05 | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| x = | -1.5 | -1.45 | -1.4 | -1.35 | … | -0.1 | -0.05 | 0 | 0.05 | 0.1 | … | 1.3 | 1.35 | 1.4 | 1.45 | 1.5 |
| $\frac{f(x + \delta x) - f(x)}{\delta x}$ | 6.5275 | 6.0925 | 5.6725 | 5.2675 | … | 0.0175 | 0.0025 | 0.0025 | 0.0175 | 0.0475 | … | 5.2675 | 5.6725 | 6.0925 | 6.5275 | 6.9775 |

*Slope data for $f(x) = x^3$ when $\delta x = 0.05$*




*Distribution of slopes of $f(x) = x^3$ when $\delta x = 0.05$*

The graph above definitely resembles an $x^2$ type of graph. It could resemble an $x^4$ type of graph or a graph of another even power of *x*, but this is unlikely since these graphs would have appeared much more tightly packed towards the *y*-axis. In any case we can confirm our assumption by testing various functions. Ultimately we see that the set of data most closely resembles the function $3x^2$, as illsutrated by the solid curve in the graph below:



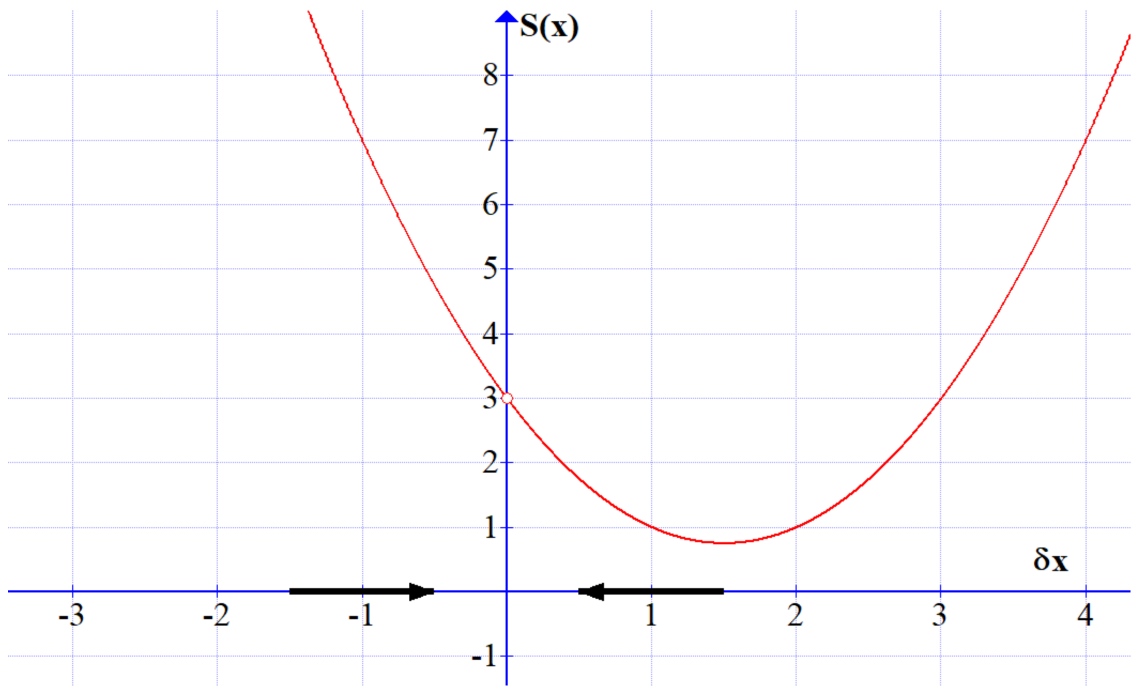

In fact, as $\delta x$ approaches 0 the set of slope data would take the exact same path as the function $3x^2$, so our derivative function is $S(x) = 3x^2$.

So, the slope graph, which started off as a collection of data with no coherent pattern of data points has ended up looking like a coherent pattern of data points satisfying the function $S(x) = 3x^2$. This happened because of a continual shortening of the interval $\delta x$ over which we built the triangles we used to calculate our slope. This continual interval shortening therefore seems to be the way forward when it comes to getting the correct function which represents the derivative of $y = f(x)$. We are now in a position to formally develop the full and proper version of this approach to finding the derivative of a function. This we will do in the next section.

To finish this section we can repeat the activity of plotting the difference-quotient function $S(\delta x)$, for a given value of $x$, by treating $S(\delta x)$ as a function of $\delta x$. Hence, if we want to find the slope of $f(x) = x^3$ at $x = -1$ we have

$$\begin{aligned} S(\delta x) &= \frac{f(-1+\delta x) - f(-1)}{\delta x}, \\ &= \frac{(-1+\delta x)^3 - (-1)^3}{\delta x}, \\ &= \frac{(\delta x)^3 - 3(\delta x)^2 + 3\delta x - 1 - (-1)}{\delta x}, \\ &= (\delta x)^2 - 3\delta x + 3\,. \end{aligned}$$

This is a quadratic in $\delta x$. Plotting this quadratic against $\delta x$ we get the following graph:




*Graph of the slope function S(x)*

As before, note that we are not taking the slope of this curve, but that this curve already represents the slope of the function $f(x) = x^3$. Looking at the graph we can then see that as $\delta x$ approaches 0 (from both the left hand side of 0 and the right hand of 0) $S(\delta x)$ approaches 3. This is what we expect since if we put $x = -1$ into $S(x) = 3x^2$ we get $S(x) = 3$.

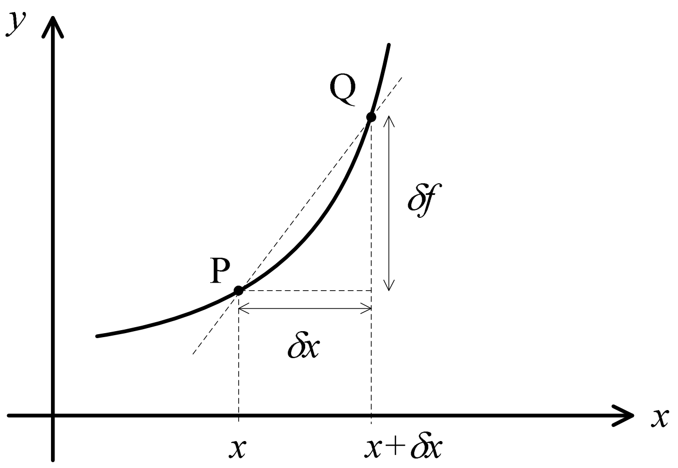
The white/empty circle at (0, 3) represents the fact that, although $\delta x$ can equal 0 in the function $S(\delta x)$, it cannot equal 0 in $\lim_{\delta x \to 0} [f(-1+\delta x) - f(-1)]/\delta x$. Nonetheless the derivative of $f(x)$ does exist at $x$ = –1 and is given by

$$\lim_{\delta x \to 0} S(x) = \lim_{\delta x \to 0} \frac{[f(-1+\delta x) - f(-1)]}{\delta x} = 3\,.$$

Note that this graph is only valid for finding the slope at $x$ = –1. We would have to draw separate graphs of $S(\delta x)$ if we wanted to find the slope of the curve at other values of $x$.

### 1.4 On the formal definition of the first derivative

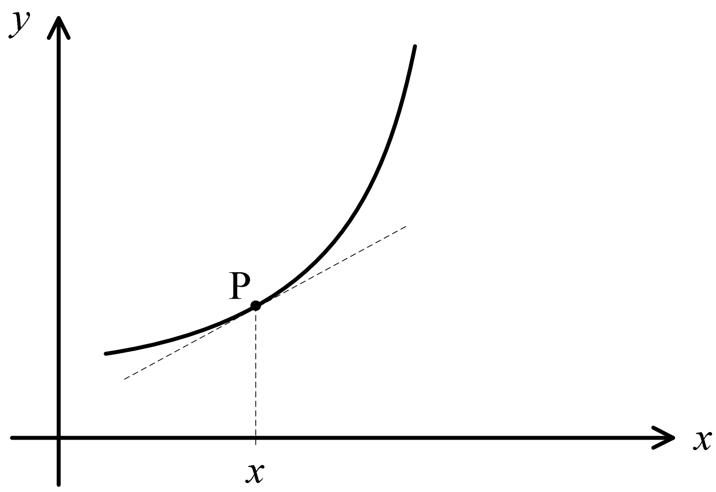
We are now in a position to more formally define the expression for the derivative of $y = f(x)$. Therefore, consider a point P($x$, $y$) of the curve of a function $y = f(x)$. Let P be fixed and let another point Q($x + \delta x$, $y + \delta y$) be located on the right hand side of P as shown below:




The slope of secant PQ can then be expressed as

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

where the right hand side is known as the difference quotient. Now let Q continually approach P. Then provided Q ≠ P (i.e. provided $\delta x \neq 0$) the secant becomes a tangent, and we have the slope of the curve at point P, as illustrated below:

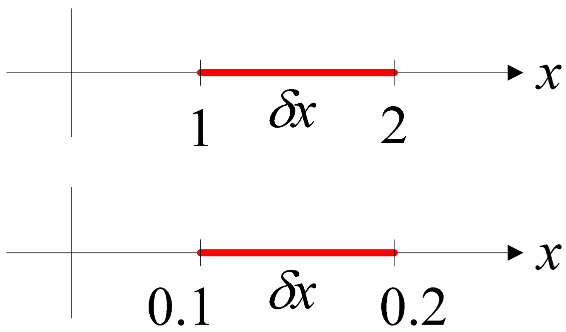


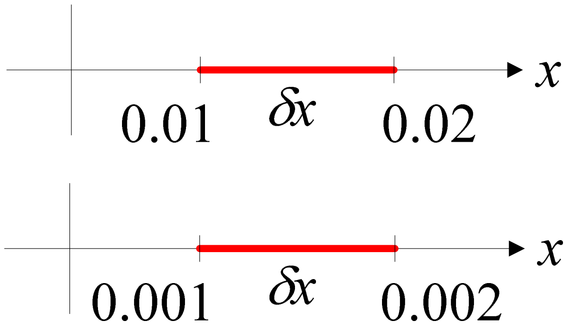

The definition of the first derivative $df/dx$ of a function $f(x)$ is then given by the limit as $\delta x \to 0$ of the difference quotient:

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

It is the concept of the limit as $\delta x \to 0$ which makes the derivative (and calculus) what it is. If there was no limit we would simply be dividing a very small, but finite, number $\delta f$ by another very small, but finite, number $\delta x$.

However, because we are applying the limit these two values become forever smaller and smaller until the end of time. Although it looks like we are going to end up with an answer of 0/0 we actually end up with a value based on a brand new way of thinking, namely that of "taking limits", i.e. the concept of forever approaching a specified number (say, 0) without actually ever reaching it. This continual approach to 0 can be illustrated by the sequence of diagrams below:






Another way of describing the concept of $\delta x \to 0$ without ever becoming 0 is by saying that $\delta x$ can get arbitrarilly close to 0. What does "arbitrarily close" mean? Well, think of a number that is as close to 0 as possible. I can always think of a number closer than this simply by dividing your number by 2. Then, you can get closer still by dividing my number by 2 again, etc. This is what it means for us to be able to choose a number $\delta x$ which is arbitrarilly close to 0 without actually being 0.

This limiting process means we actually end up with a value to $df/dx$ which represents the slope, and more generally, the *instantaneous rate of change of f(x) with respect to x*. What $df/dx$ also represents is a brand new mathematical object called a *derivative*, which is a function representing the instantaneous rate of change of *f(x)* at any point *x*.

Returning to the above definition of the derivative we might think that equation (1) applies only when Q approaches P from the right hand side (diagram 1 below). However, equation (1) is also valid when Q approaches P from the left hand side. To see why let $X = x + \delta x$ be our point P (diagram 2 below). Then our point Q is given by $X - \delta x = x$, which lies on the left hand side of P. If we now substitute this into equation (1) to get $\lim_{\delta x \to 0}[f(X) - f(X - \delta x)]/\delta x$, where the expression $f(X) - f(X - \delta x)$ is simply the vertical distance over horizontal distance as seen from the left hand side of P. What this means is that this last expression is automatialy implied by equation (1) when Q is on the left hand side of P.

Therefore in summary we have

$$\frac{f(x+\delta x)-f(x)}{(x+\delta x)-x}$$

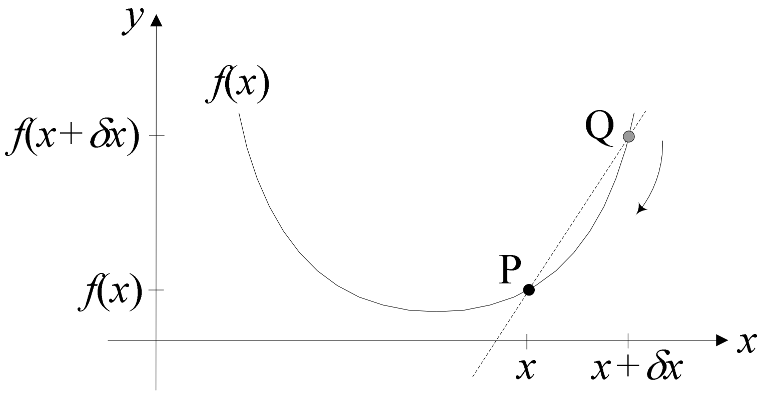



*Diagram 1: Q approaches P from the right hand side*

$$\frac{f(X)-f(X-\delta x)}{X-(X-\delta x)}$$

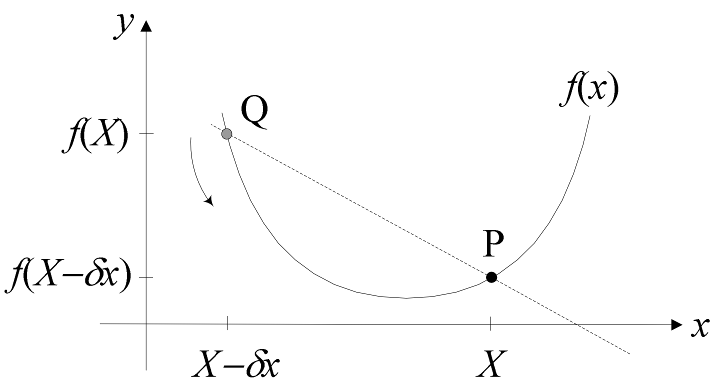
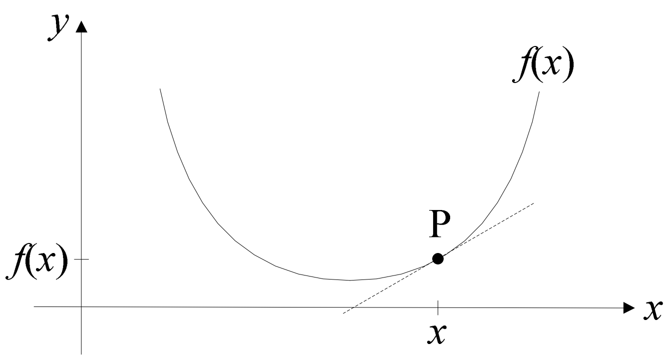


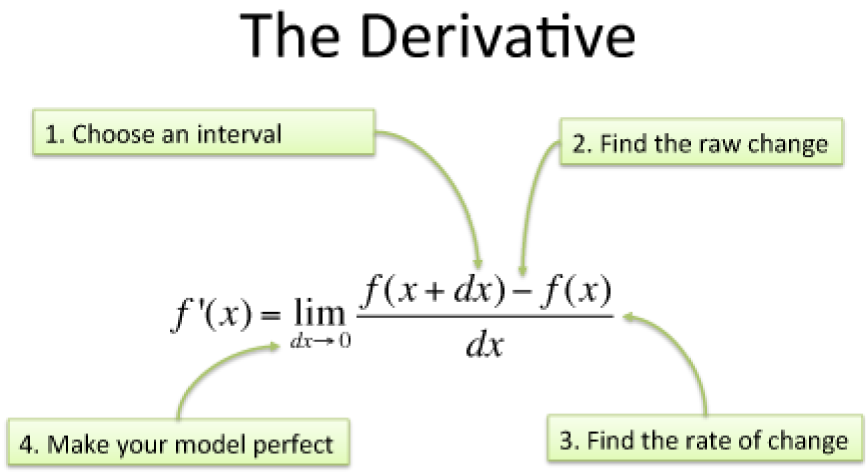

*Diagram 2: Q approaches P from the left hand side*

Therefore, approaching P from either direction leads to the same result:




$$\frac{df}{dx}=\lim_{\delta x\to 0}\frac{f(x+\delta x)-f(x)}{\delta x} \qquad \text{is the same as} \qquad \frac{df}{dx}=\lim_{\delta x\to 0}\frac{f(x)-f(x-\delta x)}{\delta x}$$

Below is a diagram I got off the internet (https://betterexplained.com/calculus/lesson-10/) which may be useful in interpreting expression (1). Here they have used the notation $f'(x)$ for the derivative which is another way of writing $df/dx$.




At this stage it is important to understand that $df/dx$ is not a division. We are not doing $df \div dx$. Rather, $df/dx$ is an operation (just like +, –, × and ÷ are operations). In this case the operation is that of differentation, and the operator is

$$\frac{d}{dx}.$$

When we get to the differentiation topic of related rates, as well as the topic of integration, we will see that *df* and *dx* can be considered separately, that we will be able to divide these, and that *df* ÷ *dx* will have its own meaning. In this case "*df*" and "*dx*" are called "differential elements" and represent infinitely small distances. However, at this stage in our learning we will interpret $df/dx$ as the operator shown above.

### 1.5 The derivative of $x^n$ from 1st principles

We are now in a position to find the general expression for the derivative of a number of basic functions, the first of which is $f(x) = x^n$ where $n \in \mathbb{R}$. O starting point for doing this will always be the definition of the derivative given as equation (1). Finding derivatives based on this expression is called *differentiating from 1st principles*.

For the reason that we haven't yet covered enough on the topic of differentiation we will first find the derivative of $f(x) = x^n$ only for the case where $n \in \mathbb{N}$ (i.e. $n$ is a positive integer only). When we ave covered the necessary extra differentiation theory (in the Differentiation II notes) we will see how we can find the derivative when $n$ is a real number.

So, applying the definition of the derivative to $x^n$ we get

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

The first question is, Can we evaluate the limit? Is there an answer to evaluating the limit as it stands? If there is then we have the answer to the derivative. Usually, however, we will not be able to evaluate the limit as it is given here (because it leads to a "0/0" situation which is not a valid result).

In the case above we cannot evaluate the limit as it stabds so we have to find a way of transforming the expresion into something whereby we can evaluate the limit. This is usually done by using algebra in a judicious way so as to get a limit which we can evaluate.

So, one of the more obvious things we can do is ot use the binomial theorem. To expand th term $(x + \delta x)^n$. We don't know if this will help us get to a point hwere we can evaluate the limit, but we have to try something. Doing this gives us

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{1}{\delta x}\left[x^n + n.\delta x.x^{n-1} + \frac{n(n-1)}{2!}(\delta x)^2 x^{n-2} + \cdots + n(\delta x)^{n-1}x + (\delta x)^n - x^n\right].$$

Cancelling the first and last terms, and dividing the rest by $\delta x$ we get

$$\frac{df}{dx} = \lim_{\delta x \to 0}\left(n.x^{n-1} + \frac{n(n-1)}{2!}\delta x.x^{n-2} + \cdots + n(\delta x)^{n-2}x + (\delta x)^{n-1}\right).$$

Note that since the first term in the limit is independent of $\delta x$ we can take it out of the limit. Hence we have

$$\frac{df}{dx} = n.x^{n-1} + \lim_{\delta x \to 0}\left(\frac{n(n-1)}{2!}\delta x.x^{n-2} + \cdots + n(\delta x)^{n-2}x + (\delta x)^{n-1}\right). \qquad (2)$$

Again we ask, Can we evaluate the limit? In this case we can. So, as $\delta x \to 0$ all terms inside the bracket become zero, and we are left with

$$\frac{df}{dx} = n.x^{n-1} \qquad (3)$$

Hence the operation of the first derivative is to transform $f(x) = x^n$, where $n \in \mathbb{N}$, into $df/dx = n.x^{n-1}$. In other words,

$$\text{If } f(x) = x \text{ then } \frac{df}{dx} = 1.x^{1-1} = 1.x^0 = 1$$

$$\text{If } f(x) = x^2 \text{ then } \frac{df}{dx} = 2.x^{2-1} = 2x$$

$$\text{If } f(x) = x^3 \text{ then } \frac{df}{dx} = 3.x^{3-1} = 3x^2$$

$$\text{If } f(x) = x^4 \text{ then } \frac{df}{dx} = 4.x^{4-1} = 4x^3$$

… etc. The mechanics of the process of differentiation is therefore to bring the power down as a multiple, and then reduce the power by 1.

But what happens if $n$ is a fraction such as ½? What happens if $n$ is negative? And finally, what happens if $n$ is a real number such as π or √2 ? What, then, is the derivative of $f(x) = x^{\pi}$ or $f(x) = x^{\sqrt{2}}$ ? Since $\frac{d}{dx}$ is a brand new operation we can't assume that expression (3) is also true when $n$ is a real number. To show that the above formula works for all real values of $n$ (i.e. $n \in \mathbb{R}$) we will need to learn some rules of differentiation which we will get to in the Differentiation II notes.

There is, now, an important point to note: it might seem that in going from (2) to (3) we have simply put $\delta x = 0$. If this is true then what is the point of the limit idea? Well, the truth is that we do NOT do $\delta x = 0$. We do do $\delta x \to 0$. It just so happens that we get the same answer as if we had done $\delta x = 0$. But this is just a coincidence. More advanced maths would be needed to prove that (2) does reduce to (3) as $\delta x \to 0$, but we can get an idea that this is so by looking at a table of values as we let $\delta x \to 0$ in (2):

| $\delta x$ | $\delta f / \delta x$ |
|---|---|
| 0.1 | $n.x^{n-1} + \frac{n(n-1)}{2!}(0.1).x^{n-2} + \frac{n(n-1)(n-2)}{3!}(0.1)^2.x^{n-3} + \ldots + (0.1)^{n-1}$ |
| 0.01 | $n.x^{n-1} + \frac{n(n-1)}{2!}(0.01).x^{n-2} + \frac{n(n-1)(n-2)}{3!}(0.01)^2.x^{n-3} + \ldots + (0.01)^{n-1}$ |

| | |
|---|---|
| 0.001 | $n.x^{n-1}+\frac{n(n-1)}{2!}(0.001).x^{n-2}+\frac{n(n-1)(n-2)}{3!}(0.001)^2.x^{n-3}+$ $\ldots+(0.001)^{n-1}$ |
| 1×10$^{-4}$ | $n.x^{n-1}+\frac{n(n-1)}{2!}(1\times10^{-4}).x^{n-2}+\frac{n(n-1)(n-2)}{3!}(1\times10^{-4})^2.x^{n-3}+$ $\ldots+(1\times10^{-4})^{n-1}$ |
| 1×10$^{-5}$ | $n.x^{n-1}+\frac{n(n-1)}{2!}(1\times10^{-5}).x^{n-2}+\frac{n(n-1)(n-2)}{3!}(1\times10^{-5})^2.x^{n-3}+$ $\ldots+(1\times10^{-5})^{n-1}$ |
| … | … |
| 1×10$^{-1000}$ | $n.x^{n-1}+\frac{n(n-1)}{2!}(1\times10^{-1000}).x^{n-2}+\frac{n(n-1)(n-2)}{3!}(1\times10^{-1000})^2.x^{n-3}$ $\ldots+(1\times10^{-1000})^{n-1}$ |
| … | … |

Therefore as $\delta x \to 0$ we will be left with only the term $n.x^{n-1}$. The results above may seem obvious but we shall see later that there is nothing obvious about evaluating limits when $\delta x \to 0$.

*Examples*

1) Despite my earlier comment about us having to wait until the Differentiation II notes before we can show the derivative formula (3) works for all real *n* and not just for integer *n*, we can still give an example of the derivative for fractional powers.

   Let us therefore consider finding the derivative of if $f(x) = x^{1/2}$. From 1st principles we have

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

Since the powers in each term of the numerator are not integers we can't used the binomial theorem to expand these (if you know about Taylor series then we could use this, but we won't go there at this point).

So what we do is to use an algebraic trick which will convert these powers into integers. Specificially we are going to use a variation on the idea of the difference of two squares which says that $(a-b)(a+b)=a^2-b^2$. So, the numerator of our fraction above is the same type of object as "$a$ – $b$". what we want is to introduce an "$a$ + $b$" term, ans we do this as follows:

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

What we have done here is to multiply our original fraction by another fraction which happens to equal 1. But this other fraction is made in such a way as to allows us to get the $(a-b)(a+b)$ structure we want in the numerator.

Multiplying out the top and bottom of this last expression we get

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

Doing this has significantly reduced the level of difficulty of the numerator. You might think that we have only increased the level of difficulty in the denominator but, as we shall in what follows, we will be able to handle this.

The above expression simplifies to

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

Since $\delta x$ approaches 0 the term $(x+\delta x)^{1/2}$ approaches $x^{1/2}$, hence we end up with

$$\frac{df}{dx}=\frac{1}{2}\frac{1}{x^{1/2}}.$$

2) Let $f(x)=x/(x+1)$. Then from 1st principles we have

$$\frac{df}{dx}=\lim_{\delta x\to 0}\left(\frac{\frac{(x+\delta x)}{[(x+\delta x)+1]}-\frac{x}{x+1}}{\delta x}\right).$$

Cross multiplying we get

$$\frac{df}{dx} = \lim_{\delta x \to 0} \left( \frac{(x+1)(x+\delta x) - x[(x+\delta x)+1]}{\delta x(x+1)[(x+\delta x)+1]} \right).$$

We could expand everything in the numerator and simplify as necessary, but a quicker way is to notice that there is an $x(x + 1)$ terms in both parts of the numerator, viz

$$\frac{df}{dx} = \lim_{\delta x \to 0} \left( \frac{x(x+1) + \delta x(x+1) - x[(x+1)+\delta x]}{\delta x(x+1)[(x+\delta x)+1]} \right).$$

Because of the minus sign these two terms cancel and we are left with

$$\frac{df}{dx} = \lim_{\delta x \to 0} \left( \frac{\delta x(x+1) - x.\delta x}{\delta x(x+1)[(x+1)+\delta x]} \right),$$

where I have regrouped terms in the squared bracket of the denominator for visual effect. Simplifying further we get

$$\frac{df}{dx} = \lim_{\delta x \to 0} \left( \frac{1}{(x+1)[(x+1)+\delta x]} \right).$$

Since $\delta x$ approaches 0 the term $[(x+1)+\delta x]$ approaches $(x+1)$, hence we end up with

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

Example 2) was one of finding the derivative of a fraction. In the Differentiation II notes we will find general formulae for finding the derivatives of fractions which will make it much easier for us to perform the operation of differentiation on fractions.

### 1.6 The derivative as a transformation from position to slope

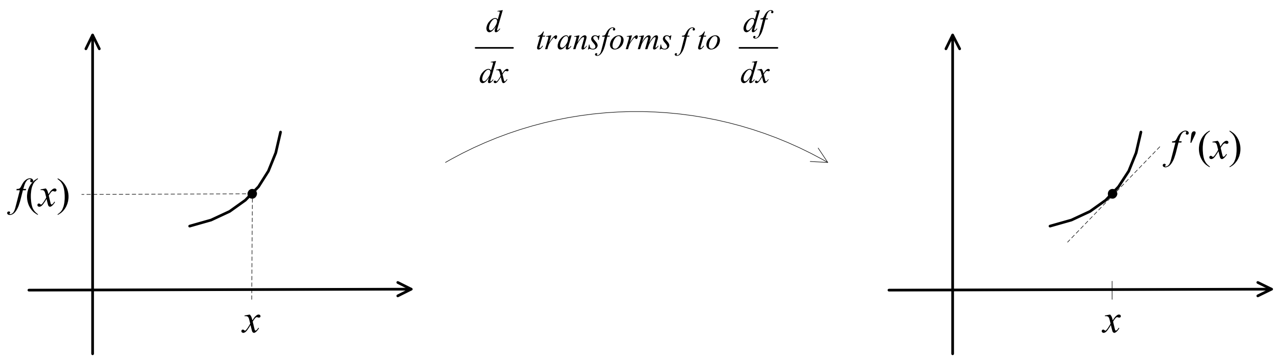
In previous sections we saw that $df/dx$ could be considered as the gradient or slope of a curve. However, there is another way of interpreting the derivative. In order to see this let us first understand what a functions is and does.

The usual way of thinking about functions is as a formula, usually written as $y = f(x)$. We put a number $x$ into our formula, and it gives us the answer $y$. We might therefore say that the formula transforms a value $x$ value into a value $y$.

In the practical context of geometry, and geometric measurement, $y = f(x)$ is considered to specify a vertical position/distance, which might be seen as height or depth, above or below the $x$-axis for a given horizontal position/distance $x$ (in the more abstract terms of advanced maths we would say that $f$ transforms a value $x$ into a value $f(x)$ according to a mapping given by $f$).

This idea of transformation applies to all functions, and since the derivative is a function it also applies to $df/dx$. So, given a function $y = f(x)$ (which itself is a transformation of $x$ into $y$), $df/dx$ is which represents the rate of change of $f(x)$ at a point $x$.

Since functions are objects which specify position ($f(x)$ as a vertical distance for a given horizontal distance $x$), and derivatives are objects which specify slope ($df/dx$ as a slope of $f(x)$ at a given position $x$), $d.../dx$ can be interpreted as an operation which transforms a "position function" $f(x)$ into a "slope function" $df/dx$. This is represented diagrammatically below:




We have already seen an example of such transformations in section 1.3.1 and 1.3.2. In this latter section we saw that, when where we were trying to find the derivative of $f(x) = x^3$, we ended up with the derivative function to be $df/dx = 3x^2$. This can be interpreted as follows:

The operation of differention $f(x) = x^3$ into $df/dx = 3x^2$

has the effect of transforming a position function into a slope function.

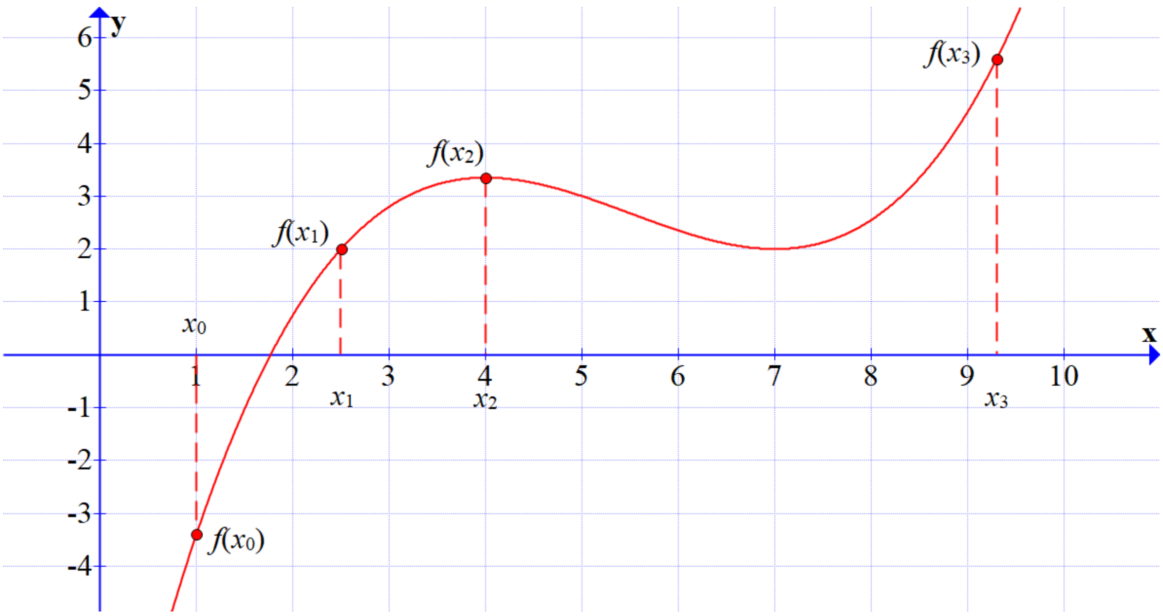
In summary, the derivative can be intrerpreted as follows:

| | **Geometric concept** | **Algebraic concept** |
|---|---|---|
| ***df/dx*** | <u>Tangent</u><br>$df/dx$ represents the slope<br>or gradient of $f(x)$<br>at one point $(x_1, y_1)$ on the curve. | <u>Function</u><br>$df/dx$ is a general expression representing the rate of change of $f(x)$ at a point $(x, y)$ anywhere on the curve.<br><br><u>Transformation</u><br>$df/dx$ represents the transformation of<br>a position function (i.e. $f(x)$)<br>to a rate of change function (i.e. $df/dx$). |

## 1.7 The derivative as a measure of sensitivity

One of the usual ways of thinking about the derivative is as a slope. For a given function $f(x)$ we can find out how steep it is at any given point $x$. Another way of interpreting the derivative is as a measure of the sensitivity of $f(x)$ at a point $x$, in other words, how sensitive $f(x)$ is when we nudge $x$ by a small amount. This section is aimed at explaining this conception of the derivative.

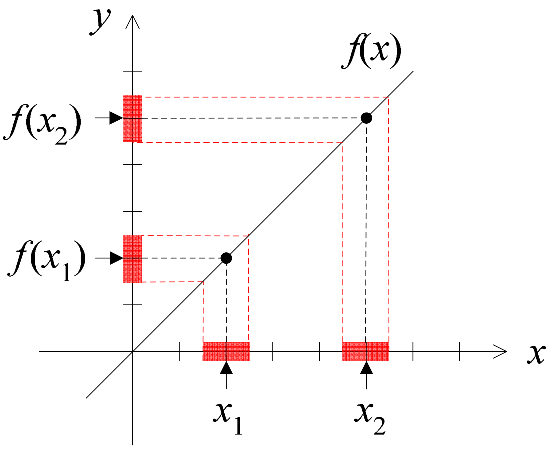
Before we get to this we first need to develop a new understanding of what functions do. The usual way of thinking about functions is as a formula which takes a number $x$ and gives us an answer $y$ according to a formula given by $f(x)$. Mathematically we write $y = f(x)$. Geometrically speaking we can say that for any horizontal distance $x$, the function $f(x)$ represents a vertical distance, or height, above or below $x$, as illustrate below for some function $f(x)$:




*A function seen geometrically as a measure of height*

The question now is, How quickly or slowly does $f(x)$ react when there is a slight perturbation or nudge in the input $x$? Does $f(x)$ react wildly or does it react mildly to these nudges in $x$? Does $f(x)$ even react at all to a nudge in $x$? This is a question about the sensitivity of $f(x)$ to changes in $x$, and ultimately to infinitely small changes in $x$.

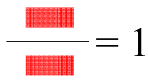
As an example consider the function $f(x) = x$. Let us now take any two points $x_1$ and $x_2$ on the $x$-axis and nudge these two values ever so slightly, by the same amount, either side of $x_1$ and $x_2$. The set up for this scenario is shown in the diagram below where the pink rectangles represent the range of nudging around the given points:



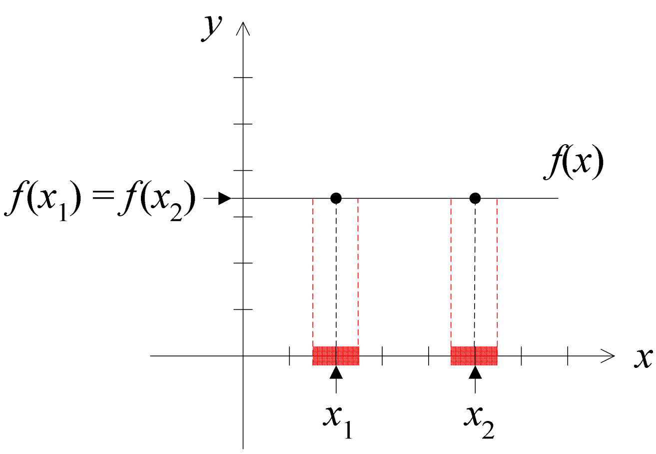

How will $f(x)$ react? Will $f(x)$ be nudged to the same extent? Will it be nudged to a greater or lesser extent, and if so by how much greater or less? Well, it can be seen from the diagram that, since the pink rectangles on the $y$-axis are of the same length as the pink rectangles on the $x$-axis, any nudging effect on $x_1$ and $x_2$ will cause $f(x)$ to react to exactly the same extent. Therefore, the nudging of $f(x)$ compared to the nudging of $x$ (i.e. the rate nudging) can be said to be 1 at both $x_1$ and $x_2$.

Comparing the nudging on $f(x)$ due to the nudging on $x$ at both $x_1$ and $x_2$ gives a measure of sensitivity of $f(x)$ at $x_1$ and $x_2$, and is another way of viewing the concept of the derivative of $f(x)$. Since $f(x)$ experiences the same extent of nudging as $x$ does over the whole of $f(x)$, the sensitivity (/derivative) of $f(x)$ is 1 for the whole of $f(x)$.

Representing visually the mathematical ratio of the vertical nudge to the horizontal nudge we could say that




What about if our function is a constant? In other words what if we have the situation below?

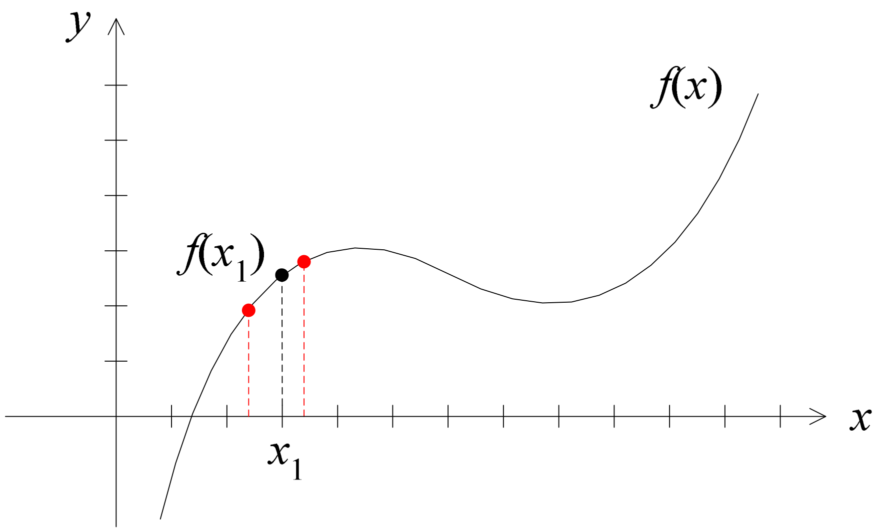


Here we see that whatever nudging occurs around $x_1$ and $x_2$ the function $f(x)$ does not react away from its $f(x_1)$ or $f(x_2)$ values. No amount of nudging, big or small, on the $x$ values will knock $f(x)$ off its perch. Therefore the nudging result on $f(x)$ compared to nudging effect due to $x$ is 0 at both $x_1$ and $x_2$. The derivative (which is the ratio of the $f(x)$ nudges to the $x$ nudges) is therefore 0.

Representing visually the mathematical ratio of the vertical nudge to the horizontal nudge we could say that

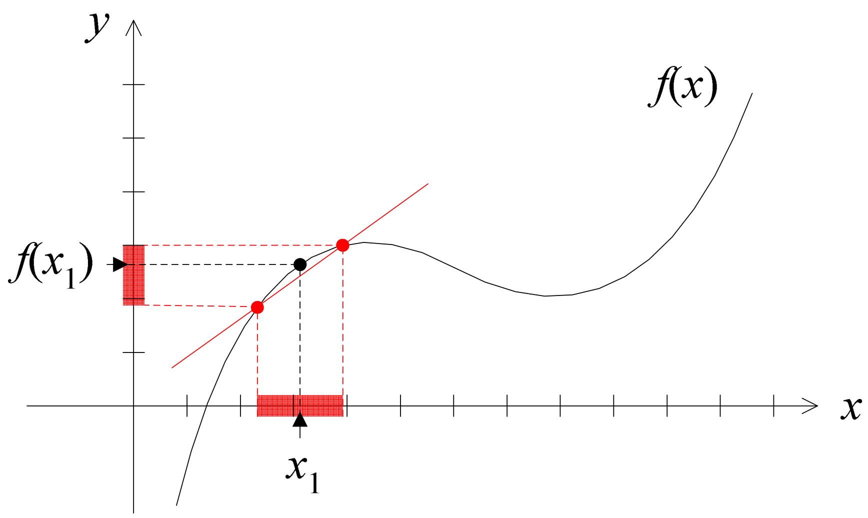
$$\frac{0}{\blacksquare} = 0$$

In general, $f(x)$ will have a curved path. Nudges of specific amounts around given values $x_1, x_2$, and $x_3$ will cause $f(x)$ to undergo a change in height. This can be seen in the diagram below when the black dot represents the height of the function for a given value $x_1$, and the red dots represents the changed heights of the function due to the amount by which $x_1$ has been nudged left and right:

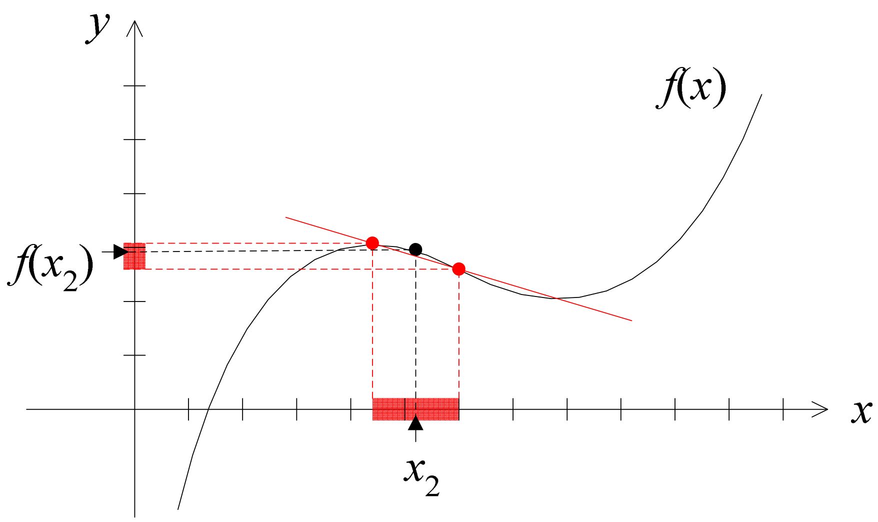



Now, $x_1$ has been nudge by the same amount left and right. But the change in height as the function moves from left red dot to black dot is not the same as the change in height from the black dot to the right red dot. As the height increase from left red dot to black dot to right red dot, so the change in height decreases. This implies that there is a rate of change of height due to the (equal amount of) nudging around $x$.

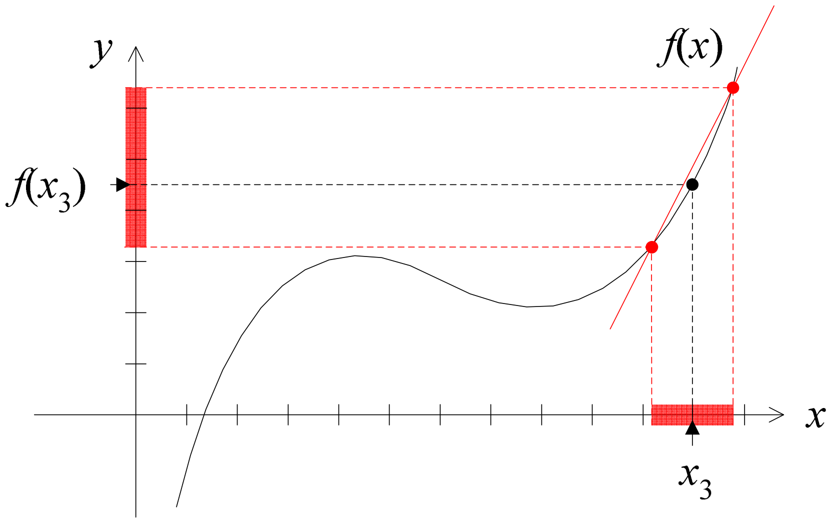
This nudging effect can be seen more easily if we draw little rectangles on the *x*-axis and *y*-axis as shown in the three diagrams below. The ratio of the range of nudge of $f(x)$ to the range of nudge of $x$ is the slope of the secant (illustrated by the solid red secant lines in the diagram).




*f(x) reacts to a medium extent for a given nudge around* $x_1$




*f(x) reacts to a small extent for a given nudge around* $x_2$

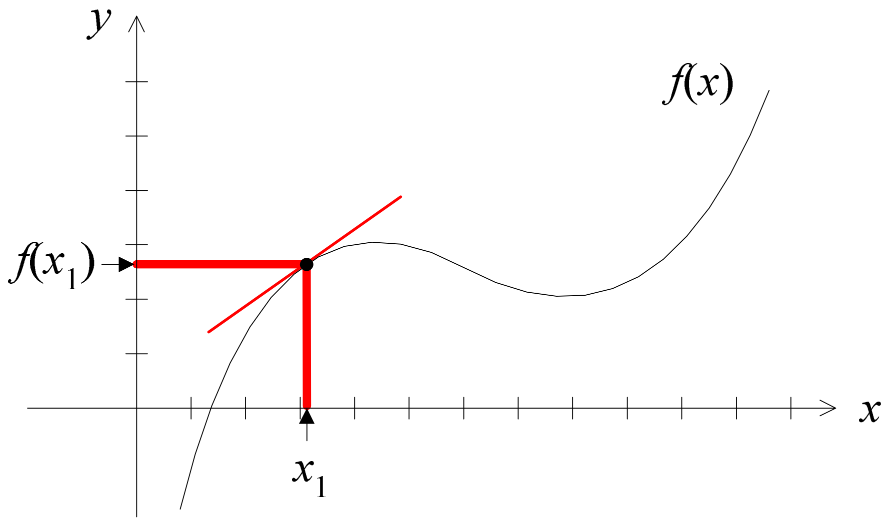


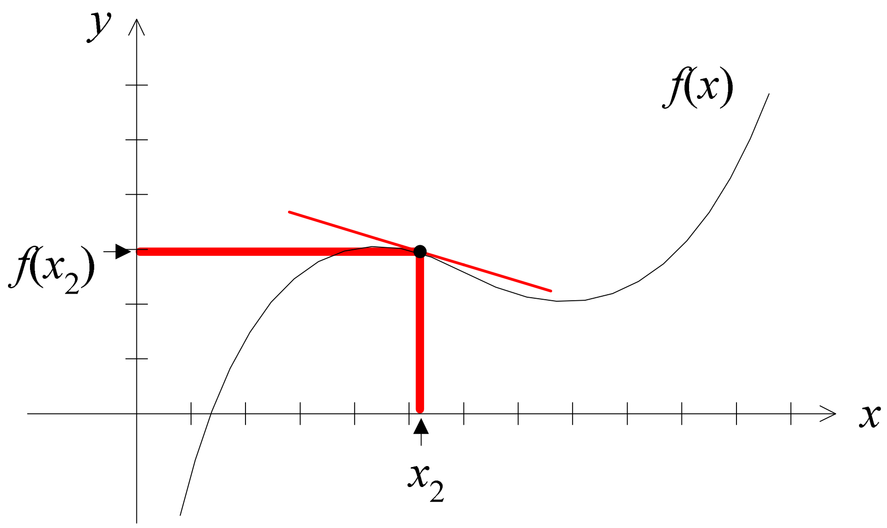

*$f(x)$ reacts to a large extent for a given nudge around $x_3$*

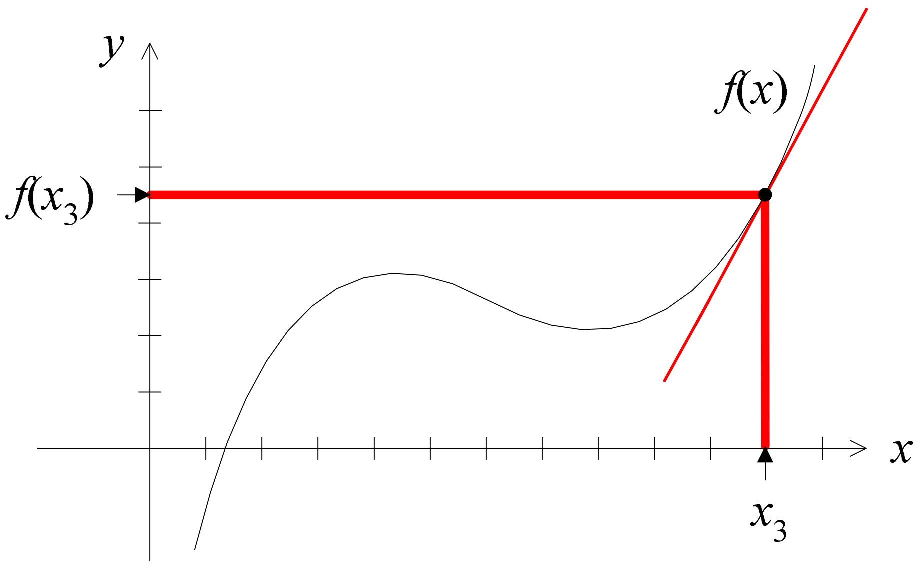
The fundamental nature of differentiation is such that, in order to study its effect, we need to consider the behaviour of points infinitely close to a given point *x*. We therefore need to consider the ratio of the nudged reaction of $f(x)$ due to nudging *x* when our pink rectangles become even more narrow than the regions depicted in the diagrams above.

So, as the nudging region around $x_1, x_2,$ and $x_3$ becomes more and more narrow, forever becoming more narrow, so the nudged region around points $f(x_1), f(x_2)$, and $f(x_3)$ also shrinks ever more. The ratio of these ever shrinking regions will lead us from a secant to a tangent, and towards the slope of $f(x)$ at $x_1, x_2$, and $x_3$ as being a measure of the sensitivity of $f(x)$ at $x_1, x_2$, and $x_3$.

In general, we can therefore interpret the derivative as describing the degree of sensitivity of a function $f(x)$ to infinitessimal nudge $\delta x$ around the input point *x*. The greater the sensitivity of $f(x)$ to nudges around a given point *x* the steeper the slope of $f(x)$ at that point *x*.


y
f(x)
f(x1)
x
x1


y
f(x)
f(x2)
x
x2


y
f(x)
f(x3)
x
x3

## 1.8 The derivative as a measure of distribution

In previous sections we have looked at the derivative from a geometric perspective. Here we will look at the derivative from a numerical perspective, and see how it can be interpreted as the rate at which the distribution of points changes as we approach a given value $y = f(x)$.

### *1.8.1 Functions as distributions of values, and how these distributions relate to rates of change*

Before we move onto studying rates of change of disgtributions we need to re-interpret how functions work, and what is is they really do. Therefore, consider the number line. The values on this line increase forever towards the right hand side and decrease forever towards the left hand side. What we will look at in this section applies to all real numbers (integers, rational and irrational numbers), but to make things easier let us look only at integer values:

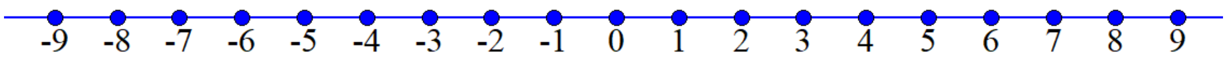



*Diagram 1: The distribution of integers on the number line*

This line shows us how the integers are distributed. In this case, integers can be seen to be evenly distributed and spaced 1 unit apart from each other. This number line is what is referred to as the $x$-axis on a graph.

One way of conceiving of functions is that

functions transform the distribution of values on the number line.

For example, the function $f(x) = x^2$ takes each value from the number line and transforms it into its squared number. Let us take –5 to 5 as a representative range, as seen by the blue dots in the diagram below:

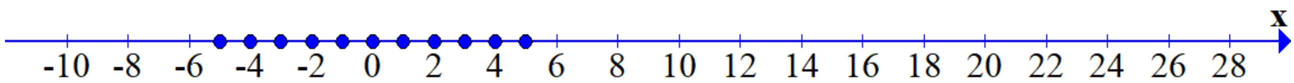



*Diagram 2: Dots showing a sample of values on the number line*

When we square these values they get transformed into those shown by the red dots in the "transformation" diagram below:

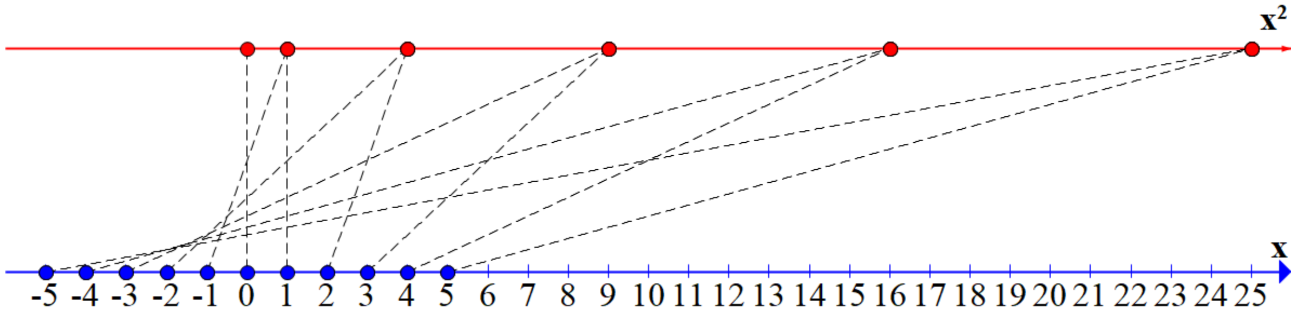



*Diagram 3: The transformation of x values to $x^2$ values*

The particular pattern shown by the distribution of red dots represents the squaring effect. Since the values of $x$ have changed from −5, −4, −3 ..., 0, ... 3, 4, 5 (and beyond) to the values of $x^2$ as 25, 16, 9, ..., 0, ... 9, 16, 25 (and beyond) there must be a rate at which these $x^2$ values have changed compared to the $x$ value. This has to be the case since the $x^2$ values are no longer evenly distributed when compared to the $x$ values but get further and further away from each other when $x < 1$ and $x > 1$.

Let us now return to the basic untransformed number line. The evenness of the distribution of values on this line means that any one integer on this line is the same distance away from its neighbour as any other integer. Let us now see what happens when we compare this distribution with other distributions given by some transformation $f(x)$.

The first transformation we will look at is $f(x) = 2$. Comparing the distribution of integers of the $x$-axis with the "distribution" of values of $f(x) = 2$ we see that when $x = 0$, $f(x) = 2$, when $x = 1$, $f(x) = 2$, when $x = 3$, $f(x) = 2$, etc. So there is no distribution in $f(x)$ values since all values of $x$ lead to the single non-distributed value 2. This is represented by the diagram below:

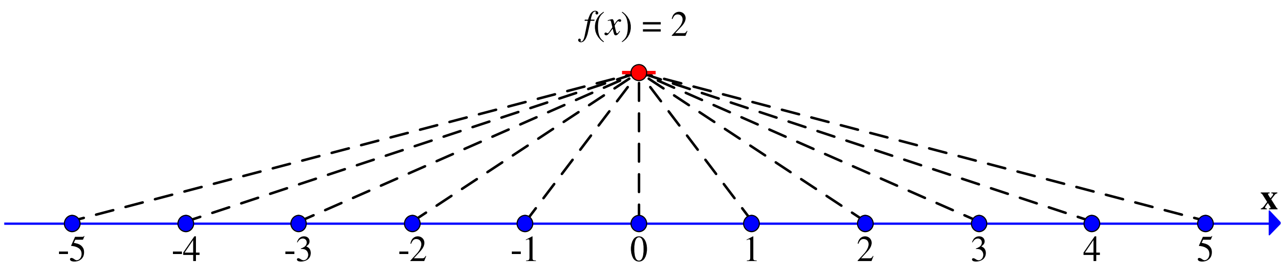



*Diagram 4: All distributed values of x lead to the non-distributed value f(x) = 2*

Since $f(x)$ has no distribution the rate of change of $f(x)$ compared to the values of $x$ is 0. Another way of saying this is that $f(x)$ is not keeping up at all with $x$. It is this which allows us to say that the derivative of $f(x)$ with respect to $x$ is 0.

The second transformation we will look at is $f(x) = x$. Comparing the distribution of integers on the $x$-axis with the distribution of values of $f(x) = x$ we see that when $x = 0$, $f(x) = 0$, when $x = 1$, $f(x) = 1$, when $x = 3$, $f(x) = 3$, etc. So there is a distribution in the values of $f(x)$ and this distribution is exactly the same as that of the untransformed $x$ values. This is represented by the diagram below:

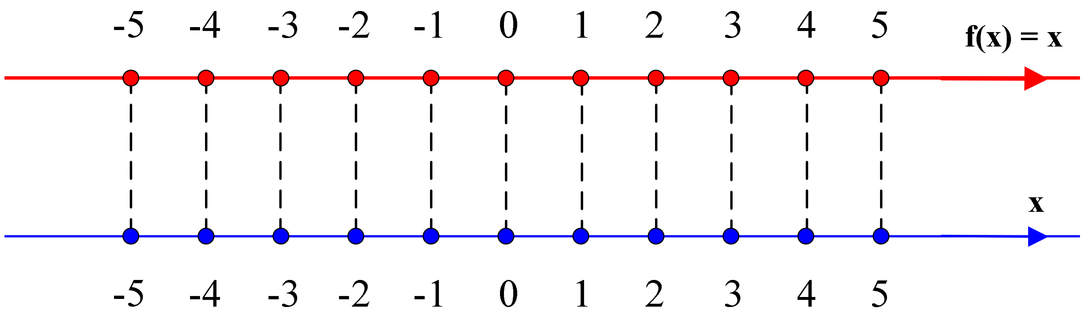



*Diagram 5: All distributed values of x lead to the distributed value f(x) = x*

Since the distribution of $f(x) = x$ is exactly the same as that of $x$ there is no change in the way the points of $f(x) = x$ are "internally" distributed compared to those of the $x$-axis. In other words, a point $c$ whose neighbouring points are a certain distance away from $c$ on the $x$-axis will have neighbouring points exactly the same distance away from c on the $f(x)$ axis. This can be interpreted as meaning that $f(x)$ is keeping up at exactly the same pace as $x$. It is this which allows us to say that the derivative of $f(x)$ with respect to $x$ is 1.

The third transformation we will look at is $f(x) = x + 2$. The only thing that has happened here compared to $f(x) = x$ is that every point of this former transformation has been shifted equally to the right by 2 points, as can be seen in the diagram below.

As can be seen, there is no change in the way the points of $f(x) = x + 2$ are "internally" distributed compared to the distribution of values on the $x$-axis. In other words, a point $c$ whose neighbouring points are a certain distance away from $c$ on the $x$-axis will have neighbouring points exactly the same distance away from c on the $f(x)$ axis, save for the fact that $c$ and it neighbouring point have all shifted to the right by 2 units.

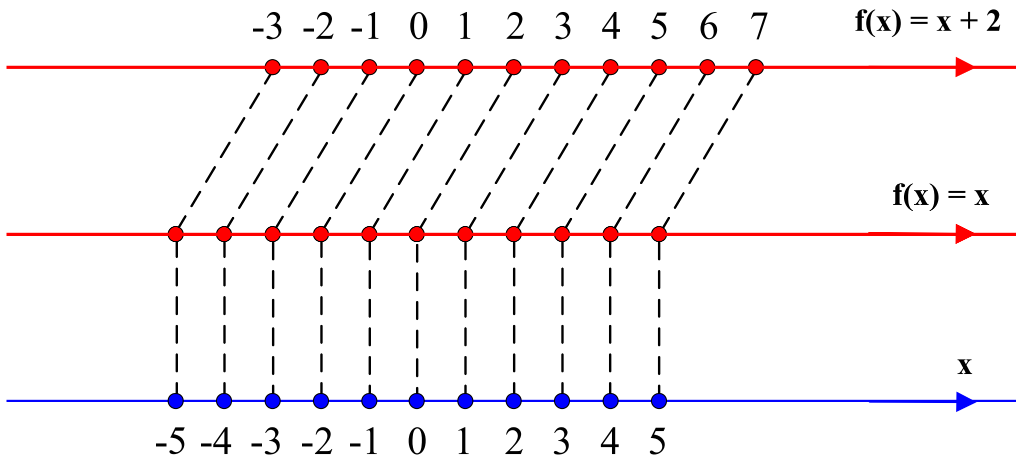



*Diagram 6: All distributed values of x lead to the distributed value f(x) = x + 2*

So the function $f(x) = x + 2$ is distributed in the same way as $f(x) = x$. The transformed distribution $f(x)$ will still be keeping with the untransformed distribution $x$ at exactly the same pace, hence the derivative of $f(x) = x + 2$ with respect to $x$ is also 1.

What if our transformation is now $f(x) = 2x$ as shown in the diagram below? Comparing the distribution of integers on the $x$-axis with the distribution of values of $f(x) = 2x$ we see that when $x = 0$, $f(x) = 0$, when $x = 1$, $f(x) = 2$, when $x = 2$, $f(x) = 4$, when $x = 3$, $f(x) = 6$, etc. What this means is that not only have the $x$ values shifted rightward and leftward (which we have seen by the previous example does not, in and of itself, affect the derivative) but the “internal” distribution of $f(x) = 2x$ values (i.e. the distribution of any two neighbouring values with respect to each other) compared to the “internal” distribution of $x$ values has changed.

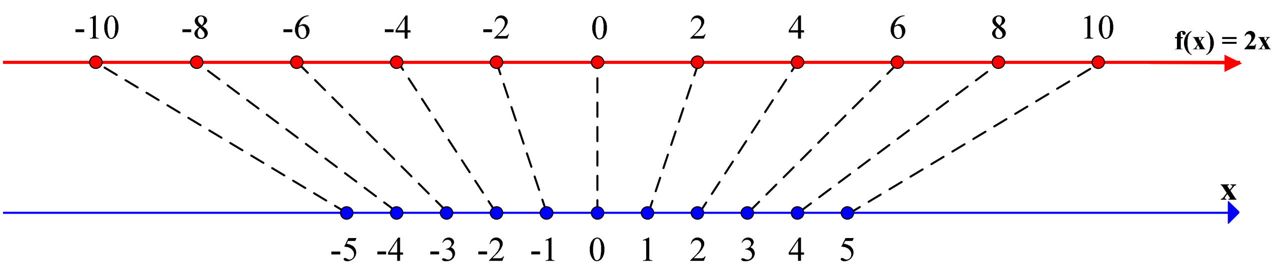



*Diagram 7: All distributed values of x lead to the distributed value f(x) = 2x*

The change in the spreading out or distribution of $f(x) = 2x$ values happens at twice the rate of that of the $x$ values. It also happens to be doing this consistently over the whole range of $2x$ values. So the pace of the transformed distribution $f(x)$ outstrips the pace of the untransformed distribution $x$ by a factor of 2 It is this which allows us to say that the derivative of $f(x)$ with respect to $x$ is 2.

So the concept of the derivative is seen to be the result of comparing the distribution of values given by a transformation $f(x)$ with the distribution of values of the $x$-axis. Although the $x$-axis values have a distribution, and therefore a rate of change, the $x$-axis is always taken as the reference distribution with which all other distributions (i.e. functions $f(x)$) are compared. And it is this comparison of the transformed distributions to the untransformed distribution which leads to the concept of the derivative.

*1.8.2 The rate of change of distribution in more detail*

The rate of change of distribution of transformed values was fairly straightforward to see for the previous examples because all values of $f(x) = x$ or $f(x) = 2x$ are evenly distributed, meaning that any one value is equally distant from its neighbours as any other value. Therefore, every single point in the distribution is running at the same pace as every other point in the distribution. Hence, the rate of change of $f(x)$ is exactly the same at every single point along the distribution.

But what if our transformation is $f(x) = x^2$ as shown in the diagram below? Here we have that (amongst other values) when $x = 0$, $f(x) = 0$, when $x = 1$, $f(x) = 1$, when $x = 2$, $f(x) = 4$, when $x = 3$, $f(x) = 9$, etc. So the distribution of $x^2$ values is not evenly spaced out compared to that of the standard $x$-axis values. In fact, when $x > 0$ values of $x^2$ get further and further away from their preceding values as $x$ increases.

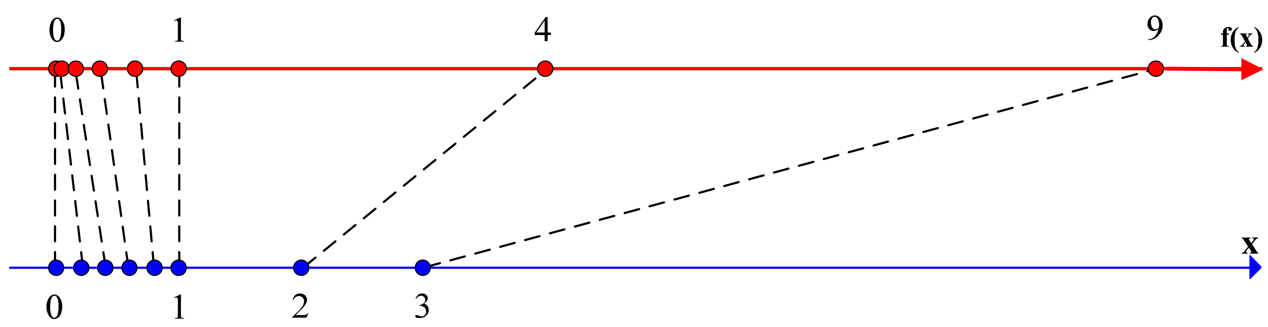



*Diagram 8: All distributed values of x lead to the distributed value f(x) = x²*

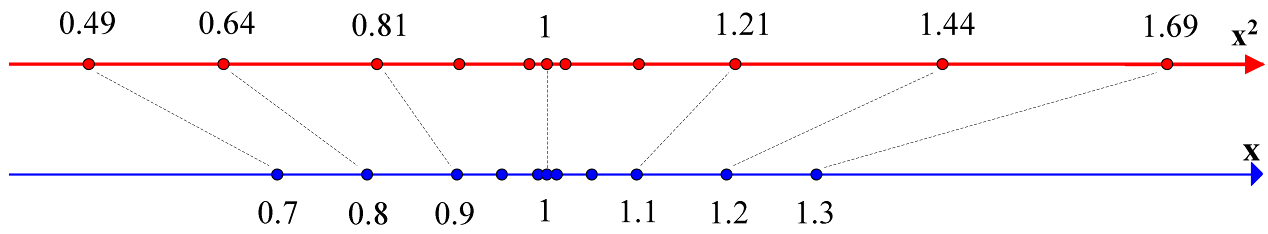
How then do we analyse the rate of change of the distribution given by $f(x)$? The question now is, At what rate are the $x^2$ values changing? In order to answer this we will have to look more closely at what is happening to the distribution of points around a specific value of $x^2$ compared to the distribution of points around the respective value of $x$.

Specifically, taking (1, 1) as the point of study, we will look at how different the pace of points around $x^2 = 1$ is, as we approach $x^2 = 1$, compared to the pace of points around $x = 1$ as we approach $x = 1$? More exactly, we will want to know the rate at which the $x^2$ values are changing around, or at, $x^2 = 1$ compared to the $x$ values around, or at, $x = 1$?

Therefore, consider $f(x) = x^2$ for some values in [0.7, 1.3], as shown in the table below:

| **x** | 0.7 | 0.8 | 0.9 | 0.95 | 0.99 | 1 | 1.01 | 1.05 | 1.1 | 1.2 | 1.3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{x^2}$ | 0.49 | 0.64 | 0.81 | 0.9025 | 0.9801 | 1 | 1.0201 | 1.1025 | 1.21 | 1.44 | 1.69 |

The transformation diagram for this is as below:




Here, for clarity, I have left out the values $x = 0.95, 0.99, 1.01$, and $1.05$, as well as $x^2 = 0.9025, 0.9801, 1.0201$, and $1.1025$ (it should be fairly straightforward to identify these in the diagram above).

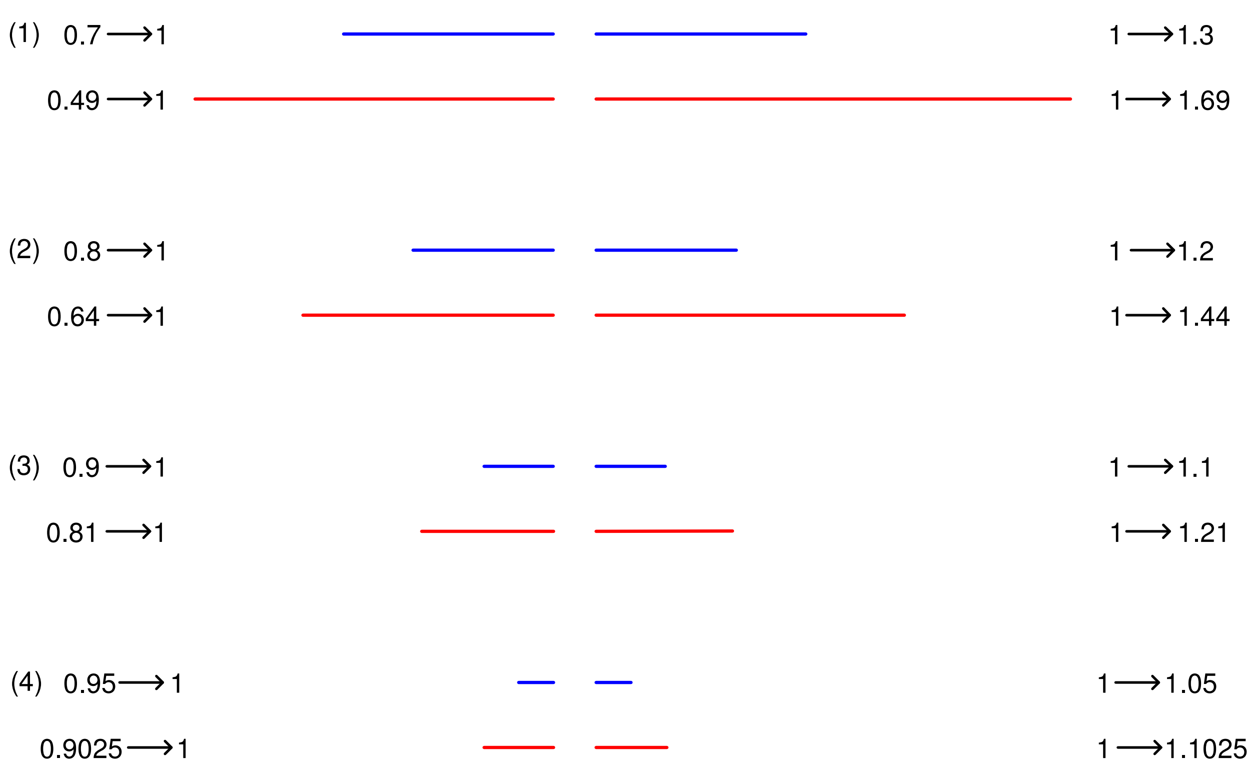
The question now is, At what rate are the $x^2$ values changing around, or at, $x^2 = 1$ compared to the $x$ values around, or at, $x = 1$? To answer this, notice that $x = 0.7$ and $x = 1.3$ are the same distance of 0.3 away from $x = 1$. However, the transformed version of these $x$ values, i.e. $x^2 = 0.49$ and $x^2 = 1.69$, are not the same distance away from $x^2 = 1$. But as $x$ approaches 1, and as $x^2$ approaches 1, values of $x^2$ become closer and closer to twice the distance away from $x^2 = 1$ compared to the distance of values away from $x = 1$.

This can be seen by the table shown below for the interval [0.7, 1.3], with the two previous examples being highlighted (blank cells to be explained shortly). In this table I have written $|x - 1|$ to represent the left and right distances of $x$ values away from $x = 1$, and $|x^2 - 1|$ to represent left and right distances of $x^2$ values away from $x^2 = 1$:

| $x$ | 0.7 | 0.8 | 0.9 | 0.95 | 0.99 | 1 | 1.01 | 1.05 | 1.1 | 1.2 | 1.3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lvert x-1\rvert$ | 0.3 | 0.2 | 0.1 | 0.05 | 0.01 | | 0.01 | 0.05 | 0.1 | 0.2 | 0.3 |
| $\lvert x^2-1\rvert$ | 0.51 | 0.36 | 0.19 | 0.0975 | 0.0199 | | 0.0201 | 0.1025 | 0.21 | 0.44 | 0.69 |
| $x^2$ | 0.49 | 0.64 | 0.81 | 0.9025 | 0.9801 | 1 | 1.0201 | 1.1025 | 1.21 | 1.44 | 1.69 |

Referring to the table above we see that on the left hand side of (1, 1), 0.51 $\not\approx$ 2×0.3, but as we get closer to the value (1, 1), 0.0199 ≈ 2×0.01. The same thing happens on the right hand side of (1, 1), so that as we get closer to the value (1, 1) the distance $|x^2 - 1|$ gets closer to two times the distance $|x - 1|$: 0.62 $\not\approx$ 2×0.3, but 0.0201 ≈ 2×0.01.

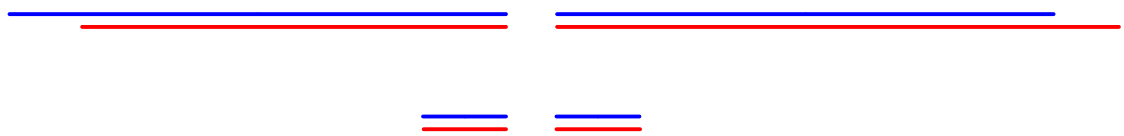
This idea can also be represented visually, for all left and right distances, as shown below:




In the diagram above, from both the left hand side and the right hand side, the blue lines are always of the same length for any given distance away from $x$ = 1. But the red lines do not have the same length for any given distance away from $x^2$ = 1.

The left and right red lines in (1) are significantly different in length, whereas the left and right red line in (4) are much closer in length, and much closer to being twice the length of the blue lines in (4). This "twice" is the derivative of $f(x)$ at $x = 1$.

What this all means in visual terms is that as the blue lines of the previous diagram get shorter and shorter so the red lines get closer and closer to being twice the length of these blue lines as we approach $x = 1$. This is illustrated in the diagram below by comparing (1) with (4) of the diagram above (with the blue lines in (1) and (4) being twice their original length):




We can see that in (1) the red lines are nowhere close to being the same length as the blue lines, but in (4) they are very close to being the same length as the blue lines.

Returing to our previous calculations of $|x^2 - 1|$ and $|x^2 - 1|$, the truth is that the calculations of $0.51 \not\approx 2\times0.3$, $0.0199 \approx 2\times0.01$, etc. previously presented are technically inaccurate. We should actually include the $x$ value of "1" into our calculations, so that we have $0.51 \not\approx 2\times1\times0.3$, $0.0199 \approx 2\times1\times0.01$, $0.62 \not\approx 2\times1\times0.3$, and $0.0201 \approx 2\times1\times0.01$. This is because it is in the nature of finding the derivative at a point that that point be included in the arithmetic calculation.

The reason as to why this has to be the case can only be seen when we do the proper mathematical analysis of differentiation, which we have done in sections ##. But given that we are looking at the derivative from a numerical-distribution perspective, we simply take as a heuristic the need to include the $x$ value into our calculation.

In general we can therefore say that, as $|x - 1|$ (the distance between $x$ and 1, or the length of the blue lines) and $|x^2 - 1|$ (the distance between $x^2$ and 1, or the length of the red lines) get smaller and smaller, we have that $|x^2 - 1| \approx 2|x - 1|$. Ultimately we have $|x^2 - 1| = 2|x - 1|$ when $x$ approaches 1. The value "2" is the derivative of $x^2$ with respect to the original (untransformed) $x$ values when $x = 1$.

In the above analysis we have actually performed a limiting process which, from our previous work, does not allows us to an perform an evaluation *at* $x = 1$. In other words, the idea of $|x^2 - 1| = 2|x - 1|$ only applies when $x \to 1$ and not when $x = 1$, and this is represented in the table above by leaving the cells blank.

We can repeat the analysis above for any value of $x$. In other words, the speed of distribution of values around $x^2 = c^2$ compared to the distribution values around $x = c$ will approach more and more a particular constant (which will be different for different values of $c$). In visual terms, and using the previous diagram as a guide, as $x$ approaches $c$ the red lines will approach a constant multiple of the blue lines given by $|x^2 - c^2| = 2c|x - c|$.

Examples for two other $x$ values are shown in the table of calculations below:

*For x = 0.7*

| **x** | 0.67 | 0.68 | 0.69 | 0.695 | 0.699 | 0.7 | 0.701 | 0.705 | 0.71 | 0.72 | 0.73 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **x²** | 0.4489 | 0.4624 | 0.4761 | 0.4830 | 0.4886 | 0.49 | 0.4914 | 0.4970 | 0.5041 | 0.5184 | 0.5329 |

| **\|x - 0.7\|** | 0.03 | 0.02 | 0.01 | 0.005 | 0.001 | | 0.001 | 0.005 | 0.01 | 0.02 | 0.03 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **\|x² - 0.49\|** | 0.0411 | 0.0276 | 0.0139 | 0.006975 | 0.001399 | | 0.001401 | 0.007025 | 0.0141 | 0.0284 | 0.0429 |

On the left hand side of (0.7, 0.49), for distances far away from (0.7, 0.49), we have $0.0411 \not\approx 2\times0.7\times0.03$, but as we get closer to the value (0.7, 0.49), $0.001399 \approx 2\times0.7\times0.001$. The same thing happens to the respective values on the right hand side of (0.7, 0.49). So, as $|x - 0.7|$ and $|x^2 - 0.49|$ get smaller and smaller we have that $|x^2 - 0.49| \approx 2 \times 0.7 \times |x - 0.7|$ as $x$ approaches 0.7. Again, the value "2" is the derivative of $x^2$ with respect to the original (untransformed) $x$ values when $x = 0.7$.

*For x = 4*

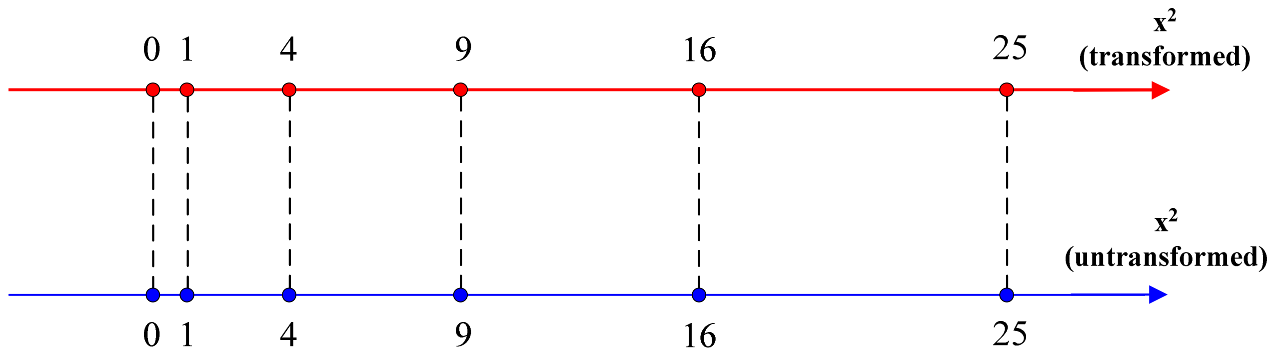
| **x** | 3.7 | 3.9 | 3.9 | 3.95 | 3.99 | 4 | 4.01 | 4.05 | 4.1 | 4.2 | 4.3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **x²** | 13.69 | 15.21 | 15.21 | 15.6025 | 15.9201 | 16 | 16.0801 | 16.4025 | 16.81 | 17.64 | 18.49 |

| **\|x - 4\|** | 0.3 | 0.1 | 0.1 | 0.05 | 0.01 | | 0.01 | 0.05 | 0.1 | 0.2 | 0.3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **\|x² - 16\|** | 2.31 | 0.79 | 0.79 | 0.3975 | 0.0799 | | 0.0801 | 0.4025 | 0.81 | 1.64 | 2.49 |

On the left hand side of (4, 16), for distances far away from (4, 16), we have 2.31 $\not\approx$ 2×4×0.3, but as we get closer to the value (4, 16), 0.0799 ≈ 2×4×0.01. The same thing happens to the respective values on the right hand side of (4, 16). So, as $|x-4|$ and $|x^2-16|$ get smaller and smaller we have that $|x^2-16| \approx 2\times 4\times |x-4|$ as $x$ approaches 4. Again, the value "2" is the derivative of $x^2$ with respect to the original (untransformed) $x$ values when $x$ = 4.

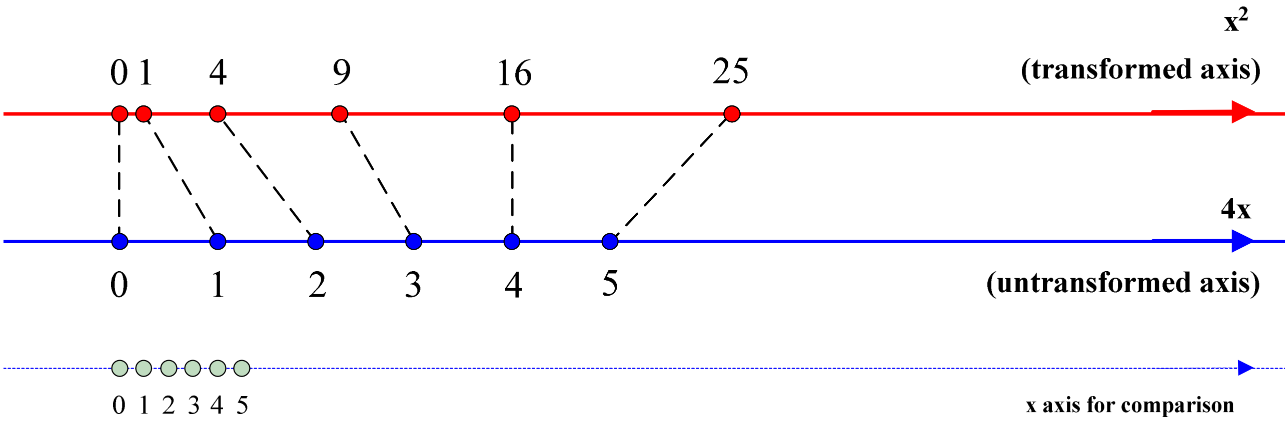
Note that the derivative of 2 at the above values of $x$ applies only when comparing the distribution of the transformed $x^2$ values to the distribution of untransformed $x$ value. If the distribution of $x^2$ values is compared with an alternative distribution of values (this acting as the original untransformed values, or alternative "$x$-axis") the derivative of $x^2$ with respect to this alternative untransformed distribution would be different.

For example, if our original untransformed distribution were $x^2$ itself then values along this "$x$-axis" will be distributed a la $x^2$, so the function $f(x)=x^2$ will not transform the "$x$-axis", since the "$x$-axis" already has the $x^2$ distribution pattern:




So the "transformed" $x^2$ values (the red dots on the red axis) are neither speeding up nor slowing down with respect to the untransformed $x^2$ values (the blue dots on the blue axis), but are actually keeping pace with the untransformed $x^2$ values. Hence the derivative of transformed $x^2$ values with respect to untransformed $x^2$ values is 1 over the whole range of transformed $x^2$ values

As another example, if our original untransformed distribution (i.e. our alternative $x$-axis) is $4x$ then values along this "$x$-axis" will be stretched out by a factor of 4: $x$ = 1 becomes new-$x$ = 4, $x$ = 2 becomes new-$x$ = 8 etc. The diagram below shows the untransformed 4$x$-axis for integer values in the interval [–4, 4], along with the usual transformation $f(x)=x^2$:




And as before we can ask, At what rate are the $x^2$ values changing around, or at, a given point $x^2 = c^2$ compared to the $4x$ values around, or at, $4x = c$? For this, and for more complicated distributions, we then resort to the maths of calculus (the above situation would be dealt with via something called the chain rule).

## 1.9 The second and third derivatives (to come)

### *1.9.1 The definition of the second derivative*

### *1.9.2 The second derivative as curvature*

### *1.9.3 The definition of the third derivative*

### *1.9.4 The third derivative as abberancy*

## 1.10 Equations involving the derivative in nature: Selected examples

In general the symbolism "$df/dx$" simply refers to a rate of change of $f(x)$ with respect to $x$. However, the use of derivatives allows us to represent physical processes, and in specific situations the derivative $df/dx$ will represent a physical process having a specific physical meaning.

Examples of equations involving derivatives which represent physical processes in nature are shown in Table 3 below. I have been very biased in the example I have chosen since the areas below are those of most interest to me. If your interest lies in other subjects then ask your teachers for examples of the use of derivatives in these subjects.

| | **Function** | **First derivative** | **Second derivative** |
|---|---|---|---|
| **Mathematics** | $y(x)$<br>Position w.r.t. $x$ | $dy/dx$<br>Slope/gradient<br>(steepness of a curve) | $d^2y/dx^2$<br>Curvature<br>(degree of bending of a curve) |
| **Mechanics** | $s(t)$<br>Position at any time $t$ | $ds/dt$<br>Speed<br>(rate of change of position) | $d^2s/dt^2$<br>Acceleration<br>(rate of change of speed) |
| **Electricity** | $q(t)$<br>Electric charge | $i = dq/dt$<br>Current<br>(rate of flow of charge) | – |
| **Electricity** | $V$(r)<br>Voltage at a given distance $r$ from an electric souce | $E = dV/dr$<br>Electric field intensity $E$, for an electric field of constant strength<br>(rate of change in voltage over a distance $r$) | – |
| **Chemistry** | $c_i(t)$<br>Concentration (molar) | $dc/dt$<br>Reaction<br>(rate of change of concentration) | – |
| **Nuclear decay** | $N(t)$<br>Number of atoms at any one time | $dN/dt$<br>Radioactive decay<br>(rate of change of the number of atoms) | – |

| | | | |
|---|---|---|---|
| **Wave motion** | $x(t)$<br>Displacement of a mass at any time $t$ from equilibrium position | $dx/dt$<br>Velocity of a mass at any time $t$ with respect to the mass's equilibrium position | $d^2x/dt^2 = -\lambda x$<br>Simple harmonic motion (Acceleration relating to current position) |
| **Mechanics** | $x(t)$<br>Position at any time $t$ | $m.dx/dt = mv$<br>Momentum | $m.d^2x/dt^2 = ma$<br>Acceleration relating to an applied force |

*Table 3: The physical meaning of certain derivatives*

### 1.11 The derivative of other basic functions from 1st principles

In section 1.3 we learnt that the limit of the difference quotient produces a function which represents the derivative of $y = f(x)$. In this section we will go through developing derivative functions relating to $y = f(x)$. Our starting point will always be the definition of the derivative given as equation (1), and reproduced below:

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

Finding derivatives based on this expression is called *differentiating from 1st principles*.

#### *1.11.1 The derivative of sin(x)*

We now come to finding the derivative of the three basic trigonometric functions $f(x) = \sin x$, $f(x) = \cos x$, and $f(x) = \tan x$. As usual we will do this from first principles.

Starting with $f(x) = \sin x$, by the definition of the first derivative we have

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

The first question is, Can we evaluate the limit as it stands? If there is then we have the answer to the derivative. Usually, however, we will not be able to evaluate the limit as it is given here (because it leads to a "0/0" situation which is not a valid answer).

We could consider forming two separate limits to the expression above (which is legal in maths) to get $df/dx = \lim_{\delta x \to 0}[\sin(x + \delta x)]/\delta x - \lim_{\delta x \to 0} \sin x /\delta x$ to see if this helps us get an answer. But, doing a table of values of both limits will show they approach infinity.

So this form of algebra, i.e. simply splitting the limit into two separate limits, doesn't work for us at this stage (but this doesn't mean that such an approach won't work later).

So we now have to consider doing something else. Usually this involves doing some simplifying algebra, specifically simplifying the expression inside the limit. In this case since our limit involves trig function we should always consider using trig identities to simplify where possible. As such we could use the identity $\sin(A-B) = \sin A \cos B - \sin B \cos A$, and this would indeed lead us to the derivative of $\sin x$. However, we will see that the identity

$$\sin P - \sin Q = 2 \sin \frac{P-Q}{2} \cos \frac{P+Q}{2}$$

will give us a quicker route to the derivative. Therefore we have

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{2 \sin\left(\frac{x+\delta x - x}{2}\right) \cos\left(\frac{x + \delta x + x}{2}\right)}{\delta x}.$$

Simplifying gives

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{2 \sin\left(\frac{\delta x}{2}\right) \cos\left(x + \frac{\delta x}{2}\right)}{\delta x}. \qquad (*)$$

Having done this algebraic transformation we again ask the question, Can we evaluate the limit? The answer this time is yes. But when we are learning something totally new we don't know this. So we need to go through explaining why the above limit is now evaluatable.

To see this let use groups terms of the previous expression in the very specific way shown below:

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{\sin\left(\frac{\delta x}{2}\right)}{\frac{\delta x}{2}} \times \lim_{\delta x \to 0} \cos\left(x + \frac{\delta x}{2}\right). \qquad (**)$$

I have done several things here. Firstly I have split the expression (*) into two limits. There is mathematical theory which allows us to do this. However, because this theory is beyond our current mathematical level this is one of those moments where we will have to accept that this can be done. We can, of course, test that such a split works by setting up a table of values for limit in (*) and compare with the table of values for the product of the limits in (**). At least this would give us numerical conformation of the validity of splitting the limits (this is left as an exercise).

Secondly, I have brought the "2" down to divide the $\delta x$ of the denominator. This is because this limit can be evaluated directly in this form, something we will see in a moment.

Thirdly, and most importantly, we split the fraction as $\sin(\delta x/2)/(\delta x/2)$ and $\cos(x + \delta x/2)$, not as $\sin(\delta x/2)$ and $\cos(x + \delta x/2)\,/(\delta x/2)$ or any other combination. Again, when we are learning this for the first time we won't know why we do it this way. But once we have gone through learning about certain limits we will be in a position to know which terms to group in (*) in order to give us limits we can evaluate.

So, back to (**). The first limit expression of $\lim_{\delta x \to 0} \sin(\delta x/2)/(\delta x/2)$ will give us the answer 1. To see this we can set up a table of values as shown below. Remember that for a limit to be valid it has to give us the same answer when we approach a given value from the left hand side (in this case $\delta x$ being very small negative numbers) as from the right hand side (in this case $\delta x$ being very small positive numbers).

| **$\delta$x** | ***y* = sin($\delta$x/2) / ($\delta$x/2)** |
|---|---|
| -0.2 | 0.998334166468282 |
| -0.02 | 0.999983333416666 |
| -0.002 | 0.999999833333342 |
| -0.0002 | 0.999999998333333 |
| … | … |
| 0 | ? |
| … | … |
| 0.0002 | 0.999999998333333 |
| 0.002 | 0.999999833333342 |
| 0.02 | 0.999983333416666 |
| 0.2 | 0.998334166468282 |

*Table of values* *$\delta$x approaches 0 from the left and the right*

It is important to note that this is only true when our angles are measured in radians. If this is the case, any limit of the form

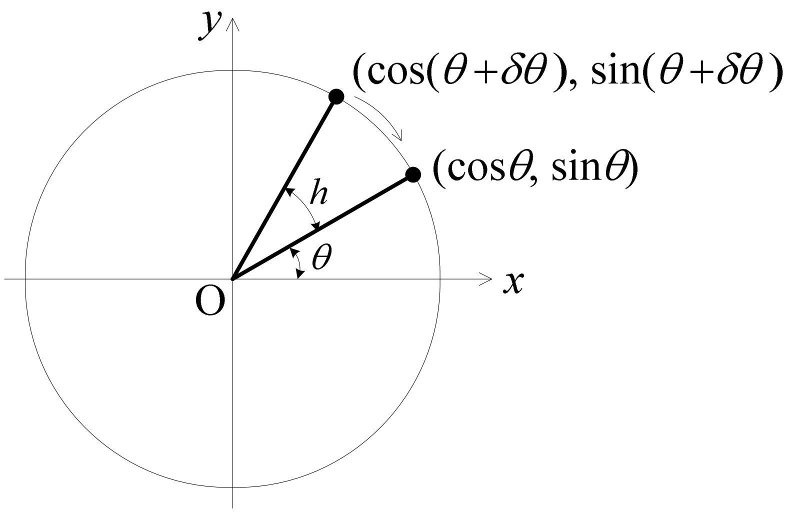
$$\lim_{\delta\theta \to 0} \frac{\sin\theta}{\theta} = 1$$

for any value of $\theta$. We just have to make sure that the denominator is exactly the same form as the variable used in the sine function.

We can prove this limit using geometry, and this will be shown in section 1.11.2 (an algebraic mathematical proof requires more advanced maths which we do not yet have so we will make do with the geometric proof to come).

Notice also that the limit equals 1 from below whether $\delta x$ approaches 0 from the left or the right. This is okay. Not all limits have to approach the limit value from above and below as was the case when defining the derivative $df/dx$. The key is that we get the same limit value when $\delta x$ approaches 0 from the left or the right.

The second limit expression of $\lim_{\delta x \to 0} \cos(x + \delta x/2)$ can be evaluated directly to give $\cos(x)$. This is not because we do $\delta x$ = 0. We can't do this since we are only ever letting $\delta x$ approach zero. So what is the difference between getting the correct answer $\cos(x + \delta x/2) = \cos x$ by incorrectly doing $\delta x$ = 0, and getting the correct answer of $\cos(x + \delta x/2) = \cos x$ when letting $\delta x$ approach 0? Well, essentially we want to see that $\cos(x + \delta x/2)$ approaches $\cos(x)$ as $\delta x$ gets smaller and smaller **so that when $\delta x$ approach 0, $\cos(x + \delta x/2) = \cos x$**, and we cannot take this for granted when dealing with limiting values. That this indeed happens can be seen visually in the diagram below, where the circle is of unit radius:




(in this case we also see that $\sin(x + \delta x/2) = \sin x$ as $\delta x$ approaches 0). If there is one thing to remember when dealing with infinitely small quantities such as $\delta x$ is it this:

the result of arithmetic or algebraic calculations involving arbitrarily small numbers is not obvious and can produce un-intuitive results.
Hence all terms involving infinitely small variables must be appropriately tested to see what results they give.

We have gone through a few table-of-values calculations to see that the limiting process does give (seemingly) unexpected results, so it seems possible that it is not obvious that $\cos(x + \delta x/2) = \cos(x)$ as $\delta x$ approaches 0.

Anyway, we ultimately have the following:

If $f(x) = \sin x$ then $\dfrac{df}{dx} = \cos x$ .

I previously mentioned that we could have used $\sin(A - B) = \sin A \cos B - \sin B \cos A$ as our simplifying trig identity. Let us go through what would have happened if we had done this:

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

$$= \lim_{\delta x \to 0} \frac{\sin x \cos \delta x + \sin \delta x \cos x - \sin x}{\delta x}. \qquad ②$$

We now need to decide how to group the terms of the fraction in ② so that we get limits we can evaluate. One thing we should be able to see is our previously seen $\lim\limits_{\delta x \to 0} \sin(\delta x)/\delta x$. So our grouping will be

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{\sin x \cos \delta x - \sin x}{\delta x} + \lim_{\delta x \to 0} \frac{\sin \delta x}{\delta x} \cos x . \qquad ③$$

Since the term $\cos x$ of the second limit is independent of $\delta x$ we can take this term out of the limit. Then applying the limit, i.e. letting $\delta x \to 0$, to $\sin(\delta x)/\delta x$ we have

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{\sin x \cos \delta x - \sin x}{\delta x} + \cos x . \qquad ⑤$$

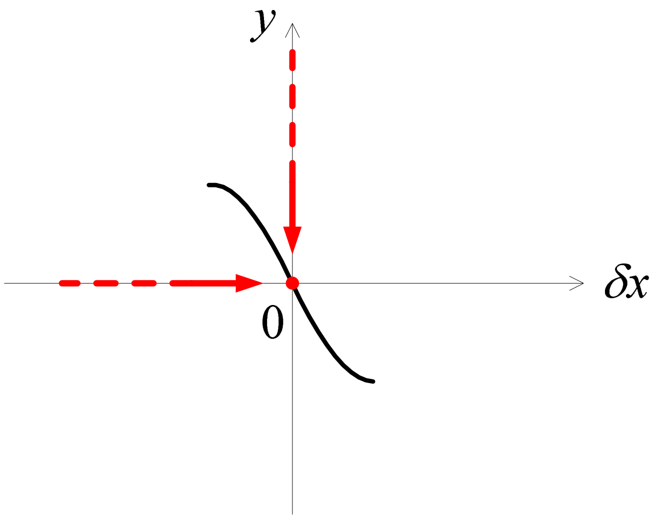
Looking at the limit in ⑤ we should see that the term $\sin x$ is independent of $\delta x$ so we can factorise this term out of the limit:

$$\frac{df}{dx} = \sin x \lim_{\delta x \to 0} \frac{\cos \delta x - 1}{\delta x} + \cos x . \qquad ⑥$$

We now have to decide if the limit in ⑥ is evaluatable. Looking at the table of values below we see that it is:

| $\delta x$ | $y = (\cos(\delta x) - 1)\,/\,\delta x$ |
|---|---|
| -0.1 | 0.049958347219742 |
| -0.01 | 0.004999958333474 |
| -0.001 | 0.000499999958326 |
| -0.0001 | 0.000049999999696 |
| … | … |

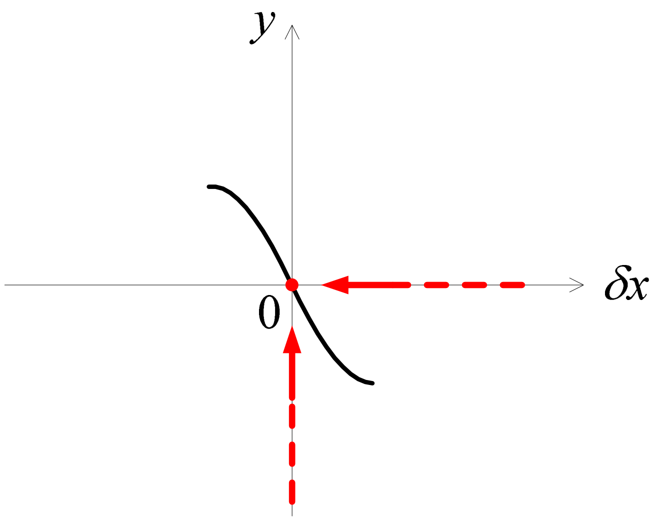
*Table of values*




*δx approaches 0 from the left*

| $\delta x$ | $y = (\cos(\delta x) - 1)\,/\,\delta x$ |
|---|---|
| … | … |
| 0.0001 | -0.000049999999696 |
| 0.001 | -0.000499999958326 |
| 0.01 | -0.004999958333474 |
| 0.1 | -0.049958347219742 |

*Table of values*




*δx approaches 0 from the right*

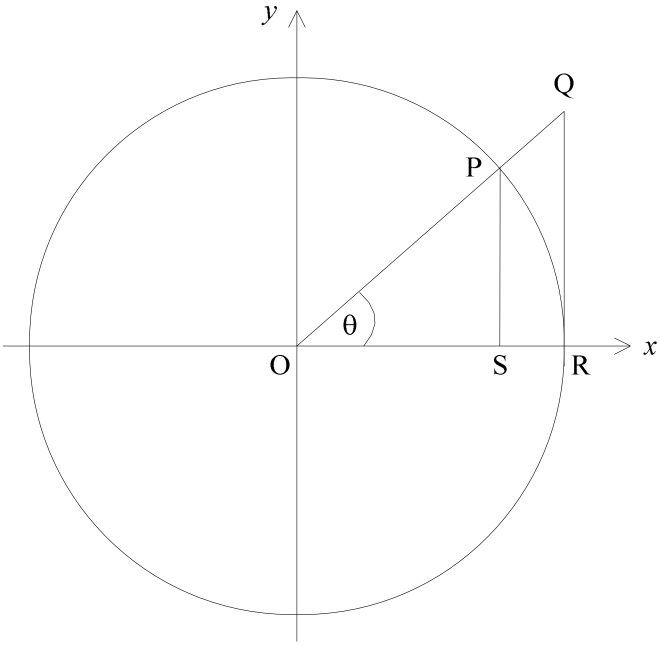
From the tables above we see that

$$\lim_{\delta x \to 0} \frac{\cos \delta x - 1}{\delta x} = 0$$

and expression ⑤ becomes $df/dx = \cos x$ as usual.

### *1.11.2 A geometric proof that $\lim_{\theta \to 0}(\sin \theta)/\theta = 1$*

Here we will go through proving the aforementioned limit. To do this we will use the diagram below which consists of a circle of radius OR = OP = $r$, an inner triangle OPS, a sector OPR and an outer triangle OQR.




Our aim will be to set up three different area formulas, one for the sector and one each of the two triangles. The significance of these area expressions is that they will allow us to get the answer we are looking for. How do you know this? Well, at the beginning, you don't. And, at the beginning, you're not supposed to know this. How can you? When this proof was originally developed the person or people who developed it did not initially know that this was the way to do it. They would have tried several different ideas and approaches before realising that one way to get to the answer was to start with he ideas of areas of geometric objects such as those in the diagram above.

However, this still doesn't help us to understand, at this moment in time, why we can use the idea of areas of the triangles and sector. So, one way around this "why do we do it this way?" gap is to simply go through the proof to make sure you understand the steps involved, after which you can go back to the first step knowing now how it is going to be used to give us the answer we now know.

To set up expressions for the areas we will have to use basic trig, and remember the area of a sector as $\frac{1}{2}r^2\theta$. So, the three areas are

$$\frac{1}{2}(OS)(PS) = \frac{1}{2}(r.\cos\theta)(r.\sin\theta)\,, \qquad \frac{1}{2}r^2\theta\,, \qquad \frac{1}{2}(OR)(RQ) = \frac{1}{2}(r)(r.\tan\theta)\,.$$

*Inner triangle OPS,* *Sector OPR,* *Outer triangle OQR.*

Notice that the area of the inner triangle OPS is smaller than the area of the sector OPR which is smaller than the area of the outer triangle OQR. Therefore, in terms of areas this can be represented pictorially as shown here:

$$\leq \qquad\qquad \leq$$

Hence

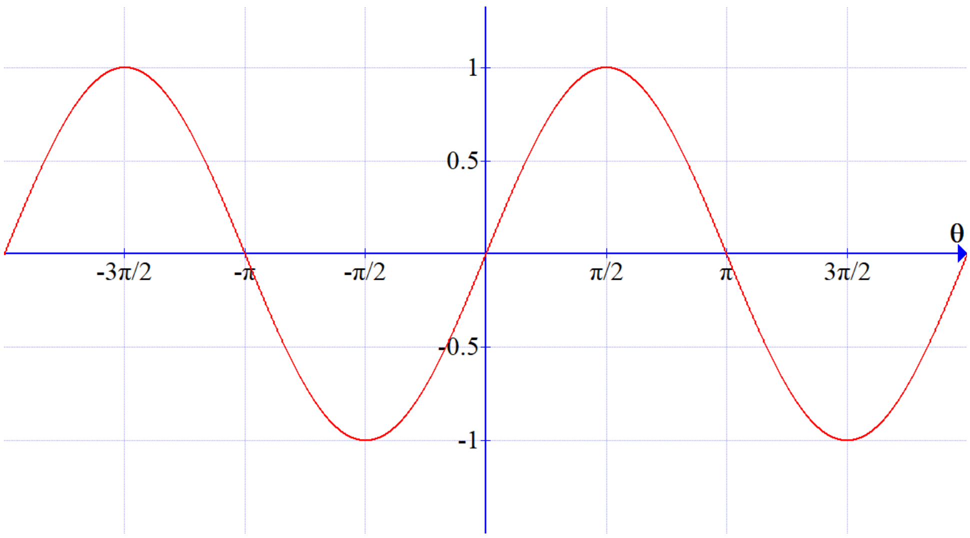
$$\frac{1}{2}(r.\cos\theta)(r.\sin\theta) \leq \frac{1}{2}r^2\theta \leq \frac{1}{2}(r)(r.\tan\theta)\ .$$

Simplifying we get

$$\cos\theta\,.\sin\theta \leq \theta \leq \tan\theta\ .$$

As discussed previously we can ask the question, Can we evaluate this inequality as $\theta$ approaches 0? No, for two reasons. Firstly we don't have the expression $(\sin\theta)/\theta$ so there is no point in evaluating the limit of this inequality. Secondly, even if we do evaluate it we get $0 \leq 0 \leq 0$ which is true but doesn't help us.

So, at this point we need to do some algebra in order to try to obtain a $(\sin\theta)/\theta$ expression. The closest we can get to this if we divide this inequality by $\sin\theta$. But we now have to be careful: is $\sin\theta$ positive or negative? Remember that if we divide an inequality by a negative value the direction of the signs change. So remembering that $\sin\theta$ looks like this …




… we see that it is non-negative (i.e. positive or zero) only in the interval [0, π]. From the original circle diagram above this means that our proof will (for the time being) only be valid when $\theta$ rotates anticlockwsie from 0 to π radians. To make our proof valid for all $\theta$ we will have to deal with the case of negative $\theta$. This we will do in a moment.

So assuming for the moment that $\theta \in [0, \pi]$ we can divide the above inequality by $\sin\theta$ to get

$$\cos\theta \leq \frac{\theta}{\sin\theta} \leq \frac{1}{\cos\theta}.$$

Having done some algebra to transform the previous inequality into this one above we can again ask the question, Can we evaluate this inequality as $\theta$ approaches 0? Yes. But before we do this we notice that we are nearly close to forming the expression $(\sin\theta)/\theta$ that we need. All we have to do is take the reciprocal of the inequality and we have our $(\sin\theta)/\theta$:

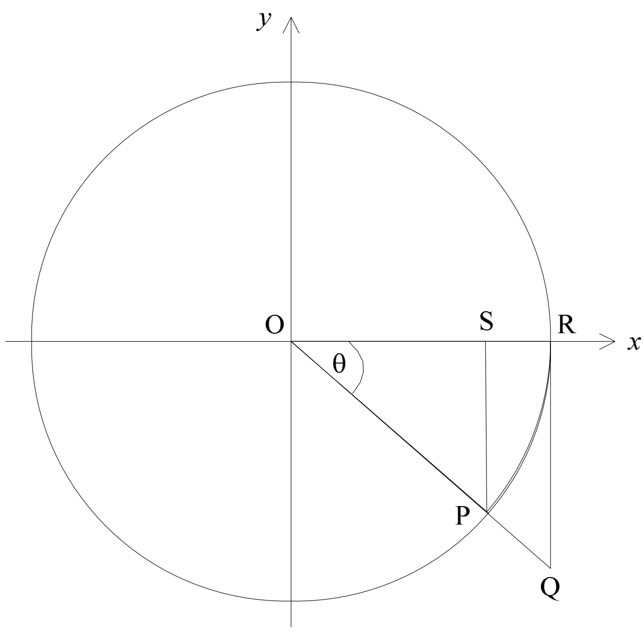
$$\frac{1}{\cos\theta} \geq \frac{\sin\theta}{\theta} \geq \cos\theta\ ,$$

remembering to change the direction of the inequality signs (why?). Now we can take limits:

$$\lim_{\theta\to 0}\frac{1}{\cos\theta} \geq \lim_{\theta\to 0}\frac{\sin\theta}{\theta} \geq \lim_{\theta\to 0}\cos\theta. \qquad (*)$$

The limit on the left hand side is 1, and the limit on the right hand side is also 1. What does this imply about the limit of $(\sin\theta)/\theta$? Well, this also has to be 1. This approach to proving a limit, by squeezing in onto our desired expression from the left and the right is a mathematically valid way of proving things.

There still remains one matter: what if $\theta$ is negative? What if $\theta$ lies in $(\pi, 2\pi)$? We could repeat the above analysis using the diagram below,




but a quicker way is to substitute $-\theta$ into (*):

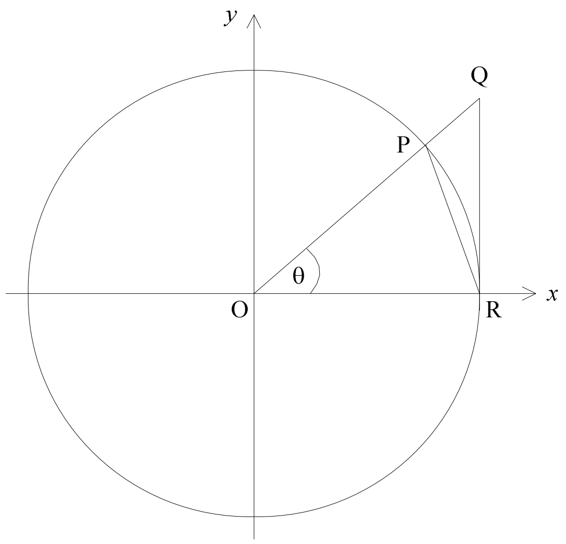
$$\lim_{\theta\to 0}\frac{1}{\cos(-\theta)} \geq \lim_{\theta\to 0}\frac{\sin(-\theta)}{-\theta} \geq \lim_{\theta\to 0}\cos(-\theta).$$

This simplifies to the same expression as (*). So it doesn't matter if $\theta$ is positive or negative, we get the same result, i.e.

$$\lim_{\theta\to 0}\frac{\sin\theta}{\theta} = 1 \quad \text{for } \theta \in [0, 2\pi].$$

So we have the amazing fact that as $\theta$ approaches 0 the area of the triangles and the area of the sector individually become 0, but the ratio of the arc length to vertical length of the inner triangle equals 1. This is the hidden feature of the geometric construction shown above, and many other geometries elsewhere in mathematics.

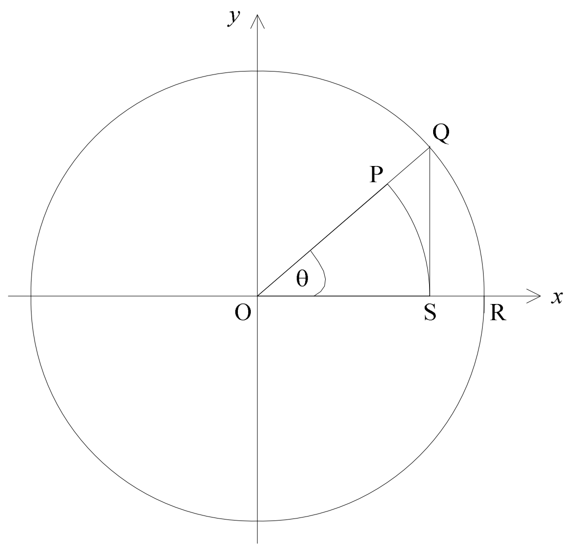
Note that in deriving the proof of the above limit we could also have used other geometric configurations such as either of those in the following two diagrams (or any other diagram which would allow us to form three area formulas from which to get the expression $(\sin\delta\theta)/\delta\theta$):




Area of triangle OPR

≤ area of sector OPR

≤ area of triangle OQR




Area of sector OPS

≤ area of triangle OQS

≤ area of sector OQR

I leave it as an exercise for you to develop the correct inequalities for both of these geometric set-ups, and to simplify these to prove the limit.

### *1.11.3 The derivative of cos(x)*

Finding the derivative of $f(x) = \cos x$ from first principles follows exactly the same approach as that for finding the derivative of $f(x) = \sin x$. We will again use first principles as our first step. Since we won't be able to evaluate this limit as it stands we will use an appropriate trig identity in order to simplify the fraction. We will then decide if we can evaluate this simplified limit (which in this case we will be able to).

So, from first principles we have

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

As ever we ask, Can the limit be evaluated as it stands? No. Technically we can evaluate the limit for any specific value of *x*, say *x* = π/3 and get an answer to the derivative at that point. But we don't know if this limit can be evaluated for all values of *x* (i.e. $-\infty < x < \infty$). Anyway, we want the general function expression for the derivative, not just the slope at one point. In this case we do not know how to evaluate the limit above as it stands.

As we did in the previous section when learning to differentiate sin*x* we could split the above expression into two separate limits to get $df/dx = \lim\limits_{\delta x \to 0} [\cos(x + \delta x)]/\delta x - \lim\limits_{\delta x \to 0} \cos x / \delta x$, but a table of values would show that both limits approach infinity. So, again, this form of algebra, i.e. simply splitting the limit into two separate limits, doesn't work for us at this stage.

So finally we decide to do some algebra to simplify the above expression in order to get a limit (or limits) we know how to evaluate. As in the previous section, so here we will use a trig identity to break the numerator apart. And, again, we will use one from the same family of trig identities, namely :

$$\cos P - \cos Q = -2 \sin \frac{P + Q}{2} \sin \frac{P - Q}{2}.$$

So our first principles expression becomes

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{-2 \sin\left(x + \frac{\delta x}{2}\right) \sin\left(\frac{\delta x}{2}\right)}{\delta x}.$$

Again we ask, Can the limit be evaluated as it stands? Based on our experience of limits from the previous section we should see that the answer is yes. We can see this by suitably re-grouping terms in this last expression in such a way as to form limits we can evaluate.

Therefore we write

$$\frac{df}{dx} = -\lim_{\delta x \to 0} \sin\left(x + \frac{\delta x}{2}\right) \times \lim_{\delta x \to 0} \frac{\sin\left(\frac{\delta x}{2}\right)}{\frac{\delta x}{2}} .$$

From our work in the previous section we know that the last limit evaluates to 1 (remembering that this applies only when our angles are measured in radians), and that the first limit simply gives us $\cos x$. So ultimately we have the following:

| | | |
|---|---|---|
| If $f(x) = \cos x$ | then | $\frac{df}{dx} = -\sin x$ . |

*1.11.4 The derivative of tan x*

Here we will see how to differentiate $f(x) = \tan x$ from first principles. So we have

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

At this point we might be tempted to use a trig identity (there are two that could be used), but a simpler way is to simply recast $\tan x$ as $\sin x \, / \cos x$, thus giving us

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

Cross-multiplying in the numerator we get

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

At this point we could use either of the trig identities $\sin(A - B) = \sin A \cos B - \sin B \cos A$ or $\sin P - \sin Q = 2\sin[(P - Q)/2] . \cos[(P + Q)/2]$. We will look at both to see what happens. The use of the former identity is shown as *version A* on the left hand side of the two solutions below, and the use of the latter identity is shown as *version B* on the right hand side of the two solutions below:

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{\sin\big((x+\delta x)-x\big)}{\delta x \cos(x+\delta x)\cos x} \,,$$

$$= \lim_{\delta x \to 0} \frac{\sin \delta x}{\delta x} \times \lim_{\delta x \to 0} \frac{1}{\cos(x+\delta x)\cos x} \,,$$

$$= \lim_{\delta x \to 0} \frac{1}{\cos(x+\delta x)\cos x} \,.$$

*Version A*

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{[\sin(2x+\delta x)+\sin \delta x]/2}{\delta x \cos(x+\delta x)\cos x} - \lim_{\delta x \to 0} \frac{[\sin(2x+\delta x)-\sin \delta x]/2}{\delta x \cos(x+\delta x)\cos x} \,,$$

$$= \lim_{\delta x \to 0} \frac{\sin \delta x}{\delta x \cos(x+\delta x)\cos x} \,,$$

$$= \lim_{\delta x \to 0} \frac{\sin \delta x}{\delta x} \times \lim_{\delta x \to 0} \frac{1}{\cos(x+\delta x)\cos x} \,,$$

$$= \lim_{\delta x \to 0} \frac{1}{\cos(x+\delta x)\cos x} \,.$$

*Version B*

In either case we are left with evaluating the final limit. From previous work we know that $\cos(x+\delta x) = \cos x$ as $\delta x$ approaches 0. Hence

$$\frac{df}{dx} = \frac{1}{\cos x \,.\, \cos x} = \frac{1}{\cos^2 x} = \sec^2 x \,.$$

So we have the following:

If $f(x) = \tan x$ then $\dfrac{df}{dx} = \sec^2 x$ .

*1.11.5 The derivative of sec x, cosec x, and cot x*

In using first principles to prove the derivative of $\sec x$ , $\operatorname{cosec} x$, and $\cot x$ we will use eactly the same thinking as we have done previously. Here I will present only the proofs of the derivatives of these. I leave it as an exercise for you to recognise why the steps have been done as they have.

For $f(x) = \sec x$, and using first principles, we have

$$\begin{aligned}
\frac{df}{dx} &= \lim_{\delta x \to 0} \frac{\sec(x+\delta x) - \sec x}{\delta x}, \\
&= \lim_{\delta x \to 0} \frac{\dfrac{1}{\cos(x+\delta x)} - \dfrac{1}{\cos x}}{\delta x}, \\
&= \lim_{\delta x \to 0} \frac{1}{\delta x} \left\{ \frac{\cos x - \cos(x+\delta x)}{\cos x \cos(x+\delta x)} \right\}, \\
&= \lim_{\delta x \to 0} \frac{1}{\delta x} \left\{ \frac{-2\sin\left(x + \frac{\delta x}{2}\right)\sin\left(-\frac{\delta x}{2}\right)}{\cos x \cos(x+\delta x)} \right\}, \\
&= \frac{1}{\cos x} \lim_{\delta x \to 0} \frac{1}{\delta x} \left\{ \frac{-2\sin\left(x + \frac{\delta x}{2}\right)\sin\left(-\frac{\delta x}{2}\right)}{\cos(x+\delta x)} \right\}, \\
&= \frac{1}{\cos x} \lim_{\delta x \to 0} \left\{ \frac{\sin\left(\frac{\delta x}{2}\right)}{\delta x/2} \right\} \lim_{\delta x \to 0} \left\{ \frac{\sin\left(x + \frac{\delta x}{2}\right)}{\cos(x+\delta x)} \right\}, \\
&= \frac{1}{\cos x} . \frac{\sin x}{\cos x},
\end{aligned}$$

as $\delta x \to 0$. So we have that

| | | |
|---|---|---|
| if $f(x) = \sec x$ | then | $\dfrac{df}{dx} = \sec x \tan x$ . |

For $f(x) = \operatorname{cosec} x$, and using first principles, we have

$$
\begin{aligned}
\frac{df}{dx} &= \lim_{\delta x \to 0} \frac{\operatorname{cosec}(x+\delta x) - \operatorname{cosec} x}{\delta x}, \\
&= \lim_{\delta x \to 0} \frac{\dfrac{1}{\sin(x+\delta x)} - \dfrac{1}{\sin x}}{\delta x}, \\
&= \lim_{\delta x \to 0} \frac{1}{\delta x} \left\{ \frac{\sin x - \sin(x+\delta x)}{\sin x \sin(x+\delta x)} \right\}, \\
&= \lim_{\delta x \to 0} \frac{1}{\delta x} \left\{ \frac{2 \sin\left(-\frac{\delta x}{2}\right) \cos\left(x + \frac{\delta x}{2}\right)}{\sin x \sin(x+\delta x)} \right\}, \\
&= -\frac{1}{\sin x} \lim_{\delta x \to 0} \frac{1}{\delta x} \left\{ \frac{2 \sin\left(\frac{\delta x}{2}\right) \cos\left(x + \frac{\delta x}{2}\right)}{\sin(x+\delta x)} \right\}, \\
&= -\frac{1}{\sin x} \lim_{\delta x \to 0} \left\{ \frac{\sin\left(\frac{\delta x}{2}\right)}{\delta x/2} \right\} \lim_{\delta x \to 0} \left\{ \frac{\cos\left(x + \frac{\delta x}{2}\right)}{\sin(x+\delta x)} \right\}, \\
&= -\frac{1}{\sin x} . \frac{\cos x}{\sin x},
\end{aligned}
$$

as $\delta x \to 0$. So we have that

| | | |
|---|---|---|
| if $f(x) = \operatorname{cosec} x$ | then | $\dfrac{df}{dx} = -\operatorname{cosec} x \cot x$ . |

For $f(x) = \cot x$, and using first principles, we have

$$
\begin{aligned}
\frac{df}{dx} &= \lim_{\delta x \to 0} \frac{\cot(x+\delta x) - \cot x}{\delta x}, \\
&= \lim_{\delta x \to 0} \frac{\dfrac{\cos(x+\delta x)}{\sin(x+\delta x)} - \dfrac{\cos x}{\sin x}}{\delta x}, \\
&= \lim_{\delta x \to 0} \frac{1}{\delta x}\left\{\frac{\sin x \cos(x+\delta x) - \cos x \sin(x+\delta x)}{\sin x \sin(x+\delta x)}\right\}, \\
&= \lim_{\delta x \to 0} \frac{1}{\delta x}\left\{\frac{\sin\big(x-(x+\delta x)\big)}{\sin x \sin(x+\delta x)}\right\}, \\
&= -\frac{1}{\sin x}\lim_{\delta x \to 0} \frac{1}{\delta x}\left\{\frac{\sin(\delta x)}{\sin(x+\delta x)}\right\}, \\
&= -\frac{1}{\sin x}.\frac{1}{\sin x},
\end{aligned}
$$

as $\delta x \to 0$. So we have that

| | | |
|---|---|---|
| if $f(x) = \cot x$ | then | $\dfrac{df}{dx} = -\text{cosec}^2\, x$ . |

For the last four derivative of $\tan x$, $\sec x$, $\text{cosec}\, x$ and $\cot x$ there are quicker way to differentiate these function by using the fact that $\tan x = \sin x \,/\cos x$, $\sec x = 1/\cos x$, etc. However, this requires the use one of the differentiation rules called the quotient rule. We will get to this rule, as well as differentiating the aformentioned functions accordingly, in the Differentiation II notes.

*1.11.6 The derivative of $a^x$*

We have seen how to differentiate $f(x) = x^n$ where $x$ is the variable and $n$ is a real number. What about differentiation $f(x) = n^x$ ? In this section we will see how to do this. For the sake of convention, and because most books show it this way, we will look at the function as $f(x) = a^x$ where $a$ is a real number.

So, by the definition of the derivative we have

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{a^{x+\delta x} - a^x}{\delta x},$$

$$= \lim_{\delta x \to 0} \frac{a^x(a^{\delta x} - 1)}{\delta x}.$$

Since $a^x$ is independent of $\delta x$ we take take this out of the limit to get

$$\frac{df}{dx} = a^x \lim_{\delta x \to 0} \frac{a^{\delta x} - 1}{\delta x}. \qquad [*]$$

Now all we have to do is evaluate the limit. If the limit is evaluatable (note that not all limit are evaluatable) then it will be some some value, $k$, say. So we can say that we have found the derivative of $f(x) = a^x$ to be $df/dx = k.a^x$ where $k$ is the answer to the limit shown in [*], and changes depending on the value of $a$. In other words, $k$ will be different when differentiating $2^x$ than when differentiating $3^x$ than when differentiating $4^x$, etc.

Let us now evaluate the limit for a few values of $a$:

**a = 1**

| $\delta x$ | $(a^{\delta x} - 1)/\delta x$ |
|---|---|
| -0.1 | 0 |
| -0.01 | 0 |
| -0.001 | 0 |
| -0.0001 | 0 |
| -0.00001 | 0 |
| … | … |
| 0 | 0 |
| … | … |
| 0.0001 | 0 |
| 0.0001 | 0 |
| 0.001 | 0 |
| 0.01 | 0 |
| 0.1 | 0 |

**a = 2**

| $\delta x$ | $(a^{\delta x} - 1)/\delta x$ |
|---|---|
| -0.1 | 0.66967008463193 |
| -0.01 | 0.690750456296496 |
| -0.001 | 0.692907009548049 |
| -0.0001 | 0.693123158469477 |
| -0.00001 | 0.693144778396437 |
| … | … |
| 0 | 0.693147180559945… |
| … | … |
| 0.00001 | 0.693149581998398 |
| 0.0001 | 0.693171203700604 |
| 0.001 | 0.693387462580075 |
| 0.01 | 0.695555005671 |
| 0.1 | 0.7177346253629 |

**a = 3**

| $\delta x$ | $(a^{\delta x} - 1)/\delta x$ |
|---|---|
| -0.1 | 1.04041540159238 |
| -0.01 | 1.092599582783 |
| -0.001 | 1.09800903512203 |
| -0.0001 | 1.09855194343034 |
| -0.00001 | 1.09860625400193 |
| … | … |
| 0 | 1.09861228866811… |
| … | … |
| 0.00001 | 1.09861832300329 |
| 0.0001 | 1.09867263829999 |
| 0.001 | 1.09921598420004 |
| 0.01 | 1.104669193785 |
| 0.1 | 1.161231740339 |

**a = 4**

| $\delta x$ | $(a^{\delta x} - 1)/\delta x$ |
|---|---|
| -0.1 | 1.29449436703875 |
| -0.01 | 1.37672955066408 |
| -0.001 | 1.38533389897107 |
| -0.0001 | 1.38619827495728 |
| -0.00001 | 1.38628475210401 |
| … | |
| 0 | 1.38629436111989… |
| … | |
| 0.00001 | 1.38630397022457 |
| 0.0001 | 1.38639045616315 |
| 0.001 | 1.38725571133453 |
| 0.01 | 1.39594797900291 |
| 0.1 | 1.48698354997035 |

So we have the following derivatives

$$f(x) = 1^x \text{ implies } \frac{df}{dx} = 0\,, \qquad f(x) = 2^x \text{ implies } \frac{df}{dx} \approx 0.69315 \times 2^x\,,$$

$$f(x) = 3^x \text{ implies } \frac{df}{dx} \approx 1.09861 \times 3^x\,, \qquad f(x) = 4^x \text{ implies } \frac{df}{dx} \approx 1.38629 \times 4^x\,.$$

But we can't keep doing this everytime we want the derivative of an exponential function. Firstly, apart for $f(x) = 1^x$, our results are not exact. Secondly, there must be a more direct and general way of finding such derivatives which also gives us exact results. And there is. We will find this in two stages. The first stage we will do here, and the second stage we will have to wait to do in the "Differentiation II" notes.

Now, looking at the tables above it seems that there is a specific value of $a$ between $a$ = 2 and $a$ = 3 whereby the value of the limit is 1. If we can find this value of $a$ then our derivative would simply be $df/dx = a^x$ for that value of $a$. So our first stages is to find $a$. Our second stage will be to find a general way of differentiating $f(x) = a^x$ when $a$ is not this specific value. The second stage will involve techniques of differentiation such as the chain rule or implicit differentiation which we will learn only in the "Differentiation II" notes.

For our first stage, let us assume that there is a value of $a$ such that

$$\lim_{\delta x \to 0} \frac{a^{\delta x} - 1}{\delta x} = 1\,.$$

At this point it might be tempting to set up a table of values of $\left(a^{\delta x} - 1\right)/\delta x$ as $\delta x$ approaches zero. But what value of $a$ do we use to do this? Do we use $a$ = 1 or $a$ = 2 or $a$ = 3? What about when $a$ = 3.141592653…. Yet another thing to understand about limits is that just because we can find a limit for one value of a variable it doesn't automatically mean that there is a limit for all values of that variable.

In that case we try a different tactic. When $\delta x$ is small (but does not approach 0) we can say

$$\frac{a^{\delta x} - 1}{\delta x} \approx 1\,.$$

Let us solve this expression for $a$:

$$\begin{aligned} a^{\delta x} - 1 &\approx \delta x\,, \\ a^{\delta x} &\approx 1 + \delta x\,, \\ a &\approx (1 + \delta x)^{1/\delta x}\,. \end{aligned}$$

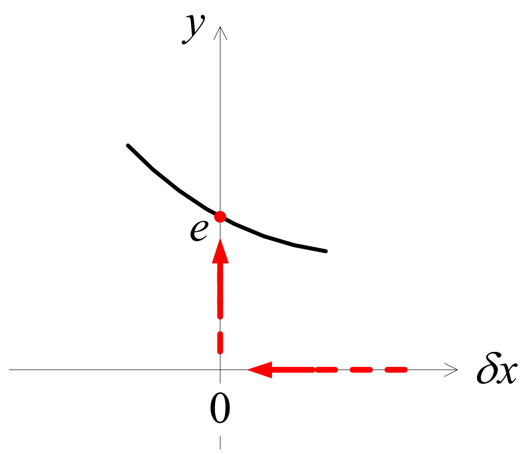
If, when evaluating the right hand side of the above equation as $\delta x$ approaches 0, we get an actual value for $a$ we can say that this is the value that will make the previous limit equal to 1. The evaluation of this limit (from both the left hand side and right hand side of 0) is here:

| **$\delta$x** | ***y* = (1 + $\delta$x)$^{1/\delta x}$** |
|---|---|
| -0.1 | 2.867971990792440 |
| -0.01 | 2.731999026429030 |
| -0.001 | 2.719642216442830 |
| -0.0001 | 2.718417755009750 |
| -0.00001 | 2.718295419980410 |
| -0.000001 | 2.718283187679370 |
| -0.0000001 | 2.718281962942360 |
| -0.00000001 | 2.718281855709170 |
| … | … |

*Table of values* *$\delta$x approaches 0 from the left*

| | |
|---|---|
| ... | ... |
| 0.00000001 | 2.718281786395800 |
| 0.0000001 | 2.718281693980370 |
| 0.000001 | 2.718280469156430 |
| 0.00001 | 2.718268237192300 |
| 0.0001 | 2.718145926824360 |
| 0.001 | 2.716923932235520 |
| 0.01 | 2.704813829421530 |
| 0.1 | 2.593742460100000 |

*Table of values*




*δx approaches 0 from the right*

So the expression $(1 + \delta x)^{1/\delta x}$ does converge, and it converges to the number $a$ = 2.718281… Because of this, our assumption that $\left(a^{\delta x} - 1\right)/\delta x$ converges to the value 1 as $\delta x \to 0$ is correct. So we now have that

$$\frac{df}{dx} = a^x$$

when $a$ = 2.718281… This last number is so special that it is given a symbol: $e$.

We can now complete the process of finding the derivative of $f(x) = a^x$. In general we have

$$\frac{df}{dx} = a^x \lim_{\delta x \to 0} \frac{a^{\delta x} - 1}{\delta x} .$$

But when $a$ = $e$ we have

$$f(x) = e^x \qquad \text{Therefore} \qquad \frac{df}{dx} = e^x .$$

This is the derivative of "the" exponential function. A quicker way of differentiating $e^x$, using something called implicit differentiation, will be shown in the Differentiation II notes.

Note that there are a number of issues relating to $\left(a^{\delta x} - 1\right)/\delta x$ which we have not addressed in terms of formal mathematical anaysis, namely something called continuity (without which the limit may never become equal to 1 even though the table of values above suggests this), and something about the function being said to be strictly increasing (without which the limit may equal 1 for several diferent values of $a$).

This would require learning some more advanced maths which is beyond the scope of these notes, so for the moment we will have to accept the intuitive notion that the above limit does exist.

### *1.11.7 The derivative of $log_a x$*

Having seen how to differentiate $a^x$, where $a$ is a real number, it would be nice to know how to differentiate the inverse of this function. Such an inverse is given by a log function. So let us go through learning how to differentiate $f(x) = \log_a x$, where $a$ is a real number and $x$ is positive only (because, for our purposes, log functions only exists when *x* is positive).

So from 1st principles we have

$$\begin{aligned} \frac{df}{dx} &= \lim_{\delta x \to 0} \frac{\log_a(x + \delta x) - \log_a x}{\delta x} \,, \\ &= \lim_{\delta x \to 0} \frac{\log_a \dfrac{(x + \delta x)}{x}}{\delta x} \,, \\ &= \lim_{\delta x \to 0} \frac{1}{\delta x} \log_a \left(1 + \frac{\delta x}{x}\right) \,, \\ &= \lim_{\delta x \to 0} \log_a \left(1 + \frac{\delta x}{x}\right)^{1/\delta x} \,. \end{aligned}$$

Recall $(1 + \delta x)^{1/\delta x}$. This was a limit we we could evaluate. It gave use the value *e* = 2.71828…. Looking carefully at the form of the limit above we see that we nearly have this form. If we had $(1 + \delta x/x)^{x/\delta x}$ this would be of the form $(1 + z)^{1/z}$ (where $z = \delta x/x$), which we know how to evaluate. So let us create the form that we want by doing an algebraic trick that will allow us to introduce an *x* as an exponent. We do this by mutliplying our limit by *x*/*x*:

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

The reason for this should become clear as we go through the following steps. Now, since *x* is independent of *δx* we can move either one of these into the limit. In our case we will move the numerator "*x*" into the limit:

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

The reason for doing this is because we can now use the rules of log to transform this multiplying "$x$" into a powering "$x$":

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

This is why we wanted to multiply our limit by $x/x$. It was to allow us to be able to introduce the necessary power of $x$ onto the bracketed term so that we could then have the exponent to be the reciprocal form of the fraction inside the bracket so that we could then evaluate the limit.

Wouldn't it be nice if the limit was equal to 1. Then we would have the derivative to simply be $1/x$. But this would depend on the value of $a$. The question is, Is there a value of $a$ for which the limit equals 1? Yes.

To see that there is, let us assume that there is such a value whereby

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

At this point it might be tempting to set up a table of values of $\lim_{\delta x \to 0} \log_a (1 + \delta x/x)^{x/\delta x}$ as $\delta x$ approaches zero. But what value of $a$ do we use to do this? Do we use $a$ = 1 or $a$ = 2 or $a$ = 3? What about when $a$ = 3.141592653….

So we use the same tactic we used in the previous section on the derivative of the exponential function $a^x$, namely that we say

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

when $\delta x$ is small (but does not approach 0). Solving this expression for $a$ we get

$$\left(1 + \frac{\delta x}{x}\right)^{x/\delta x} \approx a. \qquad \{**\}$$

If, when evaluating the right hand side of the above equation, as $\delta x$ approaches 0, we obtain an actual value we can then say that this value is what will make the previous limit equal to 1.

Note that we have already seen a table of values for {**}. This was shown in Table 2 on p15. To see how this is the same as $(1 + \delta x/x)^{x/\delta x}$ let $z = x/\delta x$. Then $(1 + \delta x/x)^{x/\delta x} = (1 + 1/z)^z$. which is the same function as that shown in Table 2 but with $x$ replaced by $z$.

Now, since we have changed our variable from $x$ to $z$, we have to put a $z$ number into our latter expression. In other words, we have to change from a $\delta x$-limiting-value to a $z$-limiting-value. So, as $\delta x$ approaches 0, $z = x/\delta x$ approaches infinity. In fact, this is true however big the values of $x$ become, since ∞ will always be miles bigger than any $x$ value I can imagine. So I don't need to worry about the size of $x$. I just need to evaluate $(1 + 1/z)^z$ based on $z$ approaching infinity.

Hence we set up a table of values for = $(1 + 1/z)^z$ as $z \to \infty$ as shown below.

| **z** | **1/z** | **(1 + 1/z)^z** |
|---|---|---|
| 1 | 1.000000000000000 | 2.000000000000000 |
| 10 | 0.100000000000000 | 2.593742460100000 |
| 100 | 0.010000000000000 | 2.704813829421530 |
| 1000 | 0.001000000000000 | 2.716923932235520 |
| 10000 | 0.000100000000000 | 2.718145926824360 |
| 100000 | 0.000010000000000 | 2.718268237197530 |
| 1000000 | 0.000001000000000 | 2.718280469156430 |
| 10000000 | 0.000000100000000 | 2.718281693980370 |
| 100000000 | 0.000000010000000 | 2.718281786395800 |
| ... | ... | ... |
| Approaching infinity | Approaching 0 | 2.718281828459045 |

*Table 4: Table of values for $(1 + 1/z)^z$ as $z \to \infty$*

From this we see that the expression $(1 + \delta x/x\,)^{x/\delta x}$ does converge, and it converges to the number $e$. Hence our assumption that $\log_a(1 + \delta x/x)^{x/\delta x}$ converged to the value 1 as $\delta x \to 0$ was correct, and the value for $a$ for the base of the log in expression {*} on p84 which makes this true is $a = e$. So we now have that

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

when $a = e = 2.718281$... Noting that $\log_e$ is in fact the ln function we can now complete the process of finding the derivative of $f(x) = \log_a x$:

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

But when $a = e$ we have

$$f(x) = \ln x \qquad \text{Therefore} \qquad \frac{df}{dx} = \frac{1}{x}.$$

This is the derivative of the natural log function. A quicker way of differentiating $\ln x$, using something called implicit differentiation, will be shown in the Differentiation II notes.

*1.11.8 A summary about the concept of limits*

In my opinion one cannot underestimate the importance and subtlety of the idea and use of limits. We have seen many occasion, via the use of table of limiting values, that as $\delta x$ approached 0 (or any variable approached any given value) we obtained specific results to certain ratio. Examples of these ratios include

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e\,, \qquad \lim_{\delta\theta\to 0}\frac{\sin\theta}{\theta} = 1\,, \qquad \lim_{\delta x\to 0}\frac{\cos\delta x - 1}{\delta x} = 0\,,$$

$$\lim_{\delta x\to 0}\frac{a^{\delta x}-1}{\delta x} = 1 \text{ when } a = e\,, \qquad \lim_{\delta x\to 0}\log_a\left(1+\frac{\delta x}{x}\right)^{x/\delta x} = 1 \text{ when } a = e\,.$$

So, the process of two numbers approaching 0 (or some other value) is not the same as the numbers actually being equal to 0 (or that other value). This is a crucial and important difference when we come to divide these two numbers. For example, both $1\times 10^{-99999}$ and $9.999\ldots\times 10^{-100000}$ are very close to 0.

However,

$$\frac{1\times 10^{-99999}}{9.999\ldots\times 10^{-100000}}$$

is not 0/0 but an actual result which is very close to 1. Nervertheless, even these two numbers are "infinitely" miles away from being "infinitely" close to 0, so we continue making these two numbers approach 0 forever. Even though we continue this "approaching 0" process for ever we still end up with an actual result to the division, and we define the ultimate final result of the division to be 1.

This concept applies to all evaluatable limits (not all limits are evaluatable). So in the proof of

$$\lim_{\theta\to 0}(\sin\theta)/\theta = 1$$

in section 1.11.2, the area of the triangles and the area of the sector individually become 0 as $\theta$ approaches 0. However, the ratio of the arc length to vertical length of the inner triangle equals 1 as $\theta$ approaches 0, even though the sector and triangle areas disappear! This is the hidden feature of the geometric construction shown in section 1.11.2, and many other geometries elsewhere in mathematics

But things are even more subtle than this. In the proof of the derivative of $\sin x$ in section 1.11.1 on p63 we saw the following limit:

$$\lim_{\delta x \to 0} \cos\left(x + \frac{\delta x}{2}\right).$$

In this case it might seem that we don't need to use limits since we are not dividing by $\delta x$, hence we will not get the situation of 0/0. Therefore, we might as well do $\delta x = 0$. However, even in this situation we cannot do this despite the fact that it would work to give us the correct answer. This is because the requirement for the use of limits isn't about whether we are dividing quantities which may each be approaching zero, but about whether the function $\cos(x + \delta x/2)$ will actually approach $\cos x$ as $\delta x$ approaches 0. In other words it is about the effect of $\delta x$ approaching 0 that is crucial, irrespective of any division that may or may not be occuring between quantities. And the fact that $\cos(x + \delta x/2) \to \cos x$ as $\delta x \to 0$ is not obvious, and has to be proved (something which is done iusing more advanced maths).

Ultimately, limits can be seen to be a hidden feature of number combinations and function combinations in general. No wonder it took thousands of years, from the time of Achilles and his tortoise, to the end of the 19th century to come to grips with infinitessimals and limits.

## 1.12 Not all functions have a derivative

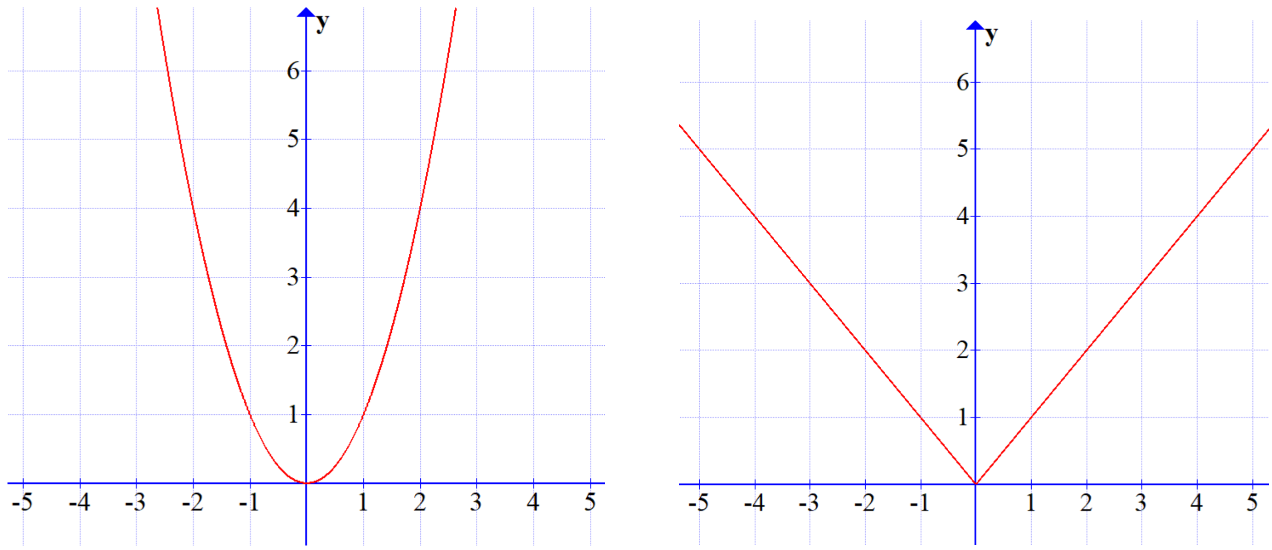
You might think that all functions have derivatives. After all, we have seen that powers of $x$ have derivatives, trig functions have derivatives, and the log and exponential functions have derivatives.

However, these are not the only functions in existence. There are some function for which we cannot evaluate a slope at certain points, and therefore these functions do not have a derivative at those points. We will see how and why this is so in a moment. There are also some functions which do not have a derivative anywhere at all along the function. This means that there is nowhere on the curve of the function where we can calculate its slope. These functions are difficult to conceptualise although they have well defined expressions.

*1.12.1 Functions continuous at a point but whose derivatives are not continuous at that point*

We have to be careful about the way we speak. When we say "A function can be differentiated" what we are really saying is that the function has a derivative (i.e. a gradient or slope) at *any and every point* we care to choose in the domain $[a, b]$ of the function. So we can choose a point $x = c$ anywhere in $[a, b]$, and the curve of $f(x)$ will have a slope at $x = c$. In other words, $f(x)$ will be differentiable at that point.

To see what it is that makes functions not have a derivative at a point let us compare the functions $f(x) = x^2$ and $g(x) = |x|$, the graphs of which are shown below:




$f(x) = x^2$ is a function which has a derivative anywhere and everywhere along $x^2$. But $g(x) = |x|$ does not have a derivative everywhere along $|x|$. There is one point where $|x|$ does not have a derivative. Or, to put it another way, there is one point on the line of $|x|$ where we cannot calculate the slope. This is because when we try to do so we end up with an infinite number of possible results, and not the unique result we need.

Why? What is it about $x^2$ that allows us to find its slope anywhere but not so for $|x|$? Well, looking at the graphs above we see that one difference between $x^2$ and $|x|$ is that the latter function has a "corner" point located at (0, 0), whereas the former does not. This corner point, or lack of it, relates to the smoothness of bending of the line/curve, and this smoothness of bending is related to how the slope changes direction: does it change direction smoothly or abruptly?

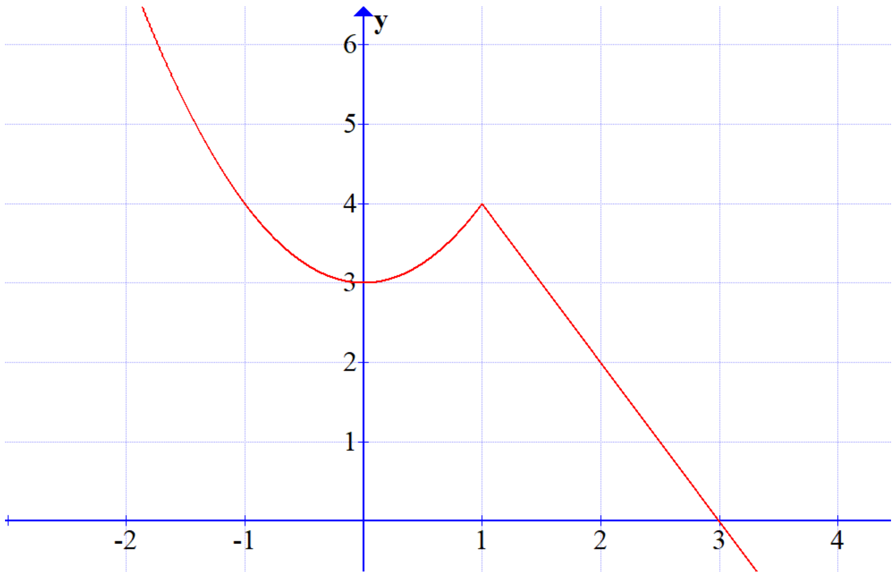
From the graphs above we notice that the change in the direction of the slope of $x^2$ on the left hand side of $x = 0$ is continuous. The change in the direction of the slope of $x^2$ on the right hand side of $x = 0$ is also continous. Visually this is seen as a smooth bending of the curve as x passes though 0 from the left to right.

Not so with |x|. In the graph above we see that there is a sudden/sharp change in the direction of the linear function: one moment it is continuously going downward, then the next moment it suddenly changes direction to go upwards. This suddenness of change in direction is so sudden that it gives rise to the aforementioned corner at $x = 0$. This is one of the reasons why |x| has no slope/derivative at $x = 0$: a complete and total lack of smoothness in the bending of the curve.

To see this effect at a more practical level consider the function

$$h(x) = \begin{cases} x^2 + 3 \text{ if } x < 1 \\ -2x + 6 \text{ if } x \geq 1 \end{cases}$$

The graph of this function is




Let say we want to find the slope of $h(x)$ at $x = 1$. We already know from our previous $h(x)$ doesn't have a slope here, but let us see what happens when we actually try to calculate the slope using the definition of the derivative. In order to do this we have to find the value of the slope as we approach $x = 1$ from the left hand side, and compare this with the value of the slope as we approach $x = 1$ from the right hand side (we have seen this idea before in section

1.2.1 and section 1.2.3 of approaching a point from the left hand side and the right hand side separately).

Also, note that approaching $x = 1$ from the left hand side is equivalent to letting $\delta x$ approach 0 from the left hand side of $x = 1$, and this is symbolised by writing a superscript minus sign on the number 0: $\delta x \to 0^-$. Similarly, approaching $x = 1$ from the right hand side is equivalent to letting $\delta x$ approach 0 from the right hand side of $x = 1$, and this is symbolised by writing a superscript plus sign on the number 0: $\delta x \to 0^+$.

So, in order to distinguish these two cases we will need to set up two separate expressions for the derivative of $h(x)$ at $x = 1$:

approaching $x = 1$ from the left hand side: $$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{h(1+\delta x) - h(1)}{\delta x};$$

approaching $x = 1$ from the right hand side: $$\frac{dh}{dx} = \lim_{\delta x \to 0^+} \frac{h(1+\delta x) - h(1)}{\delta x}.$$

Now we have to use the correct functions in these expressions. So, looking carefully at the definition of $h(x)$ we see that $x^2 + 3$ is defined for all values $x < 1$, but not for the value $x = 1$ (or greater). In this case all we can do is to *approach* $x = 1$ from the left hand side. Hence, the relevant part of $h(x)$ is $h(x) = x^2 + 3$, so $h(1 + \delta x) = (1 + \delta x)^2 + 3$ (remember that, here, $\delta x$ is negative since we are on the left hand side of $x = 1$, hence we are really looking at $1 - \delta x$ where, in this case, $\delta x$ is just the distance on the left hand side of $x = 1$. See section 1.4 where this is discussed in more detail). However, $h(1)$ represent the function evaluated *at* $x = 1$. The only function defined when $x = 1$ is $h(x) = -2x + 6$ so we will use this in order to evaluate $h(1)$.

So we have

$$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{[(1+\delta x)^2 + 3] - [-2(1) + 6]}{\delta x}.$$

All we do now is to use algebra to expand and simplify this expression:

$$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{[(1 + 2\delta x + (\delta x)^2) + 3] - [4]}{\delta x},$$

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

$$= \lim_{\delta x \to 0^-} (2 + \delta x)\,.$$

Evaluating the last limit gives us $dh/dx = 2$. This is how steep the curve is as get infinitely close to $x$ = 1 from the left hand side.

We now find the slope of the curve of the function when we approach x = 1 from the right hand side. In this case $h(1 + \delta x) = -2(1 + \delta x) + 6$. This is because we are on the right hand side of $x$ = 1 and the relevant function in this situation is $-2x + 6$.

So we have

$$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{[-2(1 + \delta x) + 6] - [-2(1) + 6]}{\delta x}.$$

Expanding and simplify this expression we get

$$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{[(-2 - 2\delta x) + 6] - [4]}{\delta x},$$

$$= \lim_{\delta x \to 0^-} \frac{-2\delta x}{\delta x},$$

Evaluating the last limit gives us $dh/dx = -2$. This is how steep the curve is as get infinitely close to $x$ = 1 from the right hand side.

But this is not the same answer as the one we got
when we approached $x$ = 1 from the left hand side.

This is what happens when a function is continuous at a point but does not have a derivative at that point: we get two different answers for our slope.

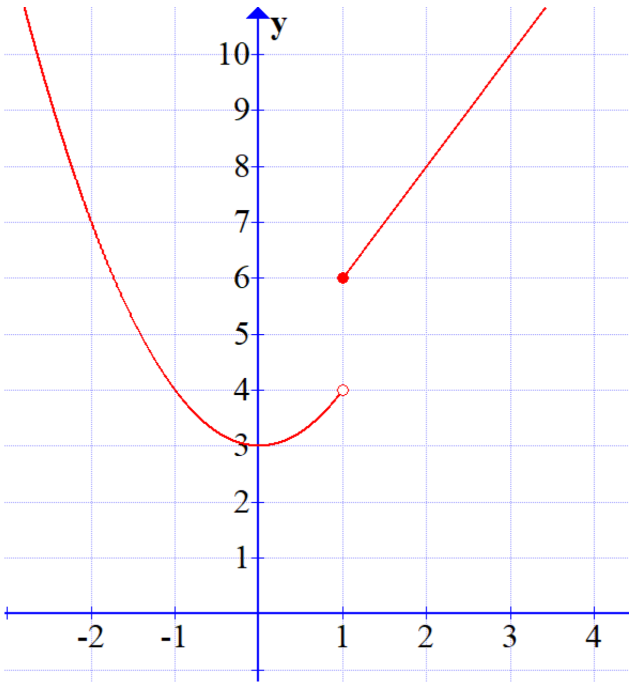
*1.12.2 Functions which are not continuous at a point*

In the previous section we saw that the reason there was no derivative at $x = 1$ was because the derivative of the function itself was not continuous there even though the function itself was continuous there. One might say that, at $x = 1$, the function is not broken but its derivative is.

However, thee is another reason why a function would not have a derivative at a given point, and that is because the function itself is broken/not continuous. To see this consider the function

$$h(x) = \begin{cases} x^2 + 3 \text{ if } x < 1 \\ -2x + 4 \text{ if } x \geq 1 \end{cases}$$

The graph of this function is




Let say we want to find the slope of $h(x)$ at $x = 1$. We know from our previous study on the derivative that in order for there to be a valid slope at x = 1 we need to get the same answer for the slope when we approach $x = 1$ from the left hand side as when we approach $x = 1$ from the right hand side. But our work on this has always presupposed that the curve was continuous, in other words that the curve had no breaks in it.

However, from the graph above, and by the way $h(x)$ is defined, we see that the curve of $h(x)$ has a break at $x$ = 1. Will this situation still allows us to find the slope of the curve at $x$ = 1? In other words will the curve be changing at the same rate, and in the same direction, as we get infinitely close to $x$ = 1 from the left and right hand side? If so then the curve of $h(x)$ will have a slope at $x$ = 1 even though it is not continuous there. Otherwise, the curve will not have a slope at $x$ = 1, and $h(x)$ will not have a derivative there. In order to find out we will have to use the definition of the derivative to calculate the value of the slope as we approach $x$ = 1 from the left and right hand side.

So, looking carefully at the definition of $h(x)$ we see that $x^2 + 3$ is defined for all values $x$ < 1, but not for the value $x$ = 1 (or greater). In this case all we can do is to *approach* $x$ = 1 from the left hand side. Hence, the relevant part of $h(x)$ is $h(x) = x^2 + 3$, so $h(1 + \delta x) = (1 + \delta x)^2 + 3$ (remember that, here, $\delta x$ is negative since we are on the left hand side of $x$ = 1, hence we are really looking at $1 - \delta x$ where, now, $\delta x$ is just the distance on the left hand side of $x$ = 1). However, $h(1)$ represent the function evaluated *at* $x$ = 1. The only function defined when $x$ = 1 is $h(x) = -2x + 4$ so we will use this in order to evaluate $h(1)$.

So we have

$$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{[(1 + \delta x)^2 + 3] - [-2(1) + 4]}{\delta x}.$$

All we do now is to use algebra to expand and simplify this expression:

$$\begin{aligned} \frac{dh}{dx} &= \lim_{\delta x \to 0^-} \frac{[(1 + 2\delta x + (\delta x)^2) + 3] - [2]}{\delta x}, \\ &= \lim_{\delta x \to 0^-} \frac{2 + 2\delta x + (\delta x)^2}{\delta x}, \\ &= \lim_{\delta x \to 0^-} \left(\frac{2}{\delta x} + 2 + \delta x\right). \end{aligned}$$

Trying to evaluate the last limit shows us that $dh/dx$ approaches infinity. What this means in practical terms is that there is no measurable slope to $h(x)$ at $x$ = 1, irrespective of the answer we get to $dh/dx$ as we approach $x$ = 1 from the right hand side.

However, it might still be instructive to find the slope of the curve of the function when we approach $x = 1$ from the right hand side. In this case $h(1+\delta x) = -2(1+\delta x)+4$.

So we have

$$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{[-2(1+\delta x)+4]-[-2(1)+4]}{\delta x}.$$

Expanding and simplify this expression we get

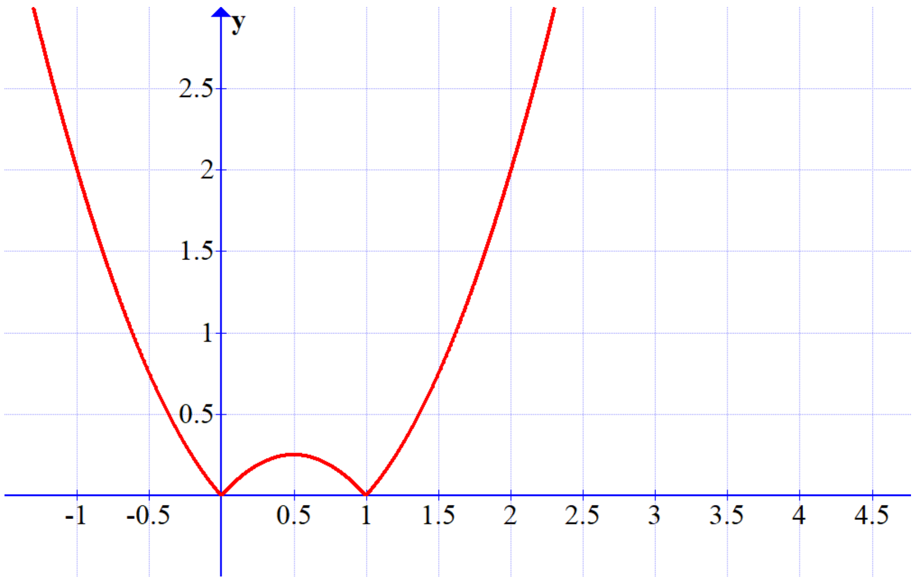
$$\frac{dh}{dx} = \lim_{\delta x \to 0^-} \frac{-2\delta x}{\delta x}.$$

Evaluating the last limit gives us $dh/dx = -2$. Ini this case we see that there is a measurable slope as we get infinitely close to $x = 1$ from the right hand side.

What this example shows us is that when a function is not continuous (i.e. broken) at a point on its curve the slope of the function on one side of that point will not exist even if it does exists on the other side of that point.

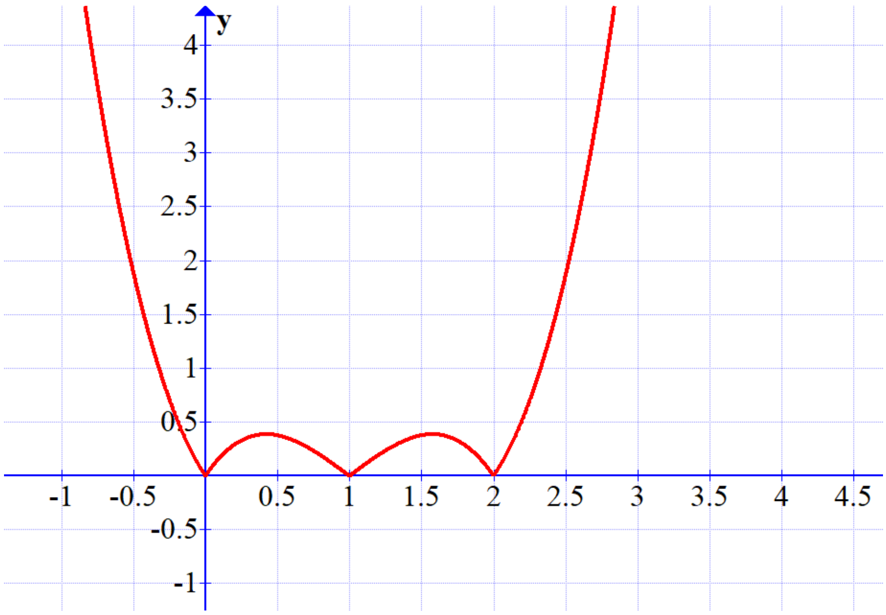
### *1.12.3 A continuous function that does not have a slope anywhere along its curve*

We have seen that continuous functions can be non-differentiable at a point. One example is $f(x) = |x|$ which cannot be differentiated at $x = 0$. What about other functions? What about trying to find the derivative of the function $f(x) = |x||x-1|$? Graphically $f(x)$ looks like this:



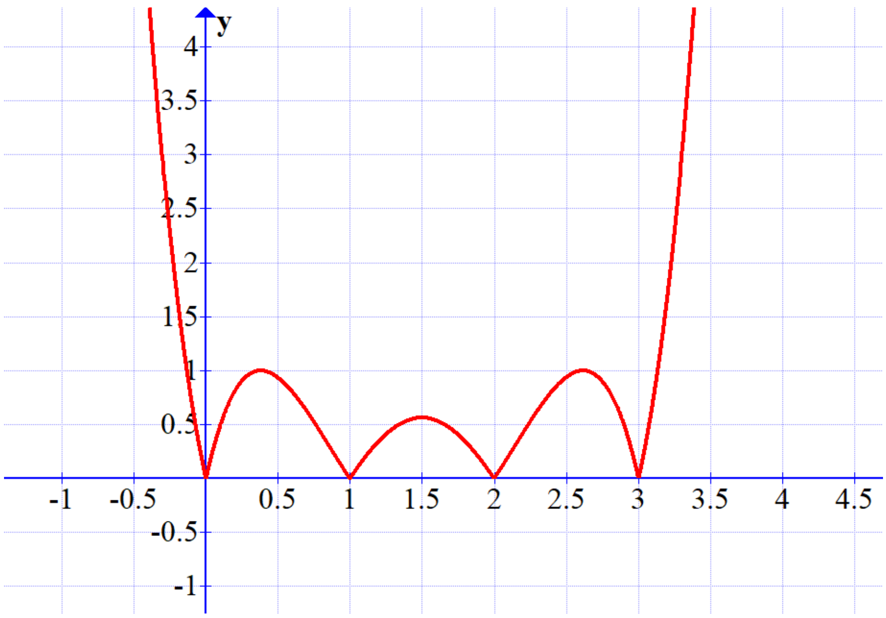

This is basically like the quadratic $f(x) = x(x-1)$ but with the negative part of the curve reflected about the $x$-axis. By this effect (which is due to the effect of the modulus function) we end up with two corners, one at $x = 0$ and one at $x = 1$.

From our previous discussion we see that the curve has no slope at $x = 0$ or $x = 1$. Another way of saying this is that the function does not have a derivative at $x = 0$ or $x = 1$. The same applies for the functions listed below:




*The function* $f(x) = |x||x-1||x-2|$.

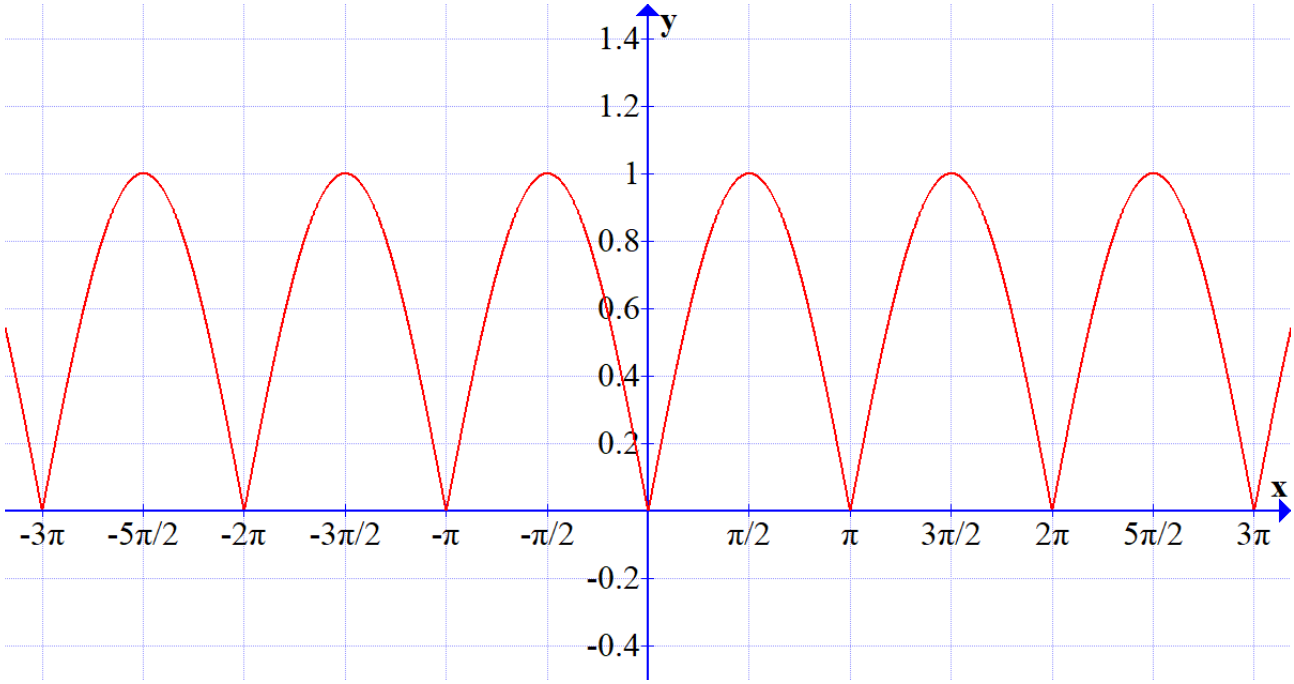
*There is no slope to the curve of f(x) at x = 0, 1, 2*




*The function* $f(x) = |x||x-1||x-2||x-3|$.

*There is no slope to the curve of f(x) at x = 0, 1, 2, 3*

In fact, any function $f(x)$ involving the modulus function will give rise to points where $f(x)$ cannot be differentiated. For example, for $f(x) = |\sin x|$ (shown below) we see that, since $\sin x$ is valid for all $-\infty < x < \infty$ (i.e. $\sin x$ extends infinitely left and right) there are an infinite number of points where we cannot find the slope of $f(x)$, and therefore where $f(x)$ cannot be differentiated at thos points.




*The graph of the function $f(x) = |\sin x|$*

*There is no slope to the curve of f(x) at x = ±nπ, for n = 1, 2, 3, …*

Even though these function have some points at which we cannot differentiate them, they are still differentiable elsewhere. We can still find the slope of these curves at other points in their domain. So it seems that however many corner points a curve may have there will always be parts of the curve between these corners where a slope can be found and the function can be differentiated.

But this is not the case. There are functions which are continuous everyhwere but for which we cannot calculate a slope anywhere along the curve of the function. Roughly speaking these function are such that their curves bend so fast that they end up being “broken” along the whole of the length of the curve. In other words, the whole of the curve is corners. The curve is made up only of corners, and there is no part of the curve that doesn’t have a corner. These functions are indeed continuous but nowhere differentiable.

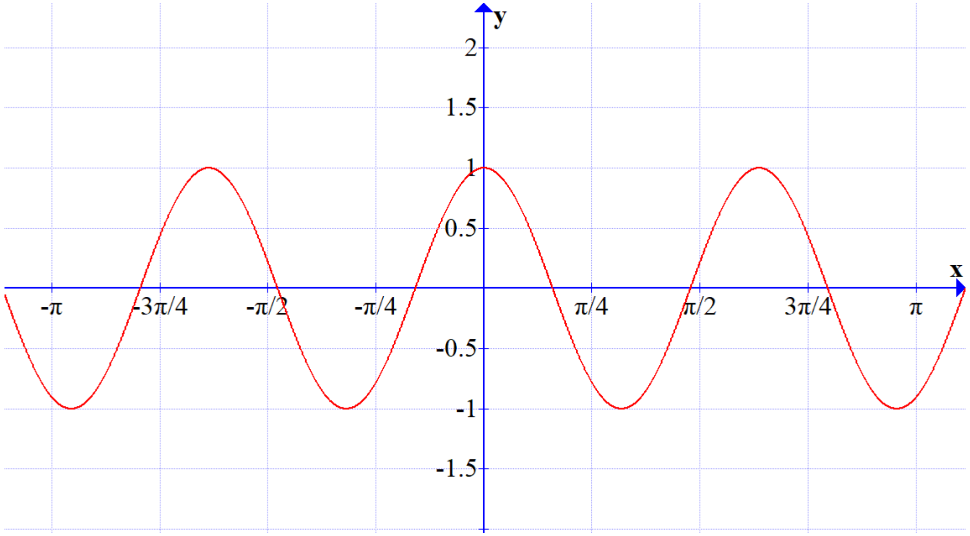
One such function is the Weierstrass function, created by the German mathematician Karl Weierstrass (1815 – 1897). It is a summation function is given by

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$$

where $0 < a < 1$, $ab \geq 1$ and $b > 1$. The key thing which makes this function nothing but corners all along the path of its curve is the fact that, because of the infinite number of terms in the series, the rate of bending of the curve ends up being so fast that the curve "breaks" everywhere along its path. This then makes the function undifferentiable anywhere along the curve.

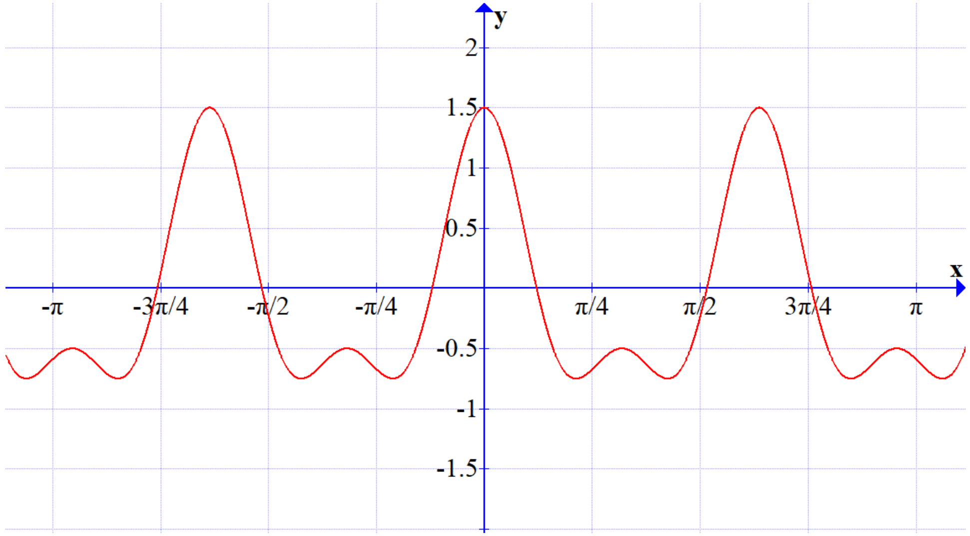
This aspect of the rate of bending of the curve being so fast can be visualised by running the summation formula above. There are two main ways we could do this. We could either fix values for $a$ and $b$ and simply increase the number of terms $n$, or we could fix values for $a$ and $n$ and increase the value of $b$.

The graphs below illustrated on p98 below are for the case when $a$ and $b$ are fixed at 0.5 and 2 respectively, with $n$ running from 0 to 2. Now, keeping $n$ fixed at $n = 2$, (with a still at $a = 0.5$), the graphs illustrated on p## are for the cases when we increase $b$ from 2 onwards in integer steps.



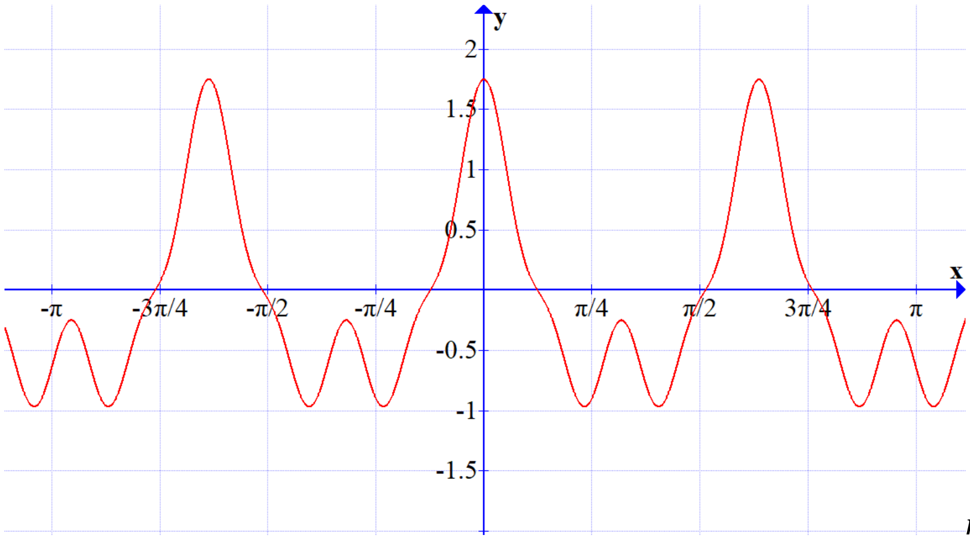

*n = 0 (i.e. one term in the summation)*

*(a = 0.5, b=2)*

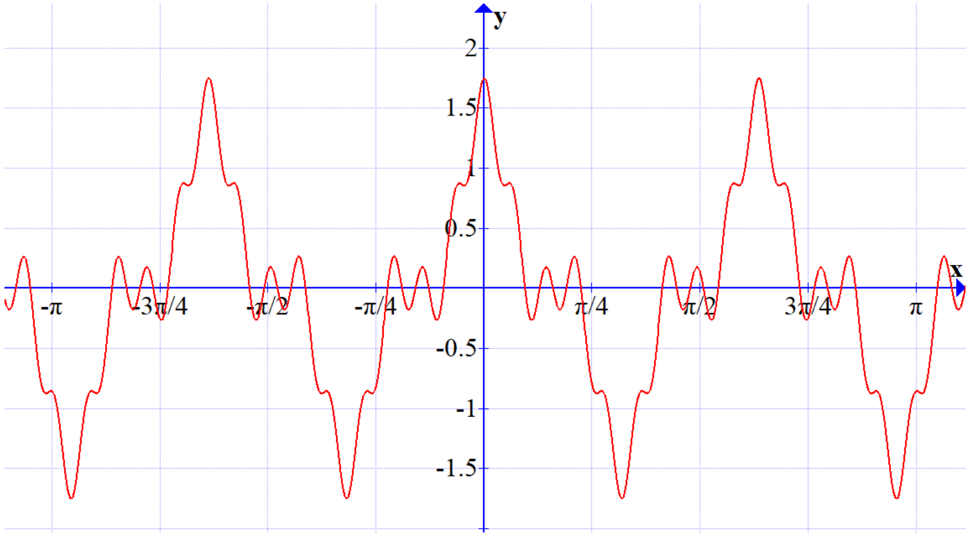



*n = 1 (i.e. two terms in the summation)*

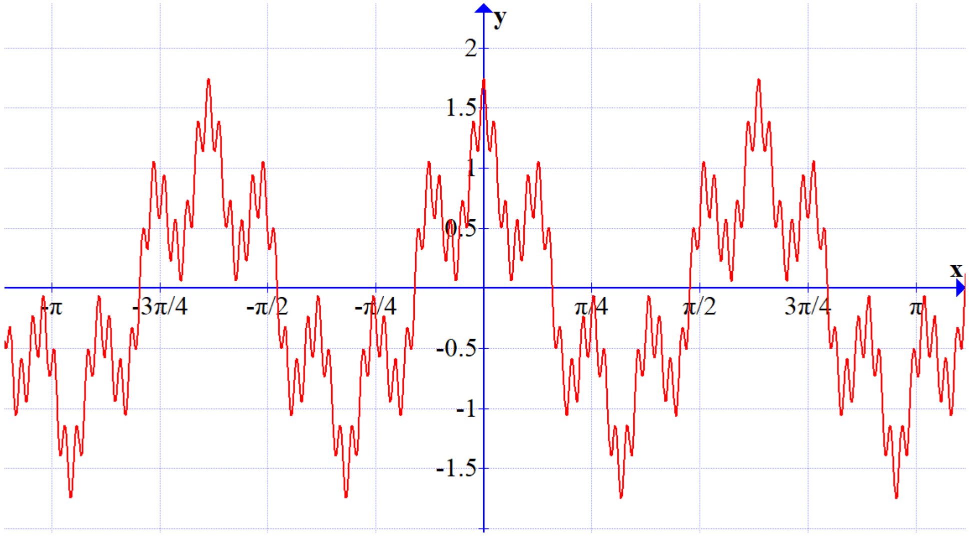
*(a = 0.5, b=2)*



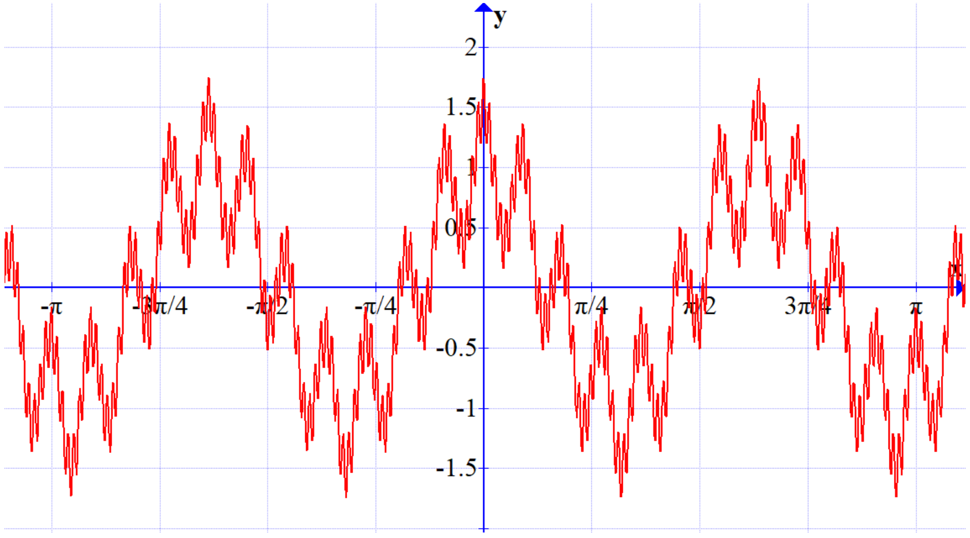

*n = 2 (i.e. three term in the summation)*



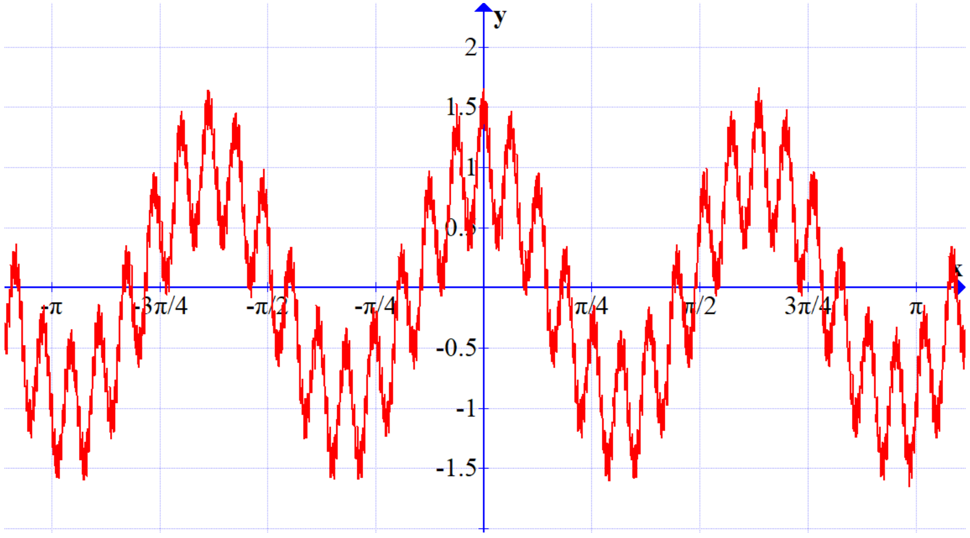

*b = 3 (a = 0.5, n = 3)*




*b = 5 (a = 0.5, n = 3)*

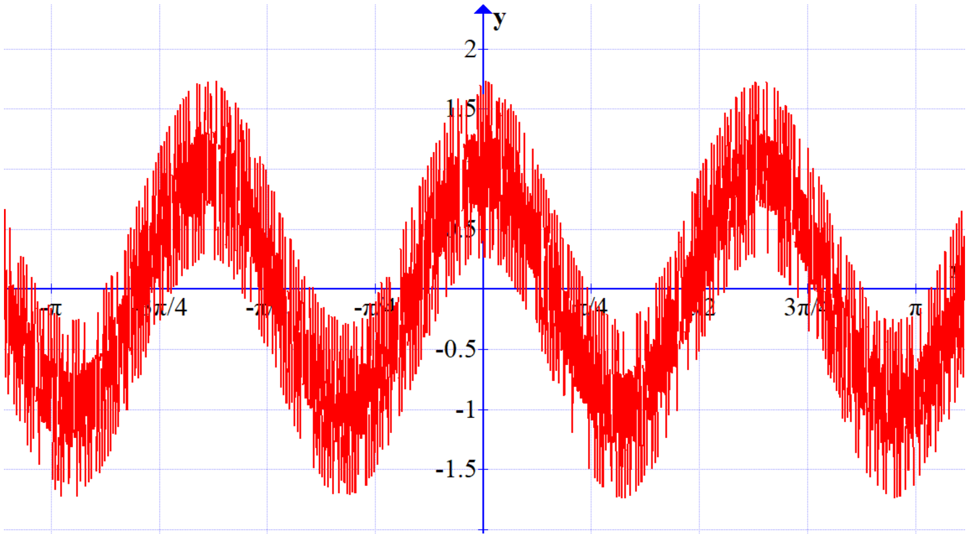



*b = 7 (a = 0.5, n = 3)*




*b = 10 (a = 0.5, n = 3)*

The more terms we add in the summation, or the more we increase the value $b$, the more spicked the function becomes until the whole function is nothing but corners: still continuous everywhere but not diffeentiable anywhere. Ultimately, for a $b$ value of $b$ = 100 we have




even though we have only three terms in the sum. Each upward and downward part of the curve becomes more and more vertical, as well becoming packed closer and closer together. Also, the apex and trough of each curve becomes more and more curved. Ultimatley, either as $n \to \infty$ and/or $b \to \infty$, the upward and downwards parts of the curve get so close together that all gaps between them disappear, and apexes and troughs are no longer curved but are sharp corners. We then have a curve which is continous everywhere but differentiable nowhere since it is made up of corners all along the length of the curve.

If you are interested in seeing other function with this non-differentiable property you can search for the Blancmange function (also known as th eTakagi function), or the saw-tooth function.

## 1.13 A study on derivatives and tangents (to come)

### *1.13.1 An algebraic definition of dy/dx*

### *1.13.2 The derivative as a linear transformation*

### *1.13.3 On subtangents*

### *1.13.4 Fixed point iteration approach to derivatives*

### *1.13.5 The Derivative through the Iteration of Linear Functions*

### *1.13.6 Limit-free differentiation*

### *1.13.7 The tangent not as a limit of secants*

### *1.13.8 The tangent parabola*

### *1.13.9 An area approach to the first derivative*

### *1.13.10 An area approach to the second derivative*