

Differentiation from first principles

mc-TY-firstppls-2009-1

In order to master the techniques explained here it is vital that you undertake plenty of practice exercises so that they become second nature.

After reading this text, and/or viewing the video tutorial on this topic, you should be able to:

- understand the process involved in differentiating from first principles
- differentiate some simple functions from first principles

Contents

1. Introduction	2
2. Differentiating a linear function	2
3. Differentiation from first principles of some simple curves	3

1. Introduction

In this unit we look at how to differentiate very simple functions from first principles. We begin by looking at the straight line.

2. Differentiating a linear function

A straight line has a constant gradient, or in other words, the rate of change of y with respect to x is a constant.

Example

Consider the straight line $y = 3x + 2$ shown in Figure 1.

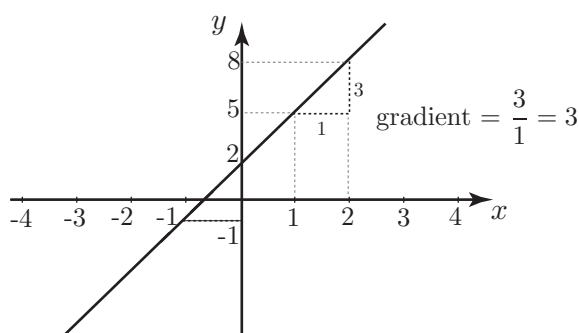


Figure 1. A graph of the straight line $y = 3x + 2$.

We can calculate the gradient of this line as follows. We take two points and calculate the change in y divided by the change in x .

When x changes from -1 to 0 , y changes from -1 to 2 , and so

$$\text{the gradient} = \frac{2 - (-1)}{0 - (-1)} = \frac{3}{1} = 3$$

No matter which pair of points we choose the value of the gradient is always 3.

Values of the function $y = 3x + 2$ are shown in Table 1.

x	-3	-2	-1	0	1	2	3
$3x$	-9	-6	-3	0	3	6	9
2	2	2	2	2	2	2	2
$y = 3x + 2$	-7	-4	-1	2	5	8	11

Table 1: Table of values of $y = 3x + 2$

Look at the table of values and note that for every unit increase in x we always get an increase of 3 units in y . In other words, y increases as a rate of 3 units, for every unit increase in x . We say that “the rate of change of y with respect to x is 3”.

Observe that the gradient of the straight line is the same as the rate of change of y with respect to x .



Key Point

For a straight line:

the rate of change of y with respect to x is the same as **the gradient** of the line.

3. Differentiation from first principles of some simple curves

For any curve it is clear that if we choose two points and join them, this produces a straight line. For different pairs of points we will get different lines, with very different gradients. We illustrate this in Figure 2.

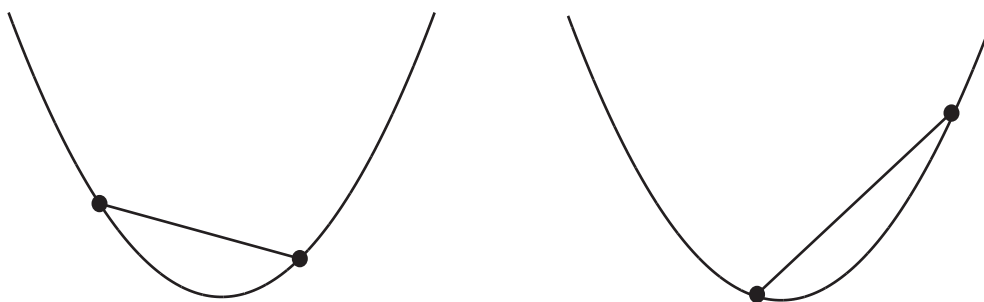


Figure 2. Joining different pairs of points on a curve produces lines with different gradients

Consider a specific example:

Suppose we look at $y = x^2$. Table 2 gives values of this function:

x	-3	-2	-1	0	1	2	3
$y = x^2$	9	4	1	0	1	4	9

Table 2: values of $y = x^2$

Note that as x increases by one unit, from -3 to -2 , the value of y decreases from 9 to 4. It has reduced by 5 units. But when x increases from -2 to -1 , y decreases from 4 to 1. It has reduced by 3. So even for a simple function like $y = x^2$ we see that y is not changing constantly with x . The rate of change of y with respect to x is not a constant.

Calculating the rate of change at a point

We now explain how to calculate the rate of change at any point on a curve $y = f(x)$. This is defined to be the gradient of the tangent drawn at that point as shown in Figure 3.

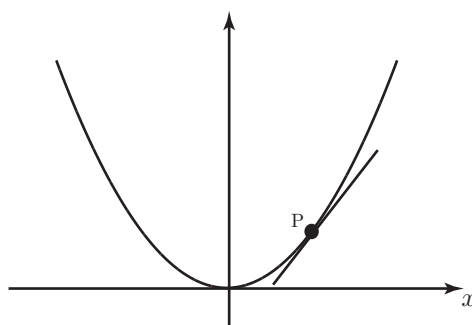


Figure 3. The rate of change at a point P is defined to be the gradient of the tangent at P.



Key Point

The gradient of a curve $y = f(x)$ at a given point is defined to be the gradient of the tangent at that point.

We use this definition to calculate the gradient at any particular point.

Consider Figure 4 which shows a fixed point P on a curve. We also show a sequence of points Q_1, Q_2, \dots getting closer and closer to P . We see that the lines from P to each of the Q 's get nearer and nearer to becoming a tangent at P as the Q 's get nearer to P .

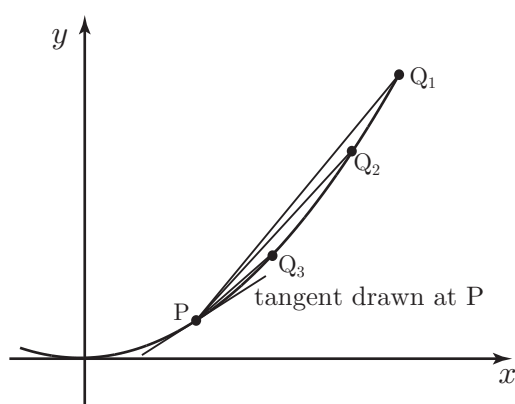


Figure 4. The lines through P and Q approach the tangent at P when Q is very close to P .

So if we calculate the gradient of one of these lines, and let the point Q approach the point P along the curve, then the gradient of the line should approach the gradient of the tangent at P , and hence the gradient of the curve.

Example

We shall perform the calculation for the curve $y = x^2$ at the point, P , where $x = 3$.

Figure 5 shows the graph of $y = x^2$ with the point P marked. We choose a nearby point Q and join P and Q with a straight line. We will choose Q so that it is quite close to P . Point R is vertically below Q , at the same height as point P , so that $\triangle PQR$ is right-angled.

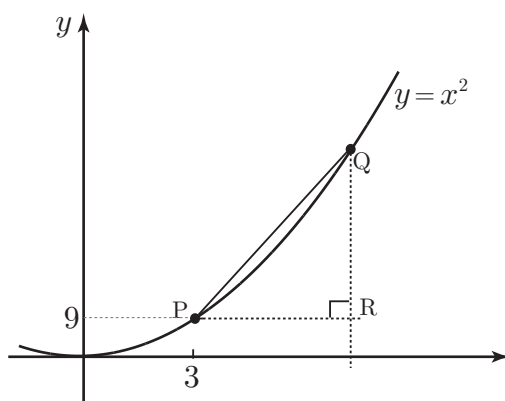


Figure 5. The graph of $y = x^2$. P is the point $(3, 9)$. Q is a nearby point.

Suppose we choose point Q so that $PR = 0.1$. The x coordinate of Q is then 3.1 and its y coordinate is 3.1^2 . Knowing these values we can calculate the change in y divided by the change in x and hence the gradient of the line PQ .

$$\text{gradient of } PQ = \frac{QR}{PR} = \frac{3.1^2 - 3^2}{0.1} = 6.1$$

We can take the gradient of PQ as an approximation to the gradient of the tangent at P , and hence the rate of change of y with respect to x at the point P .

The gradient of PQ will be a better approximation if we take Q closer to P . Table 3 shows the effect of reducing PR successively, and recalculating the gradient.

PR	0.1	0.01	0.001	0.0001
QR	0.61	0.0601	0.006001	0.00060001
$\frac{QR}{PR}$	6.1	6.01	6.001	6.0001

Table 3: The gradient of the line PQ , $\frac{QR}{PR}$ seems to approach 6 as Q approaches P .

Observe that as Q gets closer to P the gradient of PQ seems to be getting nearer and nearer to 6.

We will now repeat the calculation for a general point P which has coordinates (x, y) .

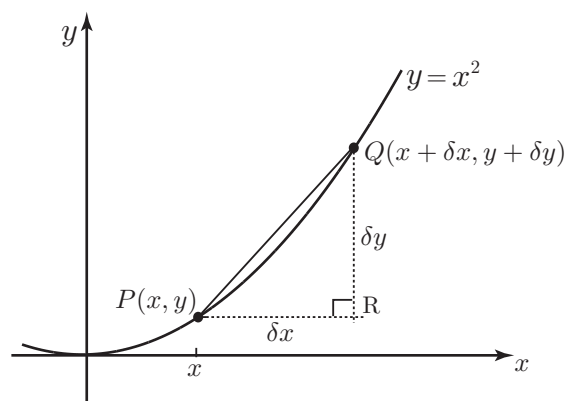


Figure 6. The graph of $y = x^2$. P is the point (x, y) . Q is a nearby point.

Point Q is chosen to be close to P on the curve. The x coordinate of Q is $x + \delta x$ where δx is the symbol we use for a small change, or small increment in x . The corresponding change in y is written as δy . So the coordinates of Q are $(x + \delta x, y + \delta y)$.

Because we are considering the graph of $y = x^2$, we know that $y + \delta y = (x + \delta x)^2$.

Squaring out the brackets:

$$\begin{aligned} y + \delta y &= (x + \delta x)^2 \\ &= x^2 + 2x(\delta x) + (\delta x)^2 \end{aligned}$$

But $y = x^2$ and so

$$\delta y = 2x(\delta x) + (\delta x)^2$$

So the gradient of PQ is

$$\begin{aligned} \frac{\delta y}{\delta x} &= \frac{2x(\delta x) + (\delta x)^2}{\delta x} \\ &= \frac{\delta x(2x + \delta x)}{\delta x} \\ &= 2x + \delta x \end{aligned}$$

As we let δx become zero we are left with just $2x$, and this is the formula for the gradient of the tangent at P . We have a concise way of expressing the fact that we are letting δx approach zero. We write

$$\text{gradient of tangent} = \lim_{\delta x \rightarrow 0} \frac{\delta y}{\delta x} = \lim_{\delta x \rightarrow 0} (2x + \delta x) = 2x$$

'lim' stands for 'limit' and we say that the limit, as δx tends to zero, of $2x + \delta x$ is $2x$. Note that when x has the value 3, $2x$ has the value 6, and so this general result agrees with the earlier result when we calculated the gradient at the point $P(3, 9)$.

We can do this calculation in the same way for lots of curves. We have a special symbol for the phrase

$$\lim_{\delta x \rightarrow 0} \frac{\delta y}{\delta x}$$

We write this as $\frac{dy}{dx}$ and say this as "dee y by dee x". This is also referred to as the **derivative** of y with respect to x .

Use of function notation

We often use function notation $y = f(x)$. Then, the point P has coordinates $(x, f(x))$. Point Q has coordinates $(x + \delta x, f(x + \delta x))$.

So, the change in y , that is δy is $f(x + \delta x) - f(x)$. Then,

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

So

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\delta x \rightarrow 0} \frac{\delta y}{\delta x} \\ &= \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} \end{aligned}$$

This is the definition, for any function $y = f(x)$, of the derivative, $\frac{dy}{dx}$.



Key Point

Given $y = f(x)$, its derivative, or rate of change of y with respect to x is defined as

$$\frac{dy}{dx} = \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x}$$

Example

Suppose we want to differentiate the function $f(x) = \frac{1}{x}$ from first principles.

A sketch of part of this graph is shown in Figure 7. We have marked point $P(x, f(x))$ and the neighbouring point $Q(x + \delta x, f(x + \delta x))$.

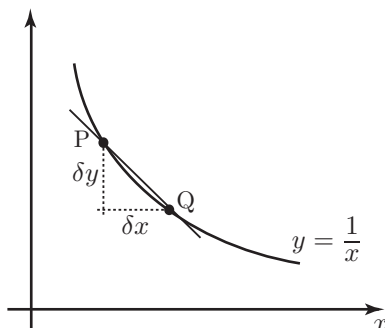


Figure 7. Graph of $y = \frac{1}{x}$.

$$\begin{aligned}\frac{\delta y}{\delta x} &= \frac{f(x + \delta x) - f(x)}{\delta x} \\ &= \frac{\frac{1}{x + \delta x} - \frac{1}{x}}{\delta x} \\ &= \frac{\frac{x - (x + \delta x)}{x(x + \delta x)}}{\delta x} \\ &= \frac{-\delta x}{x(x + \delta x)\delta x} \\ &= \frac{-1}{(x + \delta x)x}\end{aligned}$$

In the limit as $\delta x \rightarrow 0$ this becomes

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

So the derivative of $y = \frac{1}{x}$ with respect to x is $-\frac{1}{x^2}$.