3 Differentiation

3.1 Derivative as a limit

Geometrically the derivative $f'(a)$ of a function $f(x)$ at $x = a$ is equal to the slope of the tangent to the graph of $f(x)$ at $x = a$.

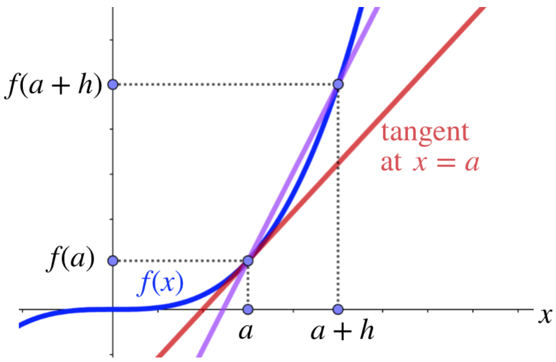

Figure 19: The graph of a function $f(x)$, the tangent at $x = a$ and the secant through the points on the graph at $x = a$ and $x = a + h$.

A secant is a line that intersects a curve at two points (the part of the secant that is between the two intersection points is called a chord).
Consider a secant that intersects the graph of $f(x)$ at the two (x, y) points $(a, f(a))$ and $(a + h, f(a + h))$. The slope of this secant is given by the difference quotient $(f(a + h) - f(a))/h$. As the two points approach each other ie. as $|h|$ decreases, the secant approaches the tangent at $x = a$. The tangent is obtained in the limit as $h \to 0$ and the derivative at a is the slope of the secant in this limit ie.

$$f'(a) = \lim_{h\to 0} \frac{f(a + h) - f(a)}{h}$$

providing this limit exists.
If this limit exists then we say that $f(x)$ **is differentiable at** $x = a$.
Eg. Use the limit definition of the derivative to calculate $f'(a)$ for $f(x) = x^2$.

$$f'(a) = \lim_{h\to 0} \frac{f(a + h) - f(a)}{h} = \lim_{h\to 0} \frac{(a + h)^2 - a^2}{h} = \lim_{h\to 0} \frac{2ah + h^2}{h} = \lim_{h\to 0}(2a + h) = 2a.$$

This is, of course, the expected result from the knowledge that $f'(x) = 2x$.

Eg. Use the limit definition of the derivative to calculate $f'(\pi)$ for $f(x) = \sin x$.

$$f'(\pi) = \lim_{h\to 0} \frac{f(\pi + h) - f(\pi)}{h} = \lim_{h\to 0} \frac{\sin(\pi + h) - \sin \pi}{h} = \lim_{h\to 0} \frac{\sin \pi \cos h + \cos \pi \sin h}{h}$$

$$= \lim_{h\to 0} \frac{-\sin h}{h} = -1$$

where the final equality uses the earlier important trigonometric limit result.
This is, of course, the expected answer from the knowledge that $f'(x) = \cos x$ giving $f'(\pi) = \cos\pi = -1$.

If $f'(a)$ exists for all a in Dom f we say that $f(x)$ **is differentiable** and then $f'(a)$ defines a function $f'(x)$ called the derivative.

Eg. Use the limit definition of the derivative to calculate the derivative of the function $f(x) = x\cos x$.

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{(x+h)\cos(x+h) - x\cos x}{h}$$

$$= \lim_{h\to 0} \frac{(x+h)(\cos x\cos h - \sin x \sin h) - x\cos x}{h}$$

$$= \lim_{h\to 0} \left(x\cos x \frac{\cos h - 1}{h} - x\sin x \frac{\sin h}{h} + \cos x\cos h - \sin x \sin h \right) = -x\sin x + \cos x$$

where we have used both of the earlier important trigonometric limit results.

The fact that the derivative is equal to the slope of the tangent of the graph allows us to determine a Cartesian equation for the tangent by calculating the derivative. Explicitly, the tangent line at the point $(a, f(a))$ is given by $y = f(a) + f'(a)(x - a)$.

Eg. For the function $f(x) = 4x^3 - x$, find a Cartesian equation for the tangent to the graph at the point $(1, 3)$.
$f'(x) = 12x^2 - 1$ so $f'(1) = 11$. The tangent line is given by
$y = f(1) + f'(1)(x-1) = 3 + 11(x-1) = 11x - 8$

A necessary condition for a function to be differentiable at a point is that it is continuous at that point, but this is not a sufficient condition.

Geometrically, there are two ways that a function can fail to be differentiable at a point where it is continuous:

(i). The tangent line is vertical at that point.

Eg. The function $f(x) = x^{1/3}$ is continuous in $\mathbb{R}$ but it is not differentiable at $x = 0$. The difference quotient at $x = 0$ is

$$\frac{f(0+h) - f(0)}{h} = \frac{h^{1/3}}{h} = \frac{1}{h^{2/3}}$$

and no limit exists as $h \to 0$ because the above grows without bound.

(ii). There is no tangent line at that point.

Eg. The function $f(x) = |x|$ is continuous in $\mathbb{R}$ but it is not differentiable at $x = 0$. The difference quotient at $x = 0$ is

$$\frac{f(0+h)-f(0)}{h} = \frac{|h|}{h} = \begin{cases} -1 & \text{if } h < 0 \\ 1 & \text{if } h > 0 \end{cases}$$

The left and right limits therefore both exist

$$\lim_{h\to 0^+} \frac{f(0+h)-f(0)}{h} = 1$$

and

$$\lim_{h\to 0^-} \frac{f(0+h)-f(0)}{h} = -1$$

but these are not equal, so

$$\lim_{h\to 0} \frac{f(0+h)-f(0)}{h}$$

does not exist.

3.2 The Leibniz and chain rules

If $f(x)$ is differentiable then its derivative $f'(x)$ may also be differentiable, in which case we denote its derivative by $f''(x)$ ie. the **second derivative** of $f(x)$. Similarly the third derivative is $f'''(x)$ and in general the n^{th} derivative is $f^{(n)}(x)$ so that $f^{(3)}(x) = f'''$. This notation is due to Lagrange (1736-1813), but other notations are also common.

$f'(x) = \frac{df}{dx},\ f''(x) = \frac{d^2f}{dx^2},\ f^{(n)}(x) = \frac{d^nf}{dx^n}$ is due to Leibniz (1646-1716).

$f'(x) = Df(x),\ f''(x) = D^2f(x),\ f^{(n)}(x) = D^nf(x)$ is due to Euler (1707-1783).

Finally, if the independent variable represents time, say $f(t)$, then $f'(t) = \dot{f}$ and $f''(t) = \ddot{f}$ is due to Newton (1642-1727). As we have seen, the derivative $\frac{df}{dx}$ is the slope of the tangent to the graph of $f(x)$. This means that the derivative $\frac{df}{dx}$ measures the rate of change of the dependent variable f with respect to changes in the independent variable x. In the case that the independent variable is time t, the derivative measures the rate of change with time, so if the dependent variable, say $X(t)$ represents the position of an object confined to a line then $\dot{X}$ is the velocity of the object, with $|\dot{X}|$ the speed, and $\ddot{X}$ is the acceleration of the object.

The product rule for differentiation is just the first case of the more general **Leibniz rule**:
If $f(x)$ and $g(x)$ are both differentiable n times then so is the product $f(x)g(x)$ with

$$D^n(fg) = \sum_{k=0}^{n} \binom{n}{k} (D^kf)(D^{n-k}g)$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is the binomial coefficient familiar from Pascal's triangle

$$\begin{array}{ccccccccc} & & & & \binom{0}{0} & & & & \\ & & & \binom{1}{0} & & \binom{1}{1} & & & \\ & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\ & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\ \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \end{array} \qquad \begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \end{array}$$

Figure 20: The binomial coefficients arranged into Pascal's triangle.

Eg. Use the Leibniz rule to calculate $D^3(x^2 \sin x)$ using
$D^3(fg) = f(D^3g) + 3(Df)(D^2g) + 3(D^2f)(Dg) + (D^3f)g.$
$f = x^2, \quad Df = 2x, \quad D^2f = 2, \quad D^3f = 0.$
$g = \sin x, \quad Dg = \cos x, \quad D^2g = -\sin x, \quad D^3g = -\cos x.$
$D^3(x^2 \sin x) = -x^2 \cos x - (3)2x \sin x + (3)2 \cos x + 0 \sin x = (6 - x^2) \cos x - 6x \sin x.$

To handle the differentiation of composite functions $(f\circ g)(x)=f(g(x))$ we turn to the following theorem:

The **chain rule theorem** states that if $g(x)$ is differentiable at x and $f(x)$ is differentiable at $g(x)$ then the composition $(f\circ g)(x)$ is differentiable at x with

$$(f\circ g)'(x)=f'(g(x))g'(x).$$

Recall that prime denotes differentiation with respect to the argument so in Leibniz notation the above formula may be written more crudely as

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

where we need to be aware that on the right hand side the argument of the first factor is $g(x)$ and the argument of the second factor is x.
This form is useful as a mnemonic (think of 'cancelling' the dg factor) but it is no more than this.

Eg. Calculate $\dfrac{d}{dx}\left(\left(x+\dfrac{1}{x}\right)^{-3}\right)$.

In terms of the above notation $g(x)=x+\frac{1}{x}$ and $f(x)=x^{-3}$
giving $g'(x)=1-\frac{1}{x^2}$ and $f'(x)=-3x^{-4}$ thus

$$\frac{d}{dx}\left(\left(x+\frac{1}{x}\right)^{-3}\right)=f'(g(x))g'(x)=-3\left(x+\frac{1}{x}\right)^{-4}\left(1-\frac{1}{x^2}\right).$$

Eg. $\frac{d}{dt}\cos(t^2)=-2t\sin(t^2)$.

3.3 L'Hôpital's rule

*** Only use this method in assignments/exam questions when told to do so ***

Recall that if $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M \neq 0$ then $\lim_{x\to a} f(x)/g(x) = L/M$. Clearly this result is not applicable to the situation where $L = M = 0$, which is called an **indeterminant form**. We have already seen how to deal with situations like this directly, but an alternative (and often easier) approach is sometimes available by making use of the following.

L'Hôpital's rule
Let $f(x)$ and $g(x)$ be differentiable on $I = (a-h, a) \cup (a, a+h)$ for some $h > 0$, with $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$.
If $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists and $g'(x) \neq 0 \ \forall\ x \in I$ then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Proof: We shall only consider the proof for the slightly easier situation in which $f(x)$ and $g(x)$ are both differentiable at $x = a$ and $g'(a) \neq 0$.
In this case

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f(x)-f(a)}{g(x)-g(a)} = \lim_{x\to a} \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} = \frac{\lim_{x\to a} \frac{f(x)-f(a)}{x-a}}{\lim_{x\to a} \frac{g(x)-g(a)}{x-a}} = \frac{f'(a)}{g'(a)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

We can apply L'Hôpital's rule to calculate the two important trigonometric limits from earlier.

Eg. Calculate $\lim_{x\to 0} \frac{\sin x}{x}$.
$f(x) = \sin x$ and $g(x) = x$ are both differentiable. Also $\lim_{x\to 0} \sin x = \sin 0 = 0$ and $\lim_{x\to 0} x = 0$. Furthermore, $f'(x) = \cos x$ and $g'(x) = 1 \neq 0$. Thus L'Hôpital's rule applies to give

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

Eg. Calculate $\lim_{x\to 0} \frac{1-\cos x}{x}$.
$f(x) = 1 - \cos x$ and $g(x) = x$ are both differentiable. Also $\lim_{x\to 0}(1 - \cos x) = 1 - \cos 0 = 0$ and $\lim_{x\to 0} x = 0$. Furthermore, $f'(x) = \sin x$ and $g'(x) = 1 \neq 0$. Thus by L'Hôpital's rule

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

If applying L'Hôpital's rule yields another indeterminant form then L'Hôpital's rule can be reapplied to this form and so on.

Eg. Calculate $\lim_{x\to 0} \frac{2\sin x - \sin(2x)}{x - \sin x}$.

$$\begin{aligned} \lim_{x\to 0} \frac{2\sin x - \sin(2x)}{x - \sin x} &= \lim_{x\to 0} \frac{2\cos x - 2\cos(2x)}{1 - \cos x} = \lim_{x\to 0} \frac{-2\sin x + 4\sin(2x)}{\sin x} \\ &= \lim_{x\to 0} \frac{-2\cos x + 8\cos(2x)}{\cos x} = \frac{-2+8}{1} = 6. \end{aligned}$$

3.4 Boundedness and monotonicity

The following definitions apply to a function $f(x)$ defined in some interval I.

Defn: If $\exists$ a constant k_1 s.t. $f(x) \leq k_1 \ \ \forall \ \ x$ in I we say that $f(x)$ is **bounded above in** I and we call k_1 an **upper bound** of $f(x)$ in I.

Furthermore, if $\exists$ a point x_1 in I s.t. $f(x_1) = k_1$ we say that the upper bound k_1 is **attained** and we call k_1 the **global maximum value** of $f(x)$ in I.

Similarly, we have the following:

Defn: If $\exists$ a constant k_2 s.t. $f(x) \geq k_2 \ \ \forall \ \ x$ in I we say that $f(x)$ is **bounded below in** I and we call k_2 a **lower bound** of $f(x)$ in I.

Furthermore, if $\exists$ a point x_2 in I s.t. $f(x_2) = k_2$ we say that the lower bound k_2 is **attained** and we call k_2 the **global minimum value** of $f(x)$ in I.

Defn: $f(x)$ is **bounded in** I if it is both bounded above and bounded below in I ie. if $\exists$ a constant k s.t. $|f(x)| \leq k \ \ \forall \ \ x$ in I.

If no interval I is specified then it is taken to be the domain of the function.

Eg. $\cos x$ is bounded (in $\mathbb{R}$) because $|\cos x| \leq 1 \ \forall x \in \mathbb{R}$. Both these upper and lower bounds are attained as $\cos 0 = 1$ and $\cos \pi = -1$. The global maximum value in $\mathbb{R}$ is 1 and the global minimum value in $\mathbb{R}$ is -1.

Eg. Consider $f(x) = \frac{\mathrm{sgn}(x)}{1+x^2}$ for $-1 \leq x \leq 1$. Then $f(x)$ is bounded in $[-1, 1]$ because $|f(x)| \leq 1$ for all x in $[-1, 1]$ but neither of the bounds is attained, so it has no global maximum value and no global minimum value in $[-1, 1]$.

Eg. On $[0, \pi/2)$ the function $\tan x$ is bounded below but not bounded above.

A condition that guarantees the existence of both a global maximum and minimum value (called **extreme values**) is provided by the following:

The extreme value theorem states that if f is a continuous function on a *closed* interval $[a, b]$ then it is bounded on that interval and has upper and lower bounds that are attained ie. $\exists$ points x_1 and x_2 in $[a, b]$ such that $f(x_2) \leq f(x) \leq f(x_1) \ \forall \ x \in [a, b]$.

Defn: $f(x)$ is **monotonic increasing** in $[a, b]$ if $f(x_1) \leq f(x_2)$ for all pairs x_1, x_2 with $a \leq x_1 < x_2 \leq b$.

Defn: $f(x)$ is **strictly monotonic increasing** in $[a, b]$ if $f(x_1) < f(x_2)$ for all pairs x_1, x_2 with $a \leq x_1 < x_2 \leq b$.

Obvious definitions apply with increasing replaced by decreasing.

Eg. sgn(x) is monotonic increasing in $[-1, 1]$ but not strictly.

Eg. x^2 is strictly monotonic increasing in $[0, b]$ for any $b > 0$.

3.5 Critical points

Defn: We say that $f(x)$ has a **local maximum** at $x = a$ if $\exists\, h > 0$ s.t. $f(a) \geq f(x) \;\; \forall\, x \in (a - h, a + h)$.

Remark: If $f(x)$ has a local maximum at $x = a$ and is differentiable at this point then $f'(a) = 0$.

Proof: As $f(x)$ is differentiable at a then

$$\lim_{x\to a^-} \frac{f(x) - f(a)}{x - a} = \lim_{x\to a^+} \frac{f(x) - f(a)}{x - a} = f'(a).$$

$f(x)$ has a local maximum at a implies

$$\frac{f(x) - f(a)}{x - a} \geq 0 \quad \text{for } x \in (a - h, a)$$

$$\text{hence} \;\; f'(a) = \lim_{x\to a^-} \frac{f(x) - f(a)}{x - a} \geq 0.$$

Similarly, $f(x)$ has a local maximum at a also implies

$$\frac{f(x) - f(a)}{x - a} \leq 0 \quad \text{for } x \in (a, a + h)$$

$$\text{hence} \;\; f'(a) = \lim_{x\to a^+} \frac{f(x) - f(a)}{x - a} \leq 0.$$

Putting these two together gives $f'(a) \leq 0 \leq f'(a)$ and thus $f'(a) = 0$.

Defn: We say that $f(x)$ has a **local minimum** at $x = a$ if $\exists\, h > 0$ s.t. $f(x) \geq f(a) \;\; \forall\, x \in (a - h, a + h)$.

Remark: If $f(x)$ has a local minimum at $x = a$ and is differentiable at this point then $f'(a) = 0$.

Defn: $f(x)$ has a **stationary point** at $x = a$ if it is differentiable at $x = a$ with $f'(a) = 0$.

Defn: An interior point $x = a$ of the domain of $f(x)$ is called a **critical point** if either $f'(a) = 0$ or $f'(a)$ does not exist.

Every local maximum and local minimum of a differentiable function is a stationary point but there may be other stationary points too.

Eg. $f(x) = x^3$ has $f'(0) = 0$ so $x = 0$ is a stationary point but it is neither a local maximum nor a local minimum because $f(x) > f(0)$ for $x \in (0, h)$ but $f(x) < f(0)$ for $x \in (-h, 0)$.

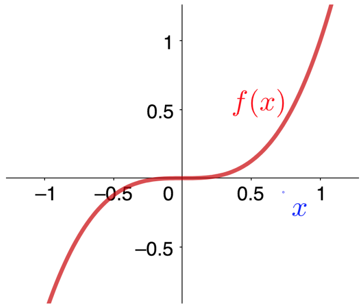

Figure 21: The graph of $f(x) = x^3$ indicating the point of inflection at $x = 0$.

Defn: If $f(x)$ is twice differentiable in an open interval around $x = a$ with $f''(a) = 0$ and if $f''(x)$ changes sign at $x = a$ then we say that $x = a$ is a **point of inflection.**

Eg. $f(x) = x^3$ has $f''(x) = 6x$ so $f''(0) = 0$. As $f''(x) < 0$ for $x < 0$ and $f''(x) > 0$ for $x > 0$ then $f''(x)$ changes sign at $x = 0$ so it is a point of inflection.

The first derivative test
Suppose $f(x)$ is continuous at a critical point $x = a$.

(i). If $\exists\, h > 0$ s.t. $f'(x) < 0\ \forall\ x \in (a - h, a)$ and $f'(x) > 0\ \forall\ x \in (a, a + h)$ then $x = a$ is a local minimum.

(ii). If $\exists\, h > 0$ s.t. $f'(x) > 0\ \forall\ x \in (a - h, a)$ and $f'(x) < 0\ \forall\ x \in (a, a + h)$ then $x = a$ is a local maximum.

(iii). If $\exists\, h > 0$ s.t. $f'(x)$ has a constant sign $\forall\ x \neq a$ in $(a - h, a + h)$ then $x = a$ is a not a local extreme value (minimum/maximum).

Eg. $f(x) = |x|$ is continuous but $f'(x) \neq 0$ so there are no stationary points. The derivative does not exist at $x = 0$ so this is a critical point. $f'(x) < 0$ for $x \in (-1, 0)$ and $f'(x) > 0$ for $x \in (0, 1)$ so by the first derivative test $x = 0$ is a local minimum. In fact $f(0) = 0$ is a global minimum as $|x| \geq 0$.

Eg. $f(x) = x^4 - 2x^3$ has $f'(x) = 4x^3 - 6x^2 = 2x^2(2x - 3)$. The only critical points are $x = 0, 3/2$. Now $f'(x) < 0$ for $x < 3/2$ and $f'(x) > 0$ for $x > 3/2$. Thus $f'(x)$ has constant sign on both (sufficiently close) sides of $x = 0$ so this is not a local extreme value. However, $f'(x)$ changes sign from negative to positive as x passes through $x = 3/2$ so this is a local minimum.

The second derivative test
Suppose $f(x)$ is twice differentiable at $x = a$ with $f'(a) = 0$.
(i). If $f''(a) > 0$ then $x = a$ is a local minimum.
(ii). If $f''(a) < 0$ then $x = a$ is a local maximum.

Note that the theorem doesn't say anything about the case $f''(a) = 0$.

Eg. $f(x) = 2x^3 - 9x^2 + 12x$ has $f'(x) = 6x^2 - 18x + 12 = 6(x^2 - 3x + 2) = 6(x-1)(x-2)$ giving stationary points at $x = 1$ and $x = 2$. Furthermore, $f''(x) = 6(2x - 3)$ giving $f''(1) = -6 < 0$ and $f''(2) = 6 > 0$ so by the second derivative test $x = 1$ is a local maximum and $x = 2$ is a local minimum.

Note that the first derivative test is more general than the second derivative test as it does not require the function to be differentiable at the critical point.

Determining all critical points and asymptotes is the key to sketching the graph of a function.
Eg. Sketch the graph of the function $f(x) = \frac{4x-5}{x^2-1}$.

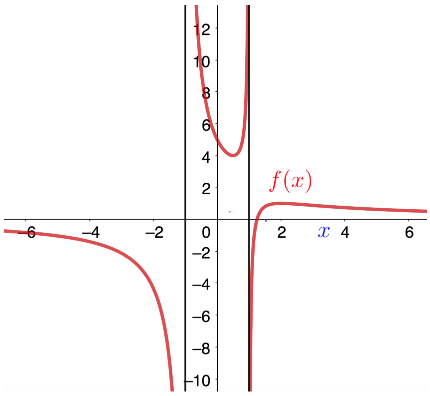

Figure 22: Graph of $f(x) = \frac{4x-5}{x^2-1}$.

$y = 0$ is a horizontal asymptote along both the positive and negative x-axis because

$$\lim_{x\to\pm\infty} \frac{4x-5}{x^2-1} = \lim_{x\to\pm\infty} \frac{1}{x}\left(\frac{4 - 5/x}{1 - \frac{1}{x^2}}\right) = 0\left(\frac{4-0}{1-0^2}\right) = 0$$

There are vertical asymptotes at $x = \pm 1$. Also

$$f'(x) = \frac{4(x^2-1) - 2x(4x-5)}{(x^2-1)^2} = \frac{-4x^2 + 10x - 4}{(x^2-1)^2} = \frac{-(4x-2)(x-2)}{(x^2-1)^2}$$

so there are stationary points at $x = \frac{1}{2}, 2$ with $f(\frac{1}{2}) = 4$ and $f(2) = 1$.

x		-1		$\frac{1}{2}$		1		2	
f'	$-$		$-$		$+$		$+$		$-$

As x passes through $\frac{1}{2}$ then $f'(x)$ changes sign from negative to positive so this is a local minimum. As x passes through 2 then $f'(x)$ changes sign from positive to negative so this is a local maximum.

If a function is defined on an interval then extreme values can occur at the endpoints of the interval. Here an **endpoint** $x = c$ is a point at which the function is defined but where the function is undefined either to the left or right of this point. For example, if the domain of a function is $[a, b]$ then a and b are both endpoints.

Defn: If c is an endpoint of $f(x)$ then f has an **endpoint maximum** at $x = c$ if $f(x) \leq f(c)$ for x sufficiently close to c.

The obvious similar definition applies for an **endpoint minimum**.

If a function is differentiable sufficiently close to an endpoint then examining the sign of the derivative can determine if there is an endpoint maximum or minimum.

If $f(x)$ is continuous on an interval $[a, b]$ then the global extreme values in this interval are attained at either critical points or endpoints, so we can determine them by examining all the possibilities.

Eg. Find the global extreme values of $f(x) = 1 + 4x^2 - \frac{1}{2}x^4$ for $x \in [-1, 3]$.

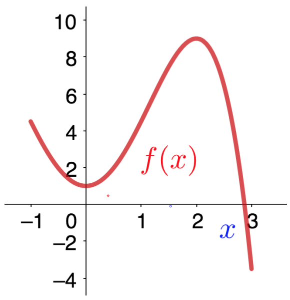

Figure 23: Graph of $f(x) = 1 + 4x^2 - \frac{1}{2}x^4$

$f'(x) = 8x - 2x^3 = 2x(4 - x^2)$ which is zero at $x = 0, 2$; recall we only consider $x \in [-1, 3]$.
Critical point: $f(0) = 1$ and $f(2) = 9$
End points: $f(-1) = \frac{9}{2}$ and $f(3) = -\frac{7}{2}$.
The smallest of these four values is $-\frac{7}{2}$, which is the global minimum, whereas the largest is 9 which is the global maximum.

Eg. Find the global extreme values of $x^2 - 2|x| + 2$ for $x \in [-\frac{1}{2}, 2]$.
First of all we can write

$$f(x) = \begin{cases} x^2 + 2x + 2 & \text{if} \quad -\frac{1}{2} \leq x < 0 \\ x^2 - 2x + 2 & \text{if} \quad 0 \leq x \leq 2 \end{cases}$$

$f'(x) = 2x + 2$ for $x \in [-\frac{1}{2}, 0)$ so no critical points here.
$f'(x) = 2x - 2$ for $x \in (0, 2]$ so a critical point at $x = 1$ with $f(1) = 1$.
$\lim_{x \to 0^+} f'(x) = -2 \neq 2 = \lim_{x \to 0^-} f'(x)$ hence $f'(x)$ does not exist at $x = 0$, making this a critical point with $f(0) = 2$.
Endpoints: $f(-\frac{1}{2}) = \frac{5}{4}$ and $f(2) = 2$.
Thus the global minimum is 1 and the global maximum is 2.

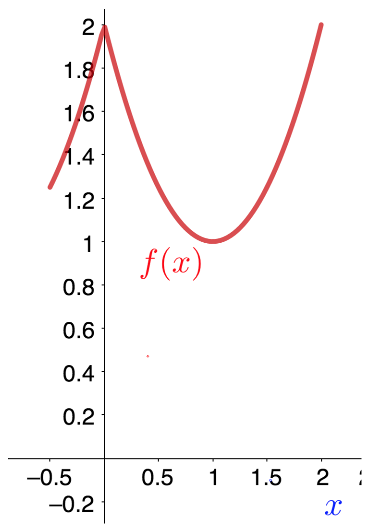

Figure 24: Graph of a piecewise function with two quadratic pieces.

3.6 Rolle's theorem

Rolle's theorem states that if f is differentiable on the open interval (a, b) and continuous on the closed interval $[a, b]$, with $f(a) = f(b)$, then there is at least one $c \in (a, b)$ for which $f'(c) = 0$.

Proof: By the extreme value theorem $\exists$ x_1 and x_2 in $[a, b]$ s.t.
$f(x_1) \le f(x) \le f(x_2)$ $\forall$ $x \in [a, b]$.
If $x_1 \in (a, b)$ then x_1 is a local minimum and $f'(x_1) = 0$ so we are done.
If $x_2 \in (a, b)$ then x_2 is a local maximum and $f'(x_2) = 0$ so we are done.
The only case that is left is if both x_1 and x_2 are endpoints, a, b. But since $f(a) = f(b)$ then in this case $f(x_1) = f(x_2) = f(a)$ so the above bound becomes $f(a) \le f(x) \le f(a)$ $\forall$ $x \in [a, b]$. Thus $f(x) = f(a)$ is constant in the interval and $f'(x) = 0$ $\forall$ $x \in [a, b]$.

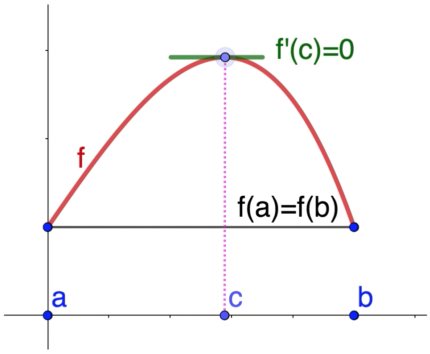

Figure 25: Illustration of Rolle's theorem.

An obvious corollary of Rolle's theorem is that if $f(x)$ is differentiable at every point of an open interval I then each pair of zeros of the function in I is separated by at least one zero of $f'(x)$. This is simply the special case of Rolle's theorem with a and b taken to be the the zeros of $f(x)$ so that $f(a) = f(b) = 0$.

An application of the above result is to provide a limit on the number of distinct real zeros of a function.

Eg. Show that $f(x) = \frac{1}{7}x^7 - \frac{1}{5}x^6 + x^3 - 3x^2 - x$ has no more than 3 distinct real roots.
Suppose the required result is false and there are (at least) 4 distinct real roots of $f(x)$. Then by Rolle's theorem there are (at least) 3 distinct real roots of $f'(x)$. Then by applying Rolle's theorem to $f'(x)$ there are (at least) 2 distinct real roots of $f''(x)$. However, $f''(x) = 6x^5 - 6x^4 + 6x - 6 = 6(x-1)(x^4+1)$ has only one real root (at $x = 1$). This contradiction proves the required result.

3.7 The mean value theorem

The mean value theorem states that if f is differentiable on the open interval (a, b) and continuous on the closed interval $[a, b]$, then there is at least one $c \in (a, b)$ for which

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

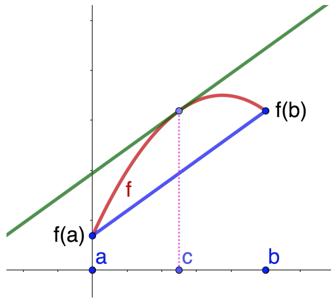

Figure 26: Illustration of the mean value theorem.

Geometrically this theorem states that there is at least one point in the interval at which the tangent to the curve is parallel to the line joining the two points $(b, f(b))$ and $(a, f(a))$ at the ends of the interval. The geometry makes the result clear but here is the proof.
Proof: Define the function

$$g(x) = (b - a)(f(x) - f(a)) - (x - a)(f(b) - f(a)).$$

Then $g(a) = g(b) = 0$ and as $g(x)$ satisfies the requirements of Rolle's theorem then $\exists\, c \in (a, b)$ for which $g'(c) = 0$. As $g'(x) = (b - a)f'(x) - (f(b) - f(a))$ then setting $x = c$ yields the required result $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Note that Rolle's theorem is just a special case of the mean value theorem with $f(a) = f(b)$.

The mean value theorem can be used to prove some obvious results relating monotonicity to the sign of the derivative. Here is an example:

Suppose $f(x)$ is continuous in $[a, b]$ and differentiable in (a, b) with $f'(x) \geq 0$ throughout this interval. Then $f(x)$ is monotonic increasing in (a, b).
Proof: If $a \leq x_1 < x_2 \leq b$ then by the mean value theorem $\exists\, c \in (a, b)$ s.t. $f'(c) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. However, since $f'(x) \geq 0$ in (a, b) then $f'(c) \geq 0$ which imples that $f(x_2) \geq f(x_1)$ ie. f is monotonic increasing.

Similar obvious results and proofs follow for the cases of monotonic decreasing and strictly monotonic increasing/decreasing.

3.8 The inverse function rule

The derivative of the inverse of a function can be obtained from the following:

The inverse function rule states that if $f(x)$ is continuous in $[a,b]$ and differentiable in (a,b) with $f'(x) > 0$ throughout this interval then its inverse function $g(y)$ (recall $g(f(x)) = x$) is differentiable for all $f(a) < y < f(b)$ with $g'(y) = 1/f'(g(y))$.

Note: A similar theorem exists for the case $f'(x) < 0$.

Eg. $f(x) = \sin x$ has $f'(x) = \cos x > 0$ for $x \in (-\pi/2, \pi/2)$.
Its inverse function $g(y) = \sin^{-1} y$ is therefore differentiable in $(-1, 1)$ with

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{\cos g(y)} = \frac{1}{\sqrt{1 - \sin^2 g(y)}}$$

but $y = f(g(y)) = \sin(g(y))$ hence $g'(y) = \frac{1}{\sqrt{1-y^2}}$. Thus we have shown that

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

Eg. $f(x) = \tan x$ has $f'(x) = \sec^2 x > 0$ for $x \in (-\pi/2, \pi/2)$.
With this domain then Ran $f = \mathbb{R}$ hence the domain of the inverse function $g(y) = \tan^{-1}(y)$ is $\mathbb{R}$.
By the inverse function rule

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{\sec^2(g(y))} = \frac{1}{1 + \tan^2(g(y))} = \frac{1}{1 + y^2}$$

where we have used $y = f(g(y)) = \tan(g(y))$.
Thus we have shown that

$$\frac{d}{dx} \tan^{-1} x = \frac{1}{1 + x^2}.$$

3.9 Partial derivatives

So far we have only considered functions of a single variable eg. $f(x)$. Later we shall study functions of two variables eg. $f(x,y)$. For this situation we shall need the concept of a **partial derivative**. The partial derivative of f with respect to x is written as $\frac{\partial f}{\partial x}$ and is obtained by differentiating f with respect to x, keeping y fixed ie. y may be treated as a constant in performing the partial differentiation with respect to x.

$$\text{Eg.}\quad f(x,y) = x^2y^3 - \sin x\cos y\, + y \quad \text{has} \quad \frac{\partial f}{\partial x} = 2xy^3 - \cos x\cos y.$$

Similarly, the partial derivative of f with respect to y is written as $\frac{\partial f}{\partial y}$ and is obtained by differentiating f with respect to y, keeping x fixed.

$$\text{Eg.}\quad f(x,y) = x^2y^3 - \sin x\cos y\, + y \quad \text{has} \quad \frac{\partial f}{\partial y} = 3x^2y^2 + \sin x\sin y\, + 1.$$

Partial derivatives are defined in terms of the following limits (which we assume exist)

$$\frac{\partial f}{\partial x}(x,y) = \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h}, \quad \frac{\partial f}{\partial y}(x,y) = \lim_{h\to 0}\frac{f(x,y+h)-f(x,y)}{h}.$$

An alternative notation for partial derivatives is to write f_x for $\frac{\partial f}{\partial x}$ etc.

By applying partial differentiation to these partial derivatives we may obtain the second order partial derivatives

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial y\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial x\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right).$$

For the earlier example $f(x,y) = x^2y^3 - \sin x\cos y\, + y$ we obtain

$$\frac{\partial^2 f}{\partial x^2} = 2y^3 + \sin x\cos y, \quad \frac{\partial^2 f}{\partial y^2} = 6x^2y + \sin x\cos y,$$
$$\frac{\partial^2 f}{\partial y\partial x} = 6xy^2 + \cos x\sin y, \quad \frac{\partial^2 f}{\partial x\partial y} = 6xy^2 + \cos x\sin y.$$

Note the equality, in this example, of the two mixed partial derivatives

$$\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\partial x}.$$

It can be shown that this result is true in general, if f and all first and second order partial derivatives are continuous.

An alternative notation for partial derivatives is

$$\frac{\partial^2 f}{\partial x^2} = f_{xx}$$

etc