

In calculus, a branch of mathematics, the **derivative** is a measurement of how a function changes when the values of its inputs change. Simply speaking, a **derivative** can be thought of as how much a quantity is changing at some given point. For example, the derivative of the position of a car at some point in time is the velocity, or speed, at which that car is travelling (conversely the integral of the velocity is the car's position or distance travelled).

The **derivative** of a function at a chosen input value describes the best linear approximation of the function near that input value.

For a real-valued function of a single real variable, **the derivative at a point equals the slope of the tangent line to the graph of the function at that point.**

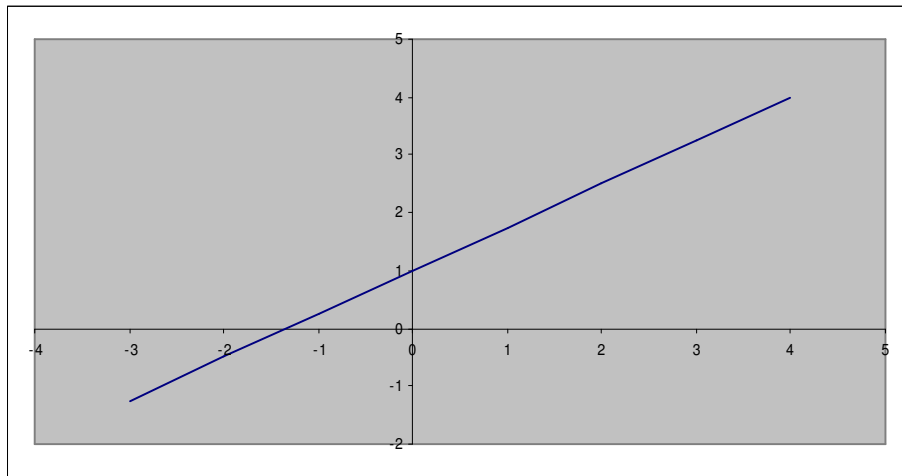
The process of finding a derivative is called **differentiation**. The fundamental theorem of calculus states that differentiation is the reverse process to integration.

Differentiation is a method to compute the rate at which a quantity, y , changes with respect to the change in another quantity, x , upon which it is dependent. This rate of change is called the **derivative** of y with respect to x . In more precise language, the dependency of y on x means that y is a function of x . If x and y are real numbers, and if the graph of y is plotted against x , *the derivative measures the slope of this graph at each point*. This functional relationship is often denoted $y = f(x)$, where f denotes the function.

The simplest case is when y is a linear function of x , meaning that the graph of y against x is a straight line. The equation of all linear graphs is of the form: **$y = f(x) = mx + c$** where **m** is the *gradient* and **c** is the *interception* point of the line with the y -axis.

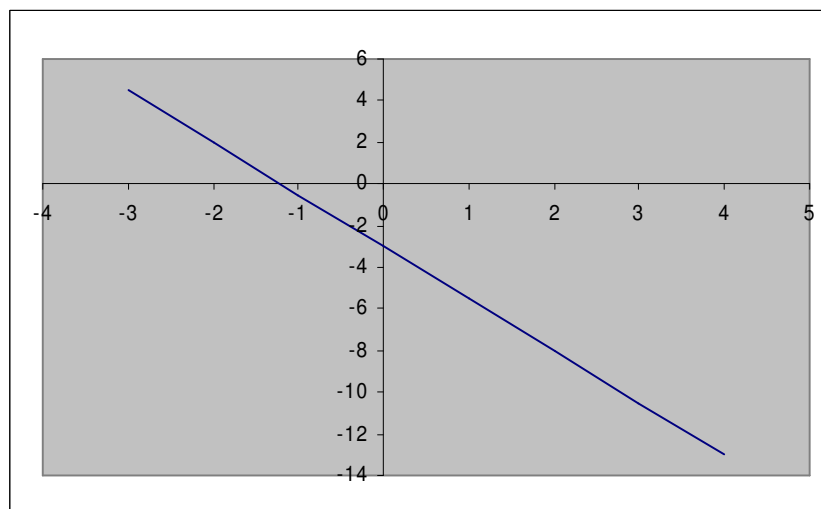
If a line forms an angle α with the positive x – axis, then we speak about the gradient of a line, i.e. the number $m = \operatorname{tg} \alpha = \frac{My - Ny}{Mx - Nx}$. The numerator shows us the function, while the denominator indicates the variable.

Example 1: $y = \frac{3}{4}x + 1$ Gradient is $\frac{3}{4}$, and an intercept with the y -axis is at 1. The gradient is positive, therefore the line is sloping upwards from left to right.



Example 2: $y = -\frac{5}{2}x - 3$ Gradient is $-\frac{5}{2}$ and an intercept with the y-axis is at -3 .

The gradient is negative, and therefore the line is sloping downwards from left to right.



Being able to identify the gradient and intercept with the y-axis from a linear equation enables you to sketch or draw a graph.

The gradient of a straight line is the same at all points on the line. With a curve however the gradient, or steepness, will depend upon where we are on the curve.

The gradient at a point P on a curve is defined as the gradient of the *tangent* drawn to the curve at this point P, i.e. the gradient of the line just touches the curve at the point P.

Suppose that we wish to find the gradient of a curve $y = f(x)$ at a point $P[x, y]$ on the curve. Consider a second point, Q , lying on the curve, near to P and with x – coordinate given by $x + \delta x$, where δx is used to denote a small increment of length in the direction of the x – axis.

$$\text{Thus the gradient of the chord } PQ = \frac{Q_y - P_y}{Q_x - P_x} = \frac{f(x + \delta x) - f(x)}{x + \delta x - x} = \frac{f(x + \delta x) - f(x)}{\delta x}$$

Now if we move Q nearer to P , say to points Q_1, Q_2, \dots then the gradient of the chords PQ, PQ_1, PQ_2, \dots will give better and better approximations for the gradient of the tangent at P and therefore, the gradient of the curve at P .

If we say that P has coordinates $[x, y]$ and Q has coordinates $[x + \delta x, y + \delta y]$ then the

$$\text{gradient } m = \lim_{\delta x \rightarrow 0} \frac{y + \delta y - y}{x + \delta x - x} = \lim_{\delta x \rightarrow 0} \frac{\delta y}{\delta x}$$

We write $\frac{dy}{dx}$ (pronounced ‘dee y by dee x’) or y' and we call it the **differential coefficient**

of y with respect to x . It gives the formula by which the gradient at any point on the line can be determined. The process of finding the differential coefficient of a function is called **differentiation**.

Differentiation Rules

Here follow the formulas and rules for differentiating the most common **functions**. In each case, the variable with respect to which the **derivative** is being taken is x .

Formulas for differentiating

FUNCTION $f(x)$	DERIVATIVE OF $f(x) =$ $= f'(x)$
$c \in \mathbb{R}$	0
x	1
x^n	$n \cdot x^{n-1}$
$\ln x$	$\frac{1}{x}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$

$\operatorname{tg} x$	$\frac{1}{\cos^2 x}$
$\operatorname{cotg} x$	$-\frac{1}{\sin^2 x}$
a^x	$a^x \cdot \ln a$
e^x	e^x

Rules for differentiating:

$$[f(x) \pm g(x)]' = f'(x) \pm g'(x)$$

$$[f(x) \cdot g(x)]' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

$$\left[\frac{f(x)}{g(x)} \right]' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x) \cdot g(x)}$$

Differentiation of a complex function:

$$y = f[g(x)], u = g(x) \Rightarrow y = f(u)$$

$$y' = f'(u) \cdot u' = f'[g(x)] \cdot g'(x)$$

Example: $y = \operatorname{tg}(3x^5 - 4x + 1)$

$$y' = \frac{1}{\cos^2(3x^5 - 4x + 1)} \cdot (15x^4 - 4)$$

The derivative as a function

Let f be a function that has a derivative at every point a in the domain of f . Because every point a has a derivative, there is a function which sends the point a to the derivative of f at a . This function is written $f'(x)$ and is called the *derivative function* or the *derivative of f* . The derivative of f collects all the derivatives of f at all the points in the domain of f . Sometimes f has a derivative at most, but not all, points of its domain. The function whose value at a equals $f'(a)$ whenever $f'(a)$ is defined and is undefined elsewhere is also called the derivative of f . It is still a function, but its domain is strictly smaller than the domain of f .

Using this idea, differentiation becomes a function of functions: The derivative is an operator whose domain is the set of all functions which have derivatives at every point of their domain and whose range is a set of functions. If we denote this operator by D , then $D(f)$ is the function $f'(x)$. Since $D(f)$ is a function, it can be evaluated at a point a . By the definition of the derivative function, $D(f)(a) = f'(a)$.

Higher derivatives

Let f be a differentiable function, and let $f'(x)$ be its derivative. The derivative of $f'(x)$ (if it has one) is written $f''(x)$ and is called the **second derivative** of f . Similarly, the derivative of a second derivative, if it exists, is written $f'''(x)$ and is called the **third derivative** of f . These repeated derivatives are called **higher-order derivatives**.

A function that has k successive derivatives is called **k times differentiable**.

On the real line, every polynomial function is infinitely differentiable. By standard differentiation rules, if a polynomial of degree n is differentiated n times, then it becomes a constant function. All of its subsequent derivatives are identically zero. In particular, they exist, so polynomials are smooth functions.

Consider a function $f(x)$. When we calculate the derivative f' of the function at a point $x = a$ say, we are finding the gradient of the tangent to the graph of that function at that point. The tangent drawn at $x = a$ has the gradient $f'(a)$.

We will use this information to calculate *the equation of the tangent to a curve at a particular point, and then the equation of the normal to a curve at a point*.

Remember: $f'(a)$ is the gradient of the tangent drawn at $x = a$.

The equation of a tangent

The equation of a straight line (we will look at tangent and later at normal as well) that passes through a point (x_0, y_0) and has gradient m is given by $m = \frac{y - y_0}{x - x_0}$, which implies

$y - y_0 = f'(x_0)(x - x_0)$, since $m = f'(x_0)$.

Example 1: Suppose we wish to find the equation of the tangent to

$f(x) = x^3 - 3x^2 + x - 1$ at the point where $x = 3$.

When $x = 3$ we note that $f(3) = 3^3 - 3 \cdot 3^2 + 3 - 1 = 27 - 27 + 3 - 1 = 2$

So the point of interest has coordinates (3, 2).

The next thing that we need is the gradient of the curve at this point. To find this, we need to differentiate $f(x)$:

$$f'(x) = 3x^2 - 6x + 1$$

We can now calculate the gradient of the curve at the point where $x = 3$.

$$f'(3) = 3 \cdot 3^2 - 6 \cdot 3 + 1 = 27 - 18 + 1 = 10$$

So we have the coordinates of the required point, (3, 2), and the gradient of the tangent at that point, that is 10.

What we want to calculate is the equation of the tangent at this point on the curve. The tangent must pass through the point and have gradient 10. The tangent is a straight line and so we use the fact that the equation of a straight line that passes through a touch-point (x_0, y_0) and has gradient m is given by the formula:

$$m = \frac{y - y_0}{x - x_0}$$

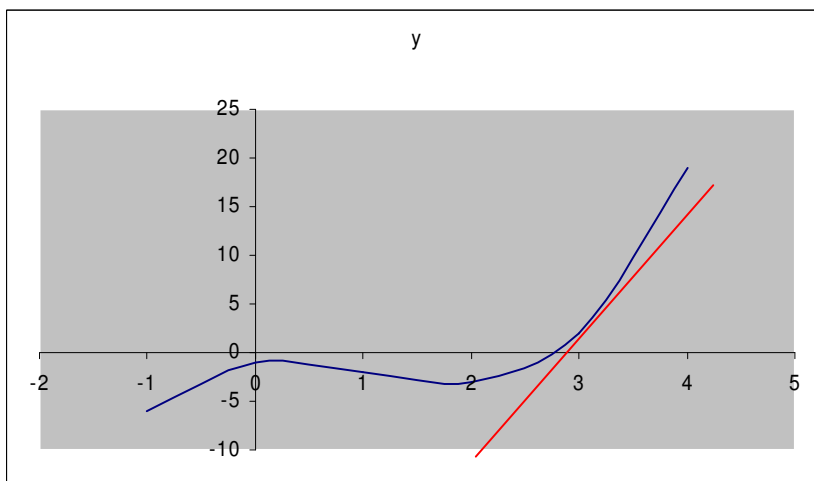
Substituting the given values $10 = \frac{y - 2}{x - 3}$ and rearranging

$$y - 2 = 10(x - 3)$$

$$y - 2 = 10x - 30$$

$$y = 10x - 28$$

This is the equation of the tangent to the curve at the point (3, 2).



Example 2:

Suppose we wish to find points on the curve $y(x)$ given by

$y = x^3 - 6x^2 + x + 3$ where the tangents are parallel to the line $y = x + 5$.

If the tangents have to be parallel to the line then they *must have the same gradient*. The standard equation for a straight line is $y = mx + c$, where m is the gradient. So what we gain from looking at this standard equation and comparing it with the straight line $y = x + 5$ is that the gradient, m , is equal to 1. Thus the gradients of the tangents we are trying to find must also have gradient 1.

We know that if we differentiate $y(x)$ we will obtain an expression for the gradients of the tangents to $y(x)$ and we can set this equal to 1. Differentiating, and setting this equal to 1 we find $y' = 3x^2 - 12x + 1 = 1$

from which $3x^2 - 12x = 0$

This is a quadratic equation which we can solve by factorisation.

$$3x^2 - 12x = 0$$

$$3x(x - 4) = 0$$

$$3x = 0 \text{ or } x - 4 = 0$$

$$x = 0 \text{ or } x = 4$$

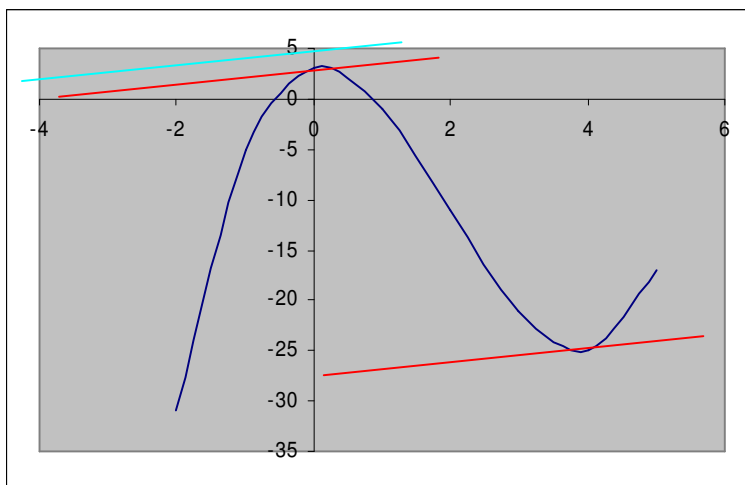
Now having found these two values of x we can calculate the corresponding y coordinates. We do this from the equation of the curve: $y = x^3 - 6x^2 + x + 3$

When $x = 0$: $y = 0^3 - 6 \cdot 0^2 + 0 + 3 = 3$.

When $x = 4$: $y = 4^3 - 6 \cdot 4^2 + 4 + 3 = 64 - 96 + 4 + 3 = -25$.

So the two points are $(0, 3)$ and $(4, -25)$

These are the two points where the gradients of the tangent are equal to 1, and so where the tangents are parallel to the line that we started out with, i.e. $y = x + 5$.



Exercise:

1. For each of the functions given below determine the equation of the tangent at the points indicated.

a) $f(x) = 3x^2 - 2x + 4$ at $x = 0$ and 3 .

b) $f(x) = 5x^3 + 12x^2 - 7x$ at $x = -1$ and 1 .

c) $f(x) = 1 - 2x$ at $x = -3, 0$ and 2 .

2. Find the equation of each tangent of the function $f(x) = x^3 - 5x^2 + 5x - 4$ which is parallel to the line $y = 2x + 1$.

3. Find the equation of each tangent of the function $f(x) = x^3 + x^2 + x + 1$ which is perpendicular to the line $2y + x + 5 = 0$.

The equation of a normal to a curve

In mathematics the word 'normal' has a very specific meaning. It means 'perpendicular' or 'at right angles'.

The *normal* is a line at right angles to the tangent.

If we have a curve we can choose a point and draw in the tangent to the curve at that point. The normal is then at right angles to the curve so it is also at right angles (perpendicular) to the tangent.

We now find the equation of the normal to a curve. There is one further piece of information that we need in order to do this. If two lines, having gradients m_1 and m_2 respectively, are at right angles to each other then the product of their gradients $m_1 \cdot m_2$, must equal -1 , i.e.

If two lines, with gradients m_1 and m_2 are at right angles then $m_1 \cdot m_2 = -1$

Example 1:

Suppose we wish to find the equation of the tangent and the equation of the normal to the

curve $y = x + \frac{1}{x}$ at the point where $x = 2$.

First of all we shall calculate the y coordinate at the point on the curve where $x = 2$:

$$y = 2 + \frac{1}{2} = \frac{5}{2}$$

Next we want the gradient of the curve at the point $x = 2$. We need to find y' .

Noting that we can write y as $y = x + x^{-1}$ then

$$y' = 1 - x^{-2} = 1 - \frac{1}{x^2}$$

Furthermore, when $x = 2$, $y'(2) = 1 - \frac{1}{4} = \frac{3}{4}$

This is the gradient of the tangent to the curve at the point $(2, \frac{5}{2})$.

We know that the standard equation for a straight line is $m = \frac{y - y_0}{x - x_0}$,

With the given values we have $y - \frac{5}{2} = \frac{3}{4} (x - 2)$

Rearranging $4y - 10 = 3x - 6$
 $4y = 3x + 4$

So the equation of the *tangent* to the curve at the point where $x = 2$ is $4y = 3x + 4$.

Now we need to find the equation of the *normal* to the curve.

Let the gradient of the normal be m_2 . Suppose the gradient of the tangent is m_1 . Recall that the normal and the tangent are perpendicular and hence $m_1 \cdot m_2 = -1$. We know

$$m_1 = \frac{3}{4}. \text{ So } \frac{3}{4} \cdot m_2 = -1, \text{ which implies that } m_2 = -\frac{4}{3}$$

So we know the gradient of the normal and we also know the point on the curve through which it passes, $(2, \frac{5}{2})$.

As before, $m = \frac{y - y_0}{x - x_0}$

$$y - \frac{5}{2} = -\frac{4}{3} (x - 2)$$

Rearranging $6y - 15 = -8x + 16$
 $6y + 8x = 31$

This is the equation of the normal to the curve at the given point.



Maxima and minima

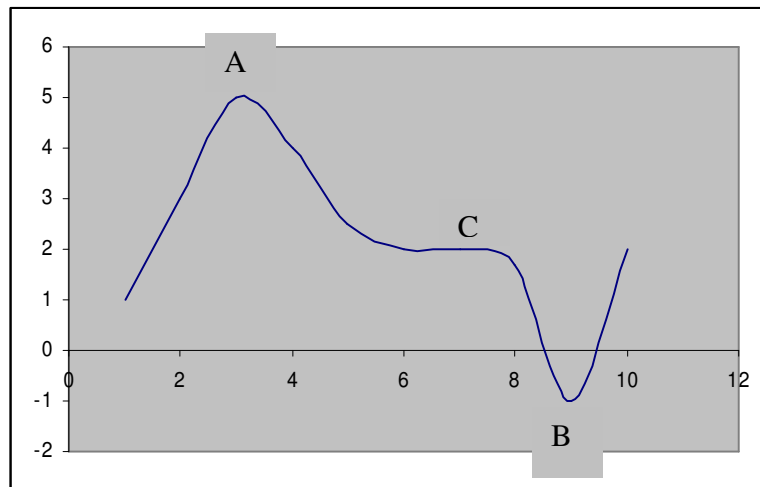
In this unit we show how differentiation can be used to find the maximum and minimum values of a function.

Because the derivative provides information about the gradient or slope of the graph of a function we can use it to locate points on a graph where the gradient is zero. We shall see that such points are often associated with the largest or smallest values of the function, at least in their immediate locality.

In many applications, a scientist, engineer, or economist for example, will be interested in such points for obvious reasons such as maximising power, or profit, or minimising losses or costs.

Stationary points

When using mathematics to model the physical world in which we live, we frequently express physical quantities in terms of variables. Then, functions are used to describe the ways in which these variables change. A scientist or engineer will be interested in the ups and downs of a function, its maximum and minimum values, its turning points. Drawing a graph of a function using a computer graph plotting package will reveal this behaviour, but if we want to know the precise location of such points we need to turn to algebra and differential calculus. In this section we look at how we can find maximum and minimum points in this way.



Consider the graph of the function, $y = f(x)$. If, at the points A, B and C, we draw tangents to the graph, note that these are parallel to the x axis, and thus the gradient is 0.

This means that at each of the points A, B and C the gradient of the graph is zero.

We know that the gradient of a graph is given by $f'(x_0)$

Consequently, $f'(x_0) = 0$ at points A, B and C. All of these points are known as **stationary points**.

Any point at which the tangent to the graph is horizontal, i.e. is parallel to the x-axis, is called a **stationary point**, and the first derivative at these points equals zero.

We can locate stationary points by looking for points at which $f'(x_0) = 0$

Turning points

Notice that at points A and B the curve actually turns. These two stationary points are referred to as **turning points**. Point C is not a turning point because, although the graph is flat for a short time, the curve continues to go down as we look from left to right.

So, all turning points are stationary points. But not all stationary points are turning points (e.g. point C).

In other words, there are points for which $f'(x_0) = 0$, which are not turning points.

At a turning point $f'(x_0) = 0$

Not all points where $f'(x_0) = 0$ are turning points, i.e. not all stationary points are turning points.

Point A is called a *local maximum* because in its immediate area it is the highest point, and so represents the greatest or maximum value of the function.

Point B is called a *local minimum* because in its immediate area it is the lowest point, and so represents the least, or minimum, value of the function.

Simply speaking, we refer to a local maximum as simply a maximum. Similarly, a local minimum is often just called a minimum.

Distinguishing maximum points from minimum points

Think about what happens to the gradient of the graph as we travel through the minimum turning point, from left to right, that is as x increases.

$f'(x)$ goes from negative through zero to positive as x increases.

Notice that to the left of the minimum point, $f'(x)$ is negative because the tangent has negative gradient. At the minimum point, $f'(x) = 0$. To the right of the minimum point $f'(x)$ is positive, because here the tangent has a positive gradient. So, $f'(x)$ goes from negative, to zero, to positive as x increases. In other words, $f'(x)$ must be increasing as x increases.

In fact, we can use this observation, once we have found a stationary point, to check if the point is a minimum. If $f'(x)$ is increasing near the stationary point then that point must be minimum.

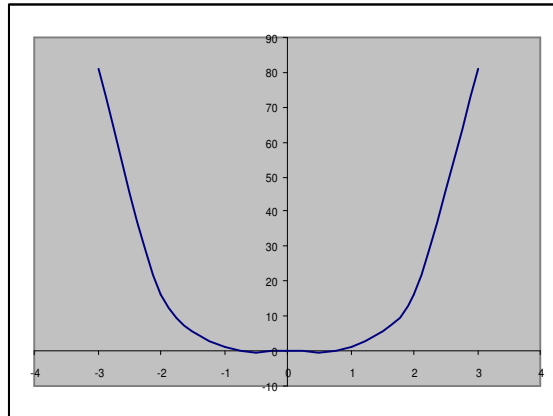
Now, if the derivative of $f'(x)$ is positive then we will know that $f'(x)$ is increasing; so we will know that the stationary point is a minimum. Now the derivative of $f'(x)$, called the second derivative, is written $f''(x)$. We conclude that if $f''(x)$ is positive at a stationary point, then that point must be a minimum turning point.

If $f'(x) = 0$ at a point, and if $f''(x) > 0$ there, then that point must be a **minimum**.

It is important to realise that this test for a minimum is not conclusive. It is possible for a stationary point to be a minimum even if $f''(x)$ equals 0, although we cannot be certain: other types of behaviour are possible. (However, we cannot have a minimum if $f''(x)$ is negative.)

To see this consider the example of the function $y = x^4$

A graph of this function is as follows:



There is clearly a minimum point when $x = 0$. But $f'(x) = 4x^3$ and this is clearly zero when $x = 0$. Differentiating again $f''(x) = 12x^2$ which is also zero when $x = 0$.

The function $y = x^4$ has a minimum at the origin where $x = 0$, but $f''(x) = 0$ and so is not greater than 0.

Now think about what happens to the gradient of the graph as we travel through the maximum turning point, from left to right, that is as x increases. $f'(x)$ is negative, $f'(x)$ is zero, and $f'(x)$ is positive. I.e. $f'(x)$ goes from positive through zero to negative as x increases.

Notice that to the left of the maximum point, $f'(x)$ is positive because the tangent has positive gradient. At the maximum point, $f'(x) = 0$. To the right of the maximum point $f'(x)$ is negative, because here the tangent has a negative gradient. So, $f'(x)$ goes from positive, to zero, to negative as x increases.

In fact, we can use this observation to check if a stationary point is a maximum. If $f'(x)$ is decreasing near a stationary point then that point must be maximum. Now, if the derivative of $f'(x)$ is negative then we will know that $f'(x)$ is decreasing; so we will know that the stationary point is a maximum. As before, the derivative of $f'(x)$, the second derivative is $f''(x)$. We conclude that if $f''(x)$ is negative at a stationary point, then that point must be a maximum turning point.

If $f'(x) = 0$ at a point, and if $f''(x) < 0$ there, then that point must be a maximum.

It is important to realise that this test for a maximum is not conclusive. It is possible for a stationary point to be a maximum even if $f''(x) = 0$, although we cannot be certain: other types of behaviour are possible. But we cannot have a maximum if $f''(x) > 0$, because, as we have already seen the point would be a minimum.

The second derivative test: summary

We can locate the position of stationary points by looking for points where $f'(x) = 0$

As we have seen, it is possible that some such points will not be turning points. We can calculate $f''(x)$ at each point we consider 'suspicious' for an extreme:

- If $f''(x)$ is positive then the stationary point is a minimum turning point.
- If $f''(x)$ is negative, then the point is a maximum turning point.
- If $f''(x) = 0$ it is possible that we have a maximum, or a minimum, or indeed other sorts of behaviour, i.e. a point of inflexion. So if $f''(x) = 0$ this second derivative test does not give us useful information and we must seek an alternative method – We need to consider the sign of $f'(x)$ on either side of the point, i.e. check the monotony of the function in the neighbourhood of that point.

Example:

Suppose we wish to find the turning points of the function $y = x^3 - 3x + 2$ and distinguish between them.

We need to find where the turning points are, and whether we have maximum or minimum points. First of all we carry out the differentiation and set $f'(x)$ equal to zero.

This will enable us to look for any stationary points, including any turning points.

$$y = x^3 - 3x + 2$$

$$y' = 3x^2 - 3$$

At stationary points, $f'(x) = 0$ and so $3x^2 - 3 = 0$

$$3(x^2 - 1) = 0 \text{ (factorising)}$$

$$3(x - 1)(x + 1) = 0 \text{ (factorising the difference of two squares)}$$

It follows that either $x - 1 = 0$ or $x + 1 = 0$ and so either $x = 1$ or $x = -1$.

We have found the x coordinates of those points on the graph, where $f'(x) = 0$, that is the *stationary points*. We need the y coordinates which are found by substituting the x values in the original function $y = x^3 - 3x + 2$.

When $x = 1$: $y = 0$.

When $x = -1$: $y = 4$.

To summarise, we have located two stationary points and these occur at $(1, 0)$ and $(-1, 4)$.

Next we need to determine whether we have maximum or minimum points, or possibly a point of inflexion which is neither maximum nor minimum.

We have seen that the first derivative $f'(x) = 3x^2 - 3$

Differentiating this we can find the second derivative $f''(x) = 6x$

We now take each point in turn and use our test.

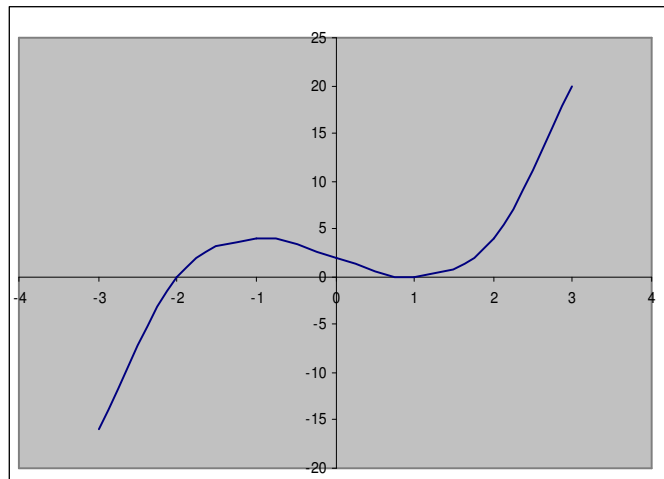
When $x = 1$: $f''(1) = 6$

We are not really interested in this value. What is important is its *sign*. Because it is *positive* we know we are dealing with a *minimum point*.

When $x = -1$: $f''(-1) = -6$

Again, what is important is its *sign*. Because it is *negative* we have a *maximum point*.

Finally, to finish this off we produce a quick sketch of the function now that we know the precise locations of its two turning points:



An example which uses *the first derivative to distinguish maxima and minima*

Example: Suppose we wish to find the turning points of the function $y = \frac{(x-1)^2}{x}$ and

distinguish between them.

First of all we need to find $f'(x)$.

In this case we need to apply the quotient rule for differentiation.

$$f'(x) = \frac{2(x-1)x - (x-1)^2}{x^2}$$

This does look complicated. Don't rush to multiply it all out if you can avoid it. Instead, look for common factors, and tidy up the expression.

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

We now set $f'(x)$ equal to zero in order to locate the stationary points including any

turning points: $\frac{(x-1)(x+1)}{x^2} = 0$

When equating a fraction to zero, it is the top line, the numerator, which must equal zero.

$$\text{Therefore } (x-1)(x+1) = 0$$

from which $x-1=0$ or $x+1=0$, and from these equations we find that $x=1$ or $x=-1$.

The y coordinates of the stationary points are found from $y = \frac{(x-1)^2}{x}$

When $x=1$: $y=0$.

When $x=-1$: $y=-4$.

We conclude that stationary points occur at $(1, 0)$ and $(-1, -4)$.

We now have to decide whether these are maximum points or minimum points. We could calculate $f''(x)$ and use the second derivative test as in the previous example. This would

involve differentiating $\frac{(x-1)(x+1)}{x^2}$ which is possible but perhaps rather fearsome! Is

there an alternative way? The answer is **yes**. We can look at how $f'(x)$ changes as we move through the stationary point. In essence, we can find out what happens to $f'(x)$ without actually calculating it.

First consider the point at $x=-1$. We look at what is happening a little bit before the point where $x=-1$, and a little bit afterwards. Often we express the idea of 'a little bit before' and 'a little bit afterwards' in the following way:

We can write $-1 - \varepsilon$ to represent a little bit less than -1 , and $-1 + \varepsilon$ to represent a little bit more. The symbol ε is the Greek letter *epsilon*. It represents a small positive quantity, say 0.1. Then $-1 - \varepsilon$ would be -1.1 , just a little less than -1 . Similarly $-1 + \varepsilon$ would be -0.9 , just a little more than -1 .

We now have a look at $f'(x)$; not its value, but its sign.

When $x = -1 - \varepsilon$, say -1.1 , $f'(x)$ is positive (i.e. before $x = -1$, the function $f(x)$ increases)

When $x = -1$ we already know that $f'(x) = 0$ (i.e. $x = -1$ is a stationary point, which may be a turning point)

When $x = -1 + \epsilon$, say -0.9 , $f'(x)$ is negative (i.e. after $x = -1$ the function $f(x)$ decreases)

We can summarise this information as follows: at $x = -1$ there is a local maximum, in other words the stationary point at $(-1, -4)$ is a maximum turning point.

Then we turn to the point $(1, 0)$. We carry out a similar analysis, looking at the sign of $f'(x)$ at $x = 1 - \epsilon$, $x = 1$, and $x = 1 + \epsilon$.

Signs of $f'(x)$ are $-$, 0 , $+$, which implies that the point is a minimum.

This, so-called first derivative test, is also the way to do it if $f''(x)$ is zero, in which case the second derivative test does not work.

