Chapter 4 Differentiable Functions

Introduction.

Differentiation is usually associated with the rate of change of a quantity or process. Suppose the process or quantity is depended on some variable x, usually this is time. We may consider it as a function $f(x)$. Fix an x, called it x_0. Then suppose at some other value of x, say x_n, the value of the process or quantity is $f(x_n)$ then we may say the *rate of change* is the quotient

$$\frac{f(x_n)-f(x_0)}{x_n-x_0}. \qquad \text{------------------------} \quad (1)$$

Usually this is the quantity we seek as in application in Revenue in Business, where it is also known as *average rate of change* from x_0 to x_n. Very often we require the so called *instantaneous rate of change* at x_0 as in marginal analysis in Business and Economics and as in velocity in mechanics. In practice, we normally calculate this quantity for some value x_n very close to x_0. Mathematically, this means we take a sequence (x_n) which converges to x_0 and consider the limit

$$\lim_{n\to\infty}\frac{f(x_n)-f(x_0)}{x_n-x_0}. \qquad \text{----------------} \quad (2)$$

If this limit exists, it is a "*derived number*" of the function f. Equivalently the limit

$$\lim_{n\to\infty}\frac{f(x_0+h_n)-f(x_0)}{h_n},$$

where $h_n = x_n - x_0$, is a derived number of f. For sufficiently small value of h_n the quotient

$$\frac{f(x_0+h_n)-f(x_0)}{h_n} \qquad \text{---------------} \ (3)$$

is an approximation of the derived number of f at x_0. For meaningful application, we assume that the limit (2) is unique for any sequence (x_n) converging to x_0 or equivalently that the limit (3) is unique for any sequence (h_n) converging to 0. Without this assumption it is possible for different sequences giving rise to different derived numbers. This assumption expresses the property that f must have, equivalently expressed by Definition 32 Chapter 3, that the limit

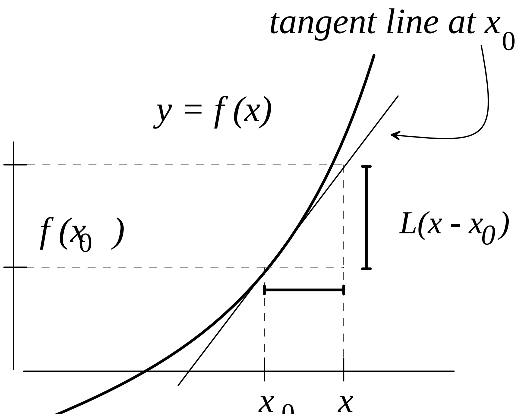
$$\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}$$

exists.

4.1 Differentiability

A subset D of $\mathbf{R}$ is a *neighbourhood* of a point x_0 if D contains an open interval I such that $x_0 \in I \subseteq D$.

Definition 1. Suppose $f : D \to \mathbf{R}$ is a function and D is a neighbourhood of x_0. (This means there exists an open interval I such that $x_0 \in I \subseteq D$.) Then we say f is differentiable at x_0 if the limit $\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}$ exists and is finite. This limit is called the derivative of f at x_0. We denote this limit by $f'(x_0)$. Note that in its equivalent form, $f'(x_0) = \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$.



Remark. Note that x_0 is a limit point (or cluster point) of $I-\{x_0\}$ and so is a limit point of D. Hence the limit above is given by Definition 32 of Chapter 3.
Suppose $f'(x_0) = L$. Then $\lim_{x\to x_0}\frac{f(x)-f(x_0)-L(x-x_0)}{x-x_0}=0$. Thus, given any $\varepsilon > 0$, there exists $\delta > 0$ such that for all x in I,

$$0<|x-x_0|<\delta \Rightarrow |f(x)-f(x_0)-L(x-x_0)|<|x-x_0|\varepsilon.$$

Thus, if we write $r(x) = f(x) - f(x_0) - L(x - x_0)$, then $f(x) = f(x_0) + L(x - x_0) + r(x)$, where $\lim_{x\to x_0}\frac{r(x)}{x-x_0}=0$. Thus, in a small neighbourhood of x_0, $f(x)$ is approximated by the linear function $Lx + f(x_0) - L x_0$ whose graph is the tangent line at x_0.

If D is open and hence is a neighbourhood of each of its points and if $f : D \to \mathbf{R}$ is differentiable at x for all x in D, then we say f is *differentiable* on D and the function $f' : D \to \mathbf{R}$ is called the derived function or the derivative of f. If the limit $\lim_{x\to x_0}\frac{f'(x)-f'(x_0)}{x-x_0}$ exists, then we say f is *twice differentiable* at x_0 and write $f''(x_0) = \lim_{x\to x_0}\frac{f'(x)-f'(x_0)}{x-x_0}$. This is called the *second derivative* of f at x_0.

Example 2.

1. Suppose the function $f : \mathbf{R} \to \mathbf{R}$ is given by $f(x) = x^2$. Take any a in $\mathbf{R}$.
   Then for any $x \neq a$, $\frac{f(x)-f(a)}{x-a}=\frac{x^2-a^2}{x-a}=\frac{(x-a)(x+a)}{x-a}=(x+a)$.
   Thus, $\lim_{x\to a}\frac{f(x)-f(a)}{x-a} = \lim_{x\to a}(x+a) = 2a$ Therefore, $f'(a) = 2a$.

Note that in the above examples, we write the function $g(x)=\frac{f(x)-f(a)}{x-a}$ for $x \neq a$ in a form for which we can easily compute the limit. It is a simple observation that if $g(x) = h(x)$ for $x \neq a$, then $\lim_{x\to a} g(x) = \lim_{x\to a} h(x)$.

2. Let the function $f : (0, \infty) \to \mathbf{R}$ be given by $f(x)=\frac{1}{x}$. Take any x_0 in $(0, \infty)$.
   Then

$$f'(x_0)=\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}=\lim_{h\to 0}\frac{\frac{1}{x_0+h}-\frac{1}{x_0}}{h}=\lim_{h\to 0}\frac{\frac{x_0-(x_0+h)}{(x_0+h)x_0}}{h}$$

$$= \lim_{h\to 0} \frac{-1}{(x_0+h)x_0} = -\frac{1}{x_0^2}.$$

An immediate consequence of differentiability is continuity.

Theorem 3. Suppose f is defined on an open interval I containing a point a. If f is *differentiable* at a, then f is *continuous* at a.

Proof. For $x \neq a$ let $f(x) = f(a) + \frac{f(x)-f(a)}{x-a}\cdot(x-a)$. Then

$$\lim_{x\to a} f(x) = f(a) + \lim_{x\to a}\left(\frac{f(x)-f(a)}{x-a}\cdot(x-a)\right).$$

As a result that f is differentiable at a,

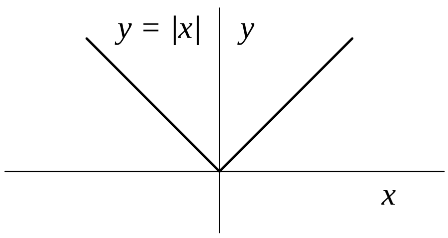
$$\lim_{x\to a}\left(\frac{f(x)-f(a)}{x-a}\cdot(x-a)\right) = \lim_{x\to a}\frac{f(x)-f(a)}{x-a}\lim_{x\to a}(x-a) = f'(a)\cdot 0 = 0.$$

Thus $\lim_{x\to a} f(x) = f(a) + 0 = f(a)$. Therefore, f is continuous at a.

Example 4.

1. The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ is *not* differentiable at x = 0 since f is not continuous at x = 0 (otherwise, by Theorem 3, f would be continuous at x = 0).

2. The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = |x|$ is *not* differentiable at x = 0 even though f is continuous at x = 0. This is deduced as follows. $\lim_{h\to 0^+}\frac{f(0+h)-f(0)}{h} = \lim_{h\to 0^+}\frac{h-0}{h} = 1$ and $\lim_{h\to 0^-}\frac{f(0+h)-f(0)}{h} = \lim_{h\to 0^-}\frac{-h-0}{h} = -1.$ Thus the right limit is not equal to the left limit. And so the limit does not exist. This means $|x|$ is not differentiable at 0.



The Graph of $y = |x|$.

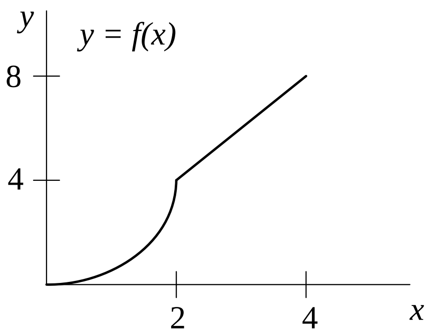
3. Let $f : \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = \begin{cases} 2x, & x \geq 2 \\ x^2, & x < 2 \end{cases}$. Then f is differentiable on $(-\infty, 2)$ and on $(2, \infty)$. f is continuous at 2 but not differentiable at x = 2. (Note that because $\lim_{x\to 2^+} f(x) = \lim_{x\to 2^+} 2x = 4$ and $\lim_{x\to 2^-} f(x) = \lim_{x\to 2^-} x^2 = 4$, both left and right limits are the same. Hence, $\lim_{x\to 2} f(x) = 4 = f(2)$. Hence f is continuous at x = 2.) For differentiability we look at the following left and right limits:

$$\lim_{h\to 0^+}\frac{f(2+h)-f(2)}{h} = \lim_{h\to 0^+}\frac{2(h+2)-4}{h} = \lim_{h\to 0^+}\frac{2h}{h} = 2$$

and

$$\lim_{h\to 0^-}\frac{f(2+h)-f(2)}{h} = \lim_{h\to 0^-}\frac{(h+2)^2-4}{h} = \lim_{h\to 0^-}\frac{4h+h^2}{h} = \lim_{h\to 0^-}(4+h) = 4.$$

Therefore, $\lim_{h\to 0}\frac{f(2+h)-f(2)}{h}$ does not exist and so f is not differentiable at $x = 2$.



Now $f'(x)=\begin{cases} 2, & x>2 \\ 2x, & x<2 \end{cases}$. Differentiating this we get $f''(x)=\begin{cases} 0, & x>2 \\ 2, & x<2 \end{cases}$. Obviously $f''(2)$ does not exist.

4. $f(x)=\begin{cases} 4x^2+1, & x\geq 1 \\ 3x^2+2x, & x<1 \end{cases}$. Then f is differentiable on $(-\infty, 1)\cup(1,\infty)$. Now

$$\lim_{h\to 0^+}\frac{f(1+h)-f(1)}{h} = \lim_{h\to 0^+}\frac{4(1+h)^2+1-5}{h} = \lim_{h\to 0^+}\frac{8h+h^2}{h} = 8$$

and

$$\lim_{h\to 0^-}\frac{f(1+h)-f(1)}{h} = \lim_{h\to 0^-}\frac{3(1+h)^2+2(1+h)-5}{h} = \lim_{h\to 0^-}\frac{6h+h^2+2h}{h}$$
$$=\lim_{h\to 0^-}(8+h) = 8 .$$

Hence, $f'(1) = 8$. Thus $f'(x)=\begin{cases} 8x, & x\geq 1 \\ 6x+2, & x<1 \end{cases}$. Similarly we deduce as follows that $f''(x)=\begin{cases} 8, & x>1 \\ 6, & x<1 \end{cases}$.

Now

$$\lim_{h\to 0^+}\frac{f'(1+h)-f'(1)}{h} = \lim_{h\to 0^+}\frac{8(1+h)-8}{h} = \lim_{h\to 0^+}\frac{8h}{h} = 8$$

and

$$\lim_{h\to 0^-}\frac{f'(1+h)-f'(1)}{h} = \lim_{h\to 0^-}\frac{6(h+1)+2-8}{h} = \lim_{h\to 0^-}\frac{6h}{h} = 6 .$$

Thus the left and right limits are not the same. Therefore, $f''(1)$ does not exist. Hence $f''(x)=\begin{cases} 8, & x>1 \\ 6, & x<1 \end{cases}$.

Sums, Products and Quotients

Theorem 5. Let f and g be defined on a neighbourhood D of x_0. L et λ and μ be any real numbers. Then if f and g are differentiable at x_0,

1. $(\lambda f + \mu g)'(x_0) = \lambda f'(x_0) + \mu g'(x_9)$,
2. (Product rule) $(f g)'(x_0) = f'(x_0)\, g(x_0) + f(x_0)g'(x_0)$,
3. (Quotient rule) if $g(x_0)\neq 0$, $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)\cdot g(x_0) - f(x_0)\cdot g'(x_0)}{(g(x_0))^2}$.

Proof.

1. $\dfrac{\lambda f(x_0+h)+\mu g(x_0+h)-\lambda f(x_0)+\mu g(x_0)}{h}$

$= \lambda \dfrac{f(x_0+h)-f(x_0)}{h} + \mu \dfrac{g(x_0+h)-g(x_0)}{h} \to \lambda f'(x_0)+\mu g'(x_0)$ as $h \to 0$.

This proves part (1).

2. $\dfrac{f(x_0+h)g(x_0+h)-f(x_0)g(x_0)}{h}$

$= \dfrac{f(x_0+h)g(x_0+h)-f(x_0)g(x_0+h)+f(x_0)g(x_0+h)-f(x_0)g(x_0)}{h}$

$= g(x_0+h)\dfrac{f(x_0+h)-f(x_0)}{h} + f(x_0)\dfrac{g(x_0+h)-g(x_0)}{h}$

$\to g(x_0)f'(x_0)+f(x_0)g'(x_0)$ as $h \to 0$ since f and g are differentiable at x_0 and g is continuous at x_0 by Theorem 3. This proves part (2).

3. $\dfrac{1}{h}\left\{\dfrac{f(x_0+h)}{g(x_0+h)} - \dfrac{f(x_0)}{g(x_0)}\right\} = \dfrac{1}{h}\left\{\dfrac{g(x_0)f(x_0+h)-f(x_0)g(x_0+h)}{g(x_0+h)g(x_0)}\right\}$

$= \dfrac{1}{h}\left\{\dfrac{g(x_0)f(x_0+h)-g(x_0)f(x_0)+g(x_0)f(x_0)-f(x_0)g(x_0+h)}{g(x_0+h)g(x_0)}\right\}$

$= \dfrac{g(x_0)\frac{f(x_0+h)-f(x_0)}{h} - f(x_0)\frac{g(x_0+h)-g(x_0)}{h}}{g(x_0+h)g(x_0)} \to \dfrac{g(x_0)f'(x_0)-f(x_0)g'(x_0)}{(g(x_0))^2}$ as $h \to 0$

since f and g are differentiable at x_0 and g is continuous at x_0 by Theorem 3 and non-zero at x_0.

Example 6.

1. If n is a natural number, let $f(x) = x^n$. Then f is differentiable on **R** and $f'(x) = n\,x^{n-1}$. Let a be a point in **R**. For all $x \neq a$,

$$\frac{f(x)-f(a)}{x-a} = \frac{x^n - x^a}{x-a} = \frac{(x-a)(x^{n-1}+x^{n-2}a+x^{n-3}a^2+\cdots+xa^{n-2}+a^{n-1})}{(x-a)}$$

$$= x^{n-1}+x^{n-2}a+x^{n-3}a^2+\cdots+xa^{n-2}+a^{n-1}. \quad \text{-----------------------} \quad (1)$$

Therefore,

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a} = \lim_{x\to a} x^{n-1}+x^{n-2}a+x^{n-3}a^2+\cdots+xa^{n-2}+a^{n-1} = na^{n-1}$$

since there are n terms on the right side of (1) and each term has the same limit a^{n-1}.

2. If $f(x) = \dfrac{1}{x}$, then $f'(x) = \dfrac{0\cdot x - 1\cdot 1}{x^2} = -\dfrac{1}{x^2}$ for $x \neq 0$.

Theorem 7. Any polynomial function is differentiable on **R**. Any rational function is differentiable on its domain of definition.

Proof. Any polynomial function is differentiable follows from Theorem 5 part (1) and (2). Since a rational function is a quotient of one polynomial function $p(x)$ by another polynomial function $q(x)$, by Theorem 5 part (3), p/q is differentiable on its domain of definition. Thus, a rational function is differentiable on its domain of definition.

The Chain Rule

Theorem 5 is a very useful tool for determining derivative but is of little use when it comes to composition of functions unless we can express composition in terms of sums of products or quotients of known differentiable functions. Indeed, composition of differentiable functions is differentiable and there is a simple formula, known as Chain Rule, giving the derivative of the composition.

Theorem 8 (Chain Rule). Let $f : I \to \mathbf{R}$ be a function defined on a neighbourhood I of x_0. Suppose $f(I) \subseteq J$ and J is an open interval. Suppose $g: J \to \mathbf{R}$ is a function defined on J. Then we have the composite $g\circ f : I \to \mathbf{R}$ defined by $g\circ f(x) = g(f(x))$. If f is differentiable at x_0 and g is differentiable at $f(x_0)$, then the composite $g\circ f$ is differentiable at x_0 and $(g\circ f)'(x_0) = g'(f(x_0))\, f'(x_0)$.

Proof.
We have to examine the quotient

$$\frac{g(f(x_0+h)) - g(f(x_0))}{h}.$$

Let $k = f(x_0 + h) - f(x_0)$. Thus k is a function of h. Since f is differentiable at x_0, f is continuous there by Theorem 3 and so $k \to 0$ as $h \to 0$ and k is continuous at 0. Suppose $k \neq 0$. Then we have

$$\begin{aligned}\frac{g(f(x_0+h)) - g(f(x_0))}{h} &= \frac{g(f(x_0+h)) - g(f(x_0))}{f(x_0+h) - f(x_0)} \cdot \frac{f(x_0+h) - f(x_0)}{h}\\ &= \frac{g(k+f(x_0)) - g(f(x_0))}{k} \cdot \frac{f(x_0+h) - f(x_0)}{h}.\end{aligned}$$

Since it is possible that k could be zero and that the above equality is only true for $k \neq 0$, to obtain a similar expression we consider the following device to get round this difficulty.

Define the function $G(k) = \begin{cases} \frac{g(k+y_0)-g(y_0)}{k}, & \text{if } k \neq 0 \\ g'(y_0), & \text{if } k = 0 \end{cases}$, where $y_0 = f(x_0)$. Then G is continuous at 0 since g is differentiable at y_0 so that

$$\lim_{k\to 0} G(k) = \lim_{k\to 0} \frac{g(k+y_0) - g(y_0)}{k} = g'(y_0) = G(0).$$

Thus if $k \neq 0$,

$$\frac{g(f(x_0+h)) - g(f(x_0))}{h} = G(k) \cdot \frac{f(x_0+h) - f(x_0)}{h}.$$

Also if $k = 0$, i.e., $f(x_0 + h) = f(x_0)$, then the above equation, being equal to zero on both sides, is also true. Thus,

$$\begin{aligned}&\frac{g(f(x_0+h)) - g(f(x_0))}{h}\\ &= G(k(h)) \cdot \frac{f(x_0+h) - f(x_0)}{h} \to G(0) \cdot f'(x_0) = g'(y_0)\, f'(x_0)\end{aligned}$$

as $h \to 0$. Note that $\lim_{h\to 0} G(k(h)) = G(0)$ by Theorem 34 part (B) Chapter 3 or by observing that $G \circ k$ is continuous at 0 since G is continuous at $k(0) = 0$ and k is continuous at 0. Hence, $(g\circ f)'(x_0) = g'(f(x_0))\, f'(x_0)$. This proves theorem 8.

Example 9. Let $f(x) = \sqrt{(x+1)}$ for $x > -1$. Then $f(x) = g \circ h(x)$, where $h(x) = x+1$ and $g(y) = \sqrt{y}$. Now $h'(x) = 1$ and $g'(y) = 1/(2\sqrt{y})$ for $y > 0$. Therefore, $f'(x) = (g \circ h)'(x) = g'(h(x))\, h'(x) = g'(x+1)\cdot 1 = 1/(2\sqrt{(x+1)})$

4.2 Mean Value Theorem.

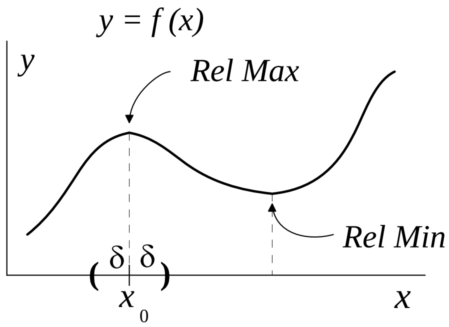
We shall now consider one of the often used theorems in Calculus, the Mean Value Theorem.

First we introduce some local definitions.

Definition 10. Let D be a neighbourhood of x_0 , i.e., there exists an open interval (a, b) such that $x_0 \in (a, b) \subseteq D$. Suppose f is a function defined on D. We say f has a *relative maximum* (*local maximum*) at x_0 if $f(x) \leq f(x_0)$ for all x in some open interval containing x_0 ,i.e., there exists $\delta > 0$ such that for all x in D,

$$|x - x_0| < \delta \Rightarrow f(x) \leq f(x_0).$$

We call such a point x_0 a *local maximizer* for f.



Similarly we say f has a *relative minimum* (*local minimum*) at x_0 if $f(x) \geq f(x_0)$ for all x in some open interval containing x_0 ,i.e., there exists $\delta > 0$ such that for all x in D, $|x - x_0| < \delta \Rightarrow f(x) \geq f(x_0)$. We call x_0 a *local minimizer* for f.

If the function f has either a relative maximum or a relative minimum at x_0 , then we say f has a *relative extremum* (*local extremum*) at x_0 .

Theorem 11. Let $f : I \to \mathbf{R}$ be a function defined on a neighbourhood I of x_0 . Suppose f is differentiable at x_0 and has a relative extremum at x_0 . Then $f'(x_0) = 0$.

Proof. Suppose that x_0 is a local maximizer. Then since f is differentiable at x_0 ,

$$f'(x_0) = \lim_{x\to x_0+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0-} \frac{f(x) - f(x_0)}{x - x_0}.$$

By definition of a local maximizer, there exists an open interval (a, b) such that $x_0 \in (a, b) \subseteq I$ and for all x in (a, b) $f(x) \leq f(x_0)$. Therefore, for all x in (a, b) $x > x_0 \Rightarrow$

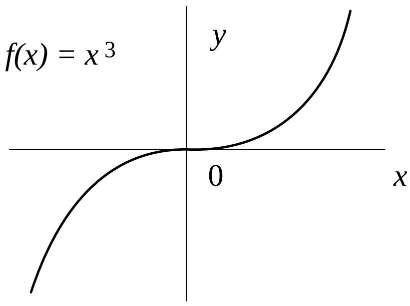
$\frac{f(x)-f(x_0)}{x-x_0} \le 0$. Hence $\lim_{x\to x_0+} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0) \le 0$. We also have that for all x in (a, b) $x < x_0 \Rightarrow \frac{f(x)-f(x_0)}{x-x_0} \ge 0$. Thus, $\lim_{x\to x_0-} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0) \ge 0$. Therefore, $f'(x_0) = 0$.

If x_0 is a local minimizer for f, then x_0 is a local maximizer for $-f$. Therefore, by what we just proved $-f'(x_0) = 0$ and so $f'(x_0) = 0$.

Remark. The converse of Theorem 11 is not true in general. One can find function f and point x_0 such that the derivative $f'(x_0) = 0$ but f does not have a relative extremum at x_0. (See the next example.)

Example 12.

1. Let $f : \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = x^3$. Then $f'(x) = 3x^2$ and so $f'(0) = 0$.

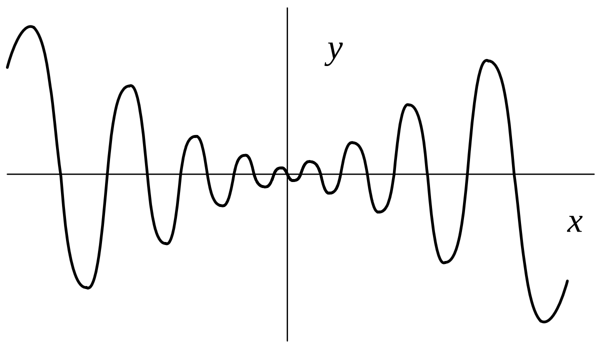


But f does not have a relative extremum at 0.

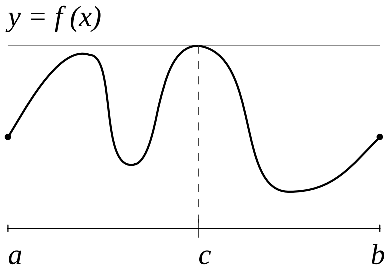
2. Let $f : \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = \begin{cases} x^2 \sin(\frac{1}{x}), & x \ne 0 \\ 0, & x = 0 \end{cases}$. Then f does not have a relative extremum at $x = 0$. f is differentiable at $x = 0$ and

$$f'(0) = \lim_{h\to 0} \frac{f(0+h)-f(0)}{h} = \lim_{h\to 0} \frac{h^2 \sin(\frac{1}{h})-0}{h} = \lim_{h\to 0} h\sin(\frac{1}{h}) = 0$$

by the Squeeze Theorem. (For $h \ne 0, -|h| \le \sin(\frac{1}{h}) \le |h|$.)



Theorem 13 (Rolle's Theorem). Suppose
1. $f:[a, b] \to \mathbf{R}$ is continuous on $[a, b]$,
2. f is differentiable on (a, b) and
3. $f(a) = f(b)$.
Then there exists a point c in (a, b) such that $f'(c) = 0$.



Proof. By Corollary 9 to the Extreme Value Theorem Chapter 3, since f is continuous on the closed interval $[a, b]$, which is compact, f attains its maximum and minimum. Then we have the following possibilities.

1. There exists c in (a,b) such that $f(c)$ is the maximum value
2. There exists c in (a,b) such that $f(c)$ is the minimum value or
3. $f(a)$ and $f(b)$ are both the absolute maximum and minimum value of f since they are equal.

Cases 1 and 2 imply that there exists a point c in (a, b) such that $f(c)$ is a relative extremum. Therefore, since f is differentiable on (a, b), by Theorem 11, $f'(c) = 0$. For case 3, since $f(a) = f(b)$ the maximum value and the minimum value of f are the same, f must be a constant function and so $f'(c) = 0$ for all c in (a, b). For this case take any value c in (a, b). This completes the proof.

Example 14. Let $f:[0, 3] \to \mathbf{R}$ be given by $f(x) = x(x - 3)$. Then $f(0) = f(3) = 0$. Thus, since f is continuous on $[0, 3]$ and differentiable on $(0, 3)$, by *Rolle's Theorem* there exists a point c in $(0, 3)$ such that $f'(c) = 0$.

Rolle's Theorem is a special case of the Mean Value Theorem. Indeed they are equivalent theorems.

Theorem 15 (Mean Value Theorem). Let $a < b$. Suppose $f:[a, b] \to \mathbf{R}$ is continuous on $[a, b]$ and f is differentiable on (a, b). Then there exists c in (a, b) such that $f'(c) = \dfrac{f(b)-f(a)}{b-a}$.

Proof. (Tilt the graph and use *Rolle's Theorem*.)
Define $g:[a, b] \to \mathbf{R}$ by $g(x) = f(x) - \dfrac{f(b)-f(a)}{b-a}(x-a)$. Then

1. g is continuous on $[a, b]$ since f is continuous on $[a, b]$,
2. g is differentiable on (a, b) since f and $(x - a)$ are differentiable on (a, b) and
3. $g(a) = g(b)$ $(= f(a))$.

Therefore, by *Rolle's Theorem*, there exists c in (a, b) such that $g'(c) = 0$. But

$$g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a} \text{ on } (a,b).$$

Therefore, $g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} = 0$. Thus $f'(c) = \frac{f(b)-f(a)}{b-a}$. This completes the proof.

The following is a consequence of the *Mean Value Theorem*.

Theorem 16. Let I be an open interval. Suppose the function $f : I \to \mathbf{R}$ is differentiable. Then f is a constant function if and only if the derivative $f'(x) = 0$ for all x in I.

Proof. Obviously, if f is a constant function, then $f'(x) = 0$ for all x in I.
Suppose now that $f'(x) = 0$ for all x in I. Fix a point x_0 in I. Take any x in I. Suppose $x > x_0$. Let x in (a, b) be such that $x > x_0$. Since f is continuous on I, f is continuous on $[x_0, x]$. Since f is differentiable on I, f is differentiable on (x_0, x). Thus, by the *Mean Value Theorem* (Theorem 15), there exists c in (x_0, x) such that $f'(c) = \frac{f(x)-f(x_0)}{x-x_0}$. But $f'(c) = 0$ and so $f(x) = f(x_0)$. Similarly, if $x < x_0$ we can use the Mean Value Theorem on the restriction of f to $[x, x_0]$ to conclude that $f(x) = f(x_0)$. Hence, $f(x) = f(x_0)$ for all x in I and so f is a constant function.

Corollary 17. If $f : [a, b] \to \mathbf{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and $f'(x) = 0$ for all x in (a, b), then f is a constant function.

Proof. By Theorem 16, f is constant on (a, b), say K, i.e., $f(x) = K$ for all x in (a, b). By the continuity at a, $f(a) = \lim_{x \to a^+} f(x) = \lim_{x \to a^+} K = K$. Also by the continuity at b, $f(b) = \lim_{x \to b^-} f(x) = \lim_{x \to b^-} K = K$. Therefore, f is a constant function on $[a, b]$.

Example 18.

1. Let $f(x) = \sqrt{x}$ on $(0, \infty)$. Then $f'(x) = \frac{1}{2\sqrt{x}}$ on $(0, \infty)$. By the *Mean Value Theorem* there exists a point c in (98,100) such that $\frac{\sqrt{100} - \sqrt{98}}{100-98} = f'(c) = \frac{1}{2\sqrt{c}}$.
But $98 < c < 100$ implies that $\frac{1}{2\sqrt{100}} < \frac{1}{2\sqrt{c}} < \frac{1}{2\sqrt{98}}$.
Hence $\frac{1}{2\sqrt{100}} < \frac{10-\sqrt{98}}{2} < \frac{1}{2\sqrt{98}}$, i.e., $\frac{1}{10} < 10 - \sqrt{98} < \frac{1}{\sqrt{98}} < \frac{1}{9}$. And so $10 - \frac{1}{9} < \sqrt{98} < 10 - \frac{1}{10}$.

2. The sine function is continuous and differentiable on $\mathbf{R}$. Take $x > 0$. Then by the *Mean Value Theorem*, there exists a point c in $(0, x)$ such that $\frac{\sin(x)-\sin(0)}{x-0} = \frac{\sin(x)}{x} = \cos(c)$. Therefore, $\left|\frac{\sin(x)}{x}\right| = |\cos(c)| \le 1$. Similarly, if we take $x < 0$, by the *Mean Value Theorem*, there exists a point d in $(x, 0)$ such that $\frac{\sin(x)-\sin(0)}{x-0} = \frac{\sin(x)}{x} = \cos(d)$. Thus $\left|\frac{\sin(x)}{x}\right| = |\cos(d)| \le 1$. Hence, for $x \ne 0$, $|\sin(x)| \le |x|$ and so $-|x| \le \sin(x) \le |x|$.

4. 3 Monotone Functions, Relative Extrema and Tests for Relative Extrema

We shall now investigate some criteria for function to be monotone.

Suppose A is a subset of $\mathbf{R}$. The interior of A is the set $Int\, A = \{x \in A :$ there exists an open interval (a, b) such that $x \in (a, b) \subseteq A\}$. Thus, $Int\, [a, b]$ is (a, b).

Theorem 19. Let I be an interval. Suppose $f : I \to \mathbf{R}$ is a continuous function and that the restriction of f to the interior of I, Int I, is differentiable.

1. If $f'(x) > 0$ for all x in $Int\, I$, then f is strictly increasing;
   if $f'(x) \geq 0$ for all x in $Int\, I$, then f is increasing.
2. If $f'(x) < 0$ for all x in $Int\, I$, then f is strictly decreasing;
   if $f'(x) \leq 0$ for all x in $Int\, I$, then f is decreasing.

Proof. (1) Suppose $f'(x) > 0$ for all x in $Int\, I$. Let the points c , d in the interval I be such that $c < d$. Then f is continuous on $[c, d]$ and differentiable on $(c, d\,) \subseteq Int\, I$. Therefore, the *Mean Value Theorem* says that there is a point η in (c, d) such that $\dfrac{f(d)-f(c)}{d-c} = f'(\eta) > 0$. Since $d - c > 0$, $f(d\,) - f(c) > 0$. Thus $f(d\,) > f(c)$. Therefore, f is strictly increasing.
(2) Suppose $f'(x) < 0$ for all x in $Int\, I$. Let $g = -f$. Then $g'(x) = -f'(x) > 0$ for all x in I. By part (1) $g = -f$ is strictly increasing. Therefore, f is strictly decreasing.
The cases when $f'(x) \geq 0$ for all x in $Int\, I$ and when $f'(x) \leq 0$ for all x in $Int\, I$ are proved in exactly the same way.

Remark.
1. Notice that in the proof of Theorem 19 part (1), if I is open it is enough to use the inequality $\dfrac{f(d)-f(c)}{d-c} \geq f'(\eta)$. It is true such a point η exists in $[c, d\,]$ with this inequality. This is one reason why we may consider the Mean Value Theorem is over rated. Indeed under the condition of the Mean Value Theorem, there exist points η and γ in $[c, d\,]$ such that

$$f'(\gamma) \geq \frac{f(d)-f(c)}{d-c} \geq f'(\eta).$$

(Ref: Theorem 2 and Theorem 3, in 'Do we need Mean Value Theorem to prove $f'(x) = 0$ on (a, b) implies that f is constant on (a, b)?')
This can only prove that f is strictly increasing on $Int\, I$ if $f'(x) > 0$ for all x in $Int\, I$. Extend to all of I by continuity. Indeed we may use this result instead of the Mean Value Theorem, whenever its application uses inequality.

2. The property of f being strictly increasing or strictly decreasing is a global property. Thus a local information like $f'(x_0) > 0$, does not necessarily imply that f *is* strictly increasing in a neighbourhood containing x_0.
For instance take the function $f(x) = \begin{cases} x + 4x^2\cos(\frac{1}{x}), x \neq 0 \\ 0, x = 0 \end{cases}$. Then $f'(0) = 1 > 0$ but f is neither increasing nor decreasing on any interval containing 0. This is because for any integer $n > 0$, $1/(2n\pi+\pi) > 1/(2n\pi+2\pi)$ but $f(1/(2n\pi+\pi)) = 1/(2n\pi+\pi)$-

$4/(2n\pi+\pi)^2 < 1/(2n\pi+2\pi) + 4/(2n\pi+2\pi)^2 = f(1/(2n\pi+2\pi))$ and that when $1/(2n\pi+\pi/2) > 1/(2n\pi+3\pi/2)$, $f(1/(2n\pi+\pi/2)) = 1/(2n\pi+\pi/2) > 1/(2n\pi+3\pi/2) = f(1/(2n\pi+3\pi/2))$. The same can be said of the property of f being increasing or decreasing.

Example 20.

1. Let $f : \mathbf{R} \to \mathbf{R}$ be given by $f(x) = x(1+\frac{x}{2})$. Then f is differentiable on **R** and is therefore continuous on **R**. Its derivative $f'(x) = 1 + x$. Hence, $f'(x) > 0$ for $x > -1$ and $f'(x) < 0$ for $x < -1$. Since f is continuous on $[-1, \infty)$ and differentiable on $(-1, \infty)$, f is increasing on $[-1, \infty)$ by Theorem 19. Similarly, since f is continuous on $(-\infty, -1]$, f is decreasing on $(-\infty, -1]$, Therefore, $f(-1)$ is the (absolute) minimum of f.

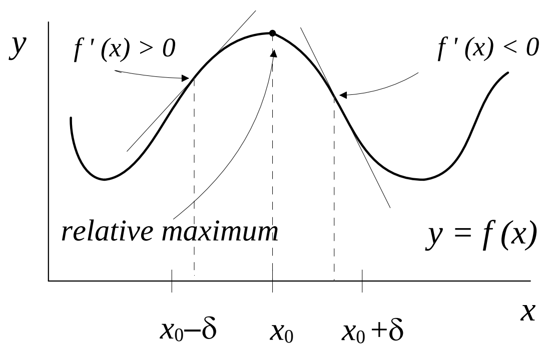
2. Let $f : (0, \infty) \to \mathbf{R}$ be the function defined by $f(x) = \frac{1}{x}$. Then f is differentiable on $(0, \infty)$ and $f'(x) = -\frac{1}{x^2} < 0$ for all $x > 0$. Thus f is decreasing on $(0, \infty)$.
(Incidentally one can deduce this directly by using property of inequality.)

Note that the point where a function changes from being increasing before it to decreasing after it or from being decreasing before it to increasing after it must be a point where the function has a relative extremum. Thus we can formulate the following test for relative extrema.

Theorem 21 (First Derivative Test for Relative Extrema). Suppose f is continuous on the open interval I containing x_0 and that f is differentiable at all points of I except possibly at x_0.

1. If there exists $\delta > 0$ such that for all x in I with $x_0 - \delta < x < x_0$, $f'(x) \geq 0$ and that for all x with $x_0 < x < x_0 + \delta$, $f'(x) \leq 0$, then f has a relative maximum value at x_0, i.e., x_0 is a local maximizer of f.

2. If there exists $\delta > 0$ such that for all x in I with $x_0 - \delta < x < x_0$, $f'(x) \leq 0$ and that for all x with $x_0 < x < x_0 + \delta$, $f'(x) \geq 0$, then f has a relative minimum value at x_0, i.e., x_0 is a local minimizer of f.



Proof. Part (1).
Since I is open there exists a $\delta' > 0$ such that $(x_0 - \delta', x_0 + \delta') \subseteq I$. By taking the intersection of $(x_0 - \delta', x_0 + \delta')$ and $(x_0 - \delta, x_0 + \delta)$ if need be, and renaming if necessary, we may assume that $(x_0 - \delta, x_0 + \delta) \subseteq I$. By Theorem 19, f is increasing on $(x_0 - \delta, x_0]$ because $f'(x) \geq 0$ on $(x_0 - \delta, x_0)$. By Theorem 19, f is decreasing on $[x_0, x_0 + \delta)$ because $f'(x) \leq 0$ on $(x_0, x_0 + \delta)$. For any x in $(x_0 - \delta, x_0 + \delta)$, $x \leq x_0$ or $x \geq x_0$. If $x \leq x_0$, then $f(x) \leq f(x_0)$ because f is increasing on $(x_0 - \delta, x_0]$. If $x \geq x_0$, then $f(x) \leq f(x_0)$ because f is decreasing on $[x_0, x_0 + \delta)$. Therefore, for all x in $(x_0 - \delta, x_0 + \delta)$, $f(x) \leq f(x_0)$. Thus $f(x_0)$ is a relative maximum and x_0 is a local maximizer of f.
The proof of part (2) is similar. We may observe that by Part (1), x_0 is a local maximizer of $-f$. Therefore, x_0 is a local minimizer of f.

Examples 22.

1. Let $f(x) = x^3 - 6x^2 + 9x + 1$. Then

$$f'(x) = 3x^2 - 12x + 9 = 3(x^2 - 4x + 3) = 3(x - 3)(x - 1).$$

Thus $f'(x) = 0$ if and only if $x = 1$ or 3. For $x < 1$, $x - 1 < 0$ and $x - 3 < 0$ so that $f'(x) > 0$. For $1 < x < 3$, $x - 1 > 0$ and $x - 3 < 0$ so that $f'(x) < 0$. For $x > 3$, $x - 1 > 0$ and $x - 3 > 0$ so that $f'(x) > 0$. Thus at $x = 1$, we have a relative maximum and the value is $f(1) = 5$. At $x = 3$, we have a relative minimum which is $f(3) = 1$.

2. Let $f(x) = \begin{cases} 2 - x^3, & x < 1 \\ x^2, & x \geq 1 \end{cases}$. Then $f'(x) = \begin{cases} -3x^2, & x < 1 \\ 2x, & x > 1 \end{cases}$.

f is not differentiable at 1 and $f'(x) = 0$ only if $x = 0$. For $0 < x < 1$, $f'(x) < 0$ and for $x > 1$, $f'(x) > 0$ and so by the first derivative test, $f(1) = 1$ is a relative minimum. Now for $x < 0$, $f'(x) < 0$. Although $f'(0) = 0$, $f(0)$ is not a relative extremum. Indeed f is decreasing on $(-\infty, 1]$ because it is decreasing on $(-\infty, 0]$ and on $[0, 1]$.

Remark. Note that we do not require that the function f be differentiable at x_0 in the First Derivative Test (Theorem 21).

We next describe a weaker test for finding local maximizers and local minimizers.
Suppose $f : D \to \mathbf{R}$ is a function. A *stationary point* of f is a point x in D where D is a neighbourhood of x and f is differentiable there with $f'(x) = 0$. A *critical point* of f is a point x, where either f is not differentiable or x is a stationary point.

Theorem 23 (Second Derivative Test for Relative Extremum). Let $f : D \to \mathbf{R}$ be a function, where D is a neighbourhood of x_0. Suppose f is differentiable on an open interval I with $I \subseteq D$. Suppose x_0 is a stationary point of f, i.e., $f'(x_0) = 0$. Suppose $f''(x_0)$ exists.
1. If $f''(x_0) < 0$, then f has a relative maximum value at x_0.
2. If $f''(x_0) > 0$, then f has a relative minimum value at x_0.

Proof .
Part (1)

$$0 > f''(x_0) = \lim_{x\to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x\to x_0} \frac{f'(x)}{x - x_0}$$

because $f'(x_0) = 0$. Therefore, by the definition of limit there exists $\delta > 0$ such that $(x_0 - \delta,\ x_0 + \delta) \subseteq I$ and for all $x \neq x_0$ in $(x_0 - \delta,\ x_0 + \delta)$

$$\frac{f'(x)}{x - x_0} < 0.$$

Therefore, for x in $(x_0 - \delta, x_0)$, $x - x_0 < 0$, $f'(x) > 0$. Also for x in $(x_0, x_0 + \delta)$, $f'(x) < 0$. Thus, by the *First Derivative Test* (Theorem 21), we have a relative maximum value at x_0.
Part (2). A similar argument as above applies to give part (2). We may also note that if $f''(x_0) > 0$, then $(-f)''(x_0) = -f''(x_0) < 0$. Therefore, by Part (1), x_0 is a local maximizer for $-f$ and hence is a local minimizer for f.

Remark.
1. Note that Theorem 23 has additional condition of twice differentiability of f at x_0 imposed, whereas the first derivative test do not need differentiability of f at x_0 but just continuity at x_0. Thus Theorem 23 is a weaker theorem.

2. If $f'(x_0) = f''(x_0) = 0$, then Theorem 23 gives no information whether $f(x_0)$ is a relative extremum. For instance, if $f(x) = x^4$, then $f''(0) = 0$ and $f(0)$ is a relative minimum. If $f(x) = x^3$, then $f''(0) = 0$ but $f(0)$ is not a relative extremum.

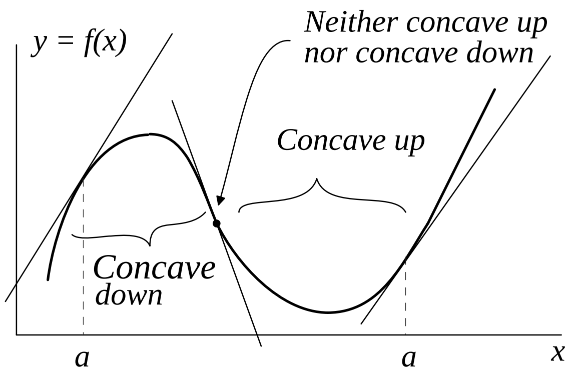
4.4 Concavity

We shall now consider the notion of concavity. There are different definitions, especially a local definition and a global definition. We shall consider the local definition.

Definition 24. Suppose $f : I \to \mathbf{R}$ is a continuous function defined on an interval I. Suppose a is a point in the interior of I. Then the graph of f is *concave upward* (respectively *concave downward*) at $x = a$ if there exists a small neighbourhood N of a such that in this small neighbourhood the graph of f lies above (respectively below) the tangent line to the graph of f at $(a,\ f(a))$ except for the point of tangency. That is to say, the graph of f is *concave upward* (respectively *concave downward*) at $x = a$ if there exists a $\delta > 0$ such that for all x not equal to a in $(a-\delta, a+\delta)$,

$f(x) > f(a) + f'(a)(x - a)$ (respectively $f(x) < f(a) + f'(a)(x - a)$).

We say the graph of f is concave upward (respectively concave downward) on an open interval I if the graph of f is concave upward (respectively concave downward) at x for all x in I.



Remark. There are other refined definitions of concavity but this one is more intuitive. Note that in the above definition no use is made of the second derivative. The next theorem gives a criterion for determining whether the graph is concave upward or downward when the second derivative does exist.

Theorem 25. Let f be a function differentiable on an open interval D containing x_0. Then

1. if $f''(x_0) > 0$, the graph of f is concave upward at $(x_0, f(x_0))$,
2. if $f''(x_0) < 0$, the graph of f is concave downward at $(x_0, f(x_0))$.

Proof . We shall prove only part (1). Part (2) is similarly proved.

Since $f''(x_0) > 0$, $\lim_{x\to x_0}\dfrac{f'(x)-f'(x_0)}{x-x_0} = f''(x_0) > 0$. Therefore, by the definition of limit, there is an open interval $I = (x_0 - \delta, x_0 + \delta) \subseteq D$ such that for all $x \neq x_0$ in I,

$$g(x) = \frac{f'(x)-f'(x_0)}{x-x_0} > 0. \quad \text{--------------------------} \quad (1)$$

The tangent line to the graph of f at $(x_0, f(x_0))$ is given by $\dfrac{y-f(x_0)}{x-x_0} = f'(x_0)$, i.e.,

$$y = f(x_0) + f'(x_0)(x - x_0) .$$

We want to show that, $x \neq x_0$ in I, $f(x) > y = f(x_0) + f'(x_0)(x - x_0)$, i.e.,

$$f(x) - f(x_0) > f'(x_0)(x - x_0).$$

Now for x in $(x_0 , x_0+ \delta)$, $x > x_0$ and so by (1),

$$f'(x) > f'(x_0). \quad \text{-------------------------------} \quad (2)$$

By the *Mean Value Theorem*, there exists a x' in (x_0, x) such that

$$\frac{f(x)-f(x_0)}{x-x_0} = f'(x') > f'(x_0)$$

by (2).

Therefore, for x in $(x_0, x_0+ \delta)$, $f(x) - f(x_0) > f'(x_0)(x - x_0)$, which is what is required to prove. Similarly, for x in $(x_0 - \delta, x_0)$,

$$f'(x) < f'(x_0) \quad \text{---------------------------------} \quad (3)$$

by (1). Now the *Mean Value Theorem* applies to give x' in (x, x_0) such that

$$\frac{f(x)-f(x_0)}{x-x_0} = f'(x') < f'(x_0)$$

by (3).

Since $x < x_0$, we have (by multiplying the above by $(x - x_0) < 0$) for x in $(x_0- \delta, x_0)$,

$$f(x) - f(x_0) > f'(x_0)(x - x_0).$$

Thus we have shown that for any $x \neq x_0$ in I, $f(x) - f(x_0) > f'(x_0)(x - x_0)$, which is what is required.

Definition 26. A point $(c,\ f(c))$ is a *point of inflection* of the graph of the function f if f is continuous at c and there is an open interval containing c such that the graph of f changes from concave upward before c to concave downward after c or from concave downward before c to concave upward after c .

Note. Our definition does not require that the function be differentiable at a point of inflection. There are other more refined definitions of a point of inflection but ours is the simplest.

Example 27.

1. Let $f(x) = x(x^2 - 1) = x^3 - x$. Then $f'(x) = 3x^2 - 1$ and $f''(x) = 6x$. Thus $f''(x) > 0$ for $x > 0$. Therefore, the graph of f is concave upward on the interval $(0, \infty)$. Also $f''(x) < 0$ for $x < 0$. Thus the graph of f is concave downward on the interval $(-\infty, 0)$. The point $(0, 0)$ is a point of inflection.
2. Let $f(x) = \begin{cases} 3x^2 + 1, x \geq 0 \\ 2x^3 + 1, x < 0 \end{cases}$. Then $f'(x) = \begin{cases} 6x\ , x \geq 0 \\ 6x^2\ , x < 0 \end{cases}$ and $f''(x) = \begin{cases} 6\ \ , x > 0 \\ 12x\ , x < 0 \end{cases}$.

$f''(0)$ does not exist. Obviously $f''(x) > 0$ for $x > 0$ and $f''(x) < 0$ for $x < 0$. Thus there is a point of inflection at $x = 0$. The graph of f is concave upward on the interval $(0, \infty)$ and is concave downward on the interval $(-\infty, 0)$.

Theorem 28. Suppose the graph of a function f is either concave upward or concave downward on an open interval I. Then any tangent line to the graph of f can only intersect the graph of f at the point of tangency.

Proof. We shall prove the theorem for the case f is concave upward on the open interval I. Note that since the graph of f is concave upward on I, the function f is differentiable on I.

Suppose there exists a point k in I such that the tangent line at $(k,\ f(k))$ meets the graph again at the point $(p, f(p))$. We may assume that $k < p$. Then the equation of the tangent line at $x = k$ is given by

$$y = f(k) + f'(k)(x - k).$$

We now proceed by "tilting the graph". Let $g : [k, p] \to \mathbf{R}$ be defined by

$$g(x) = f(x) - f(k) - f'(k)(x - k).$$

Then since f is differentiable on $[k, p]$, g is also differentiable on (k, p) and continuous on $[k, p]$. By the Extreme Value Theorem, there exists a maximum of g on $[k, p]$. Note that $g(k) = 0$ and $g(p) = 0$ since $f(p)$ lies on the tangent line to the graph of f at $x = k$ so that $f(p) = f(k) + f'(k)(p - k)$. Since the graph of f is concave upward at $x = k$, there exists $\delta > 0$ such that for all x in $(k, k + \delta)$, $f(x) > f(k) + f'(k)(x - k)$ and so $g(x) > 0$. Hence the maximum of g can only occur in the interior of $[k, p]$. Suppose that it occurs at $x = d$ in the interior of $[k, p]$. Then for all x in $[k, p]$,

$$g(x) = f(x) - f(k) - f'(k)(x - k) \leq g(d) = f(d) - f(k) - f'(k)(d - k).$$

In particular $f'(d) = f'(k)$. This is because since $g(d)$ is a relative maximum, $g'(d) = 0$ and so since $g'(x) = f'(x) - f'(k)$ for x in (k, p), $g'(d) = 0$ implies that $f'(d) = f'(k)$.
Therefore, we have that for all x in $[k, p]$,

$$f(x) \le f(d) + f'(d)(x - d),$$

which is derived from $g(x) \le g(d)$. But since the equation of the tangent line to f at the point $x = d$ is given by $y = f(d) + f'(d)(x - d)$, we therefore conclude that there exists a $\delta > 0$ such that $(d - \delta, d + \delta) \subseteq [k, p]$ and

$$f(x) > f(d) + f'(d)(x - d)$$

for $x \neq d$ in $(d - \delta, d + \delta)$. But we have just shown that for $x \neq d$ in $(d - \delta, d + \delta)$, $f(x) \le f(d) + f'(d)(x - d)$. This contradiction shows that the tangent line at any point k cannot meet the graph of f at $(p, f(p))$. If $p < k$, we can show similarly, that the tangent line cannot meet the graph of f at $(p, f(p))$ too. Hence any tangent line to the graph of f at any point $(x, f(x))$ cannot intersect the graph of f other than the point of tangency $(x, f(x))$.
(This argument also applies to the case when the graph of f is concave downward on the open interval I.) This completes the proof.

Theorem 29. (1) If the function f is concave upward on an open interval I, then the derived function f' is strictly increasing on I.
(2) If the function f is concave downward on an open interval I, then the derived function f' is strictly decreasing on I.

Proof. Part (1) Take any two points $c < d$ in I. By Theorem 28, since the tangent line at any point on the graph can only meet the graph exactly once, and since the graph of f is concave upward on I, for any point k in I,

$$f(x) > f(k) + f'(k)(x - k) \text{ for } x \neq k.$$

Thus we have

$$f(x) > f(c) + f'(c)(x - c) \text{ for } x \neq c \text{ in } I \qquad \text{---------------} \quad (1)$$

And

$$f(x) > f(d) + f'(d)(x - d) \text{ for } x \neq d \text{ in } I \qquad \text{--------------------} \quad (2)$$

Hence, from (1), putting $x = d$, $f(d) > f(c) + f'(c)(d - c)$ so that $\frac{f(d) - f(c)}{d - c} > f'(c)$.
We also have by setting $x = c$ in (2), $f(c) > f(d) + f'(d)(c - d)$ so that $\frac{f(d) - f(c)}{d - c} < f'(d)$. Therefore, $f'(c) < \frac{f(d) - f(c)}{d - c} < f'(d)$. This shows that f' is strictly increasing.
Similar argument applies to part (2).

Theorem 30. Suppose f is differentiable on some open interval containing c and $(c, f(c))$ is a point of inflection of the graph of f. If $f''(c)$ exists, then $f''(c) = 0$.

Proof. Since $(c, f(c))$ is a point of inflection of the graph of f, there exists a $\delta > 0$ such that f is concave downward (or concave upward) on $(c - \delta, c)$ and concave upward (or concave downward) on $(c, c + \delta)$. We shall assume without loss of generality that f is concave downward before c and concave upward after c. Thus f' is strictly decreasing on $(c - \delta, c)$ by Theorem 29. Also by Theorem 29, since the graph of f is concave upward on $(c, c + \delta)$, f' is strictly increasing on $(c, c + \delta)$. Thus, since f' is continuous at c, $f'(c)$ is a relative minimum. This can be deduced as follows. For any $x < c$ in $(c - \delta, c)$ $f'(j) > f'(y)$ for all y with $x < y < c$. Thus by the continuity of f at c, $f'(c) = \lim_{y \to c^-} f'(y) \le f'(x)$. Likewise for any $x > c$ in $(c, c + \delta)$, $f'(x) > f'(y)$ for all y with $x > y > c$. Again by the continuity of f at $x = c$, $f'(x) \ge \lim_{y \to c^+} f'(y) = f'(c)$. Hence $f'(x) \ge f'(c)$ for all x in $(c - \delta, c + \delta)$ and so $f'(c)$ is a relative minimum. Therefore, by Theorem 11, since f' is differentiable at c, $f''(c) = 0$.

Examples 31.

1. Let $f(x) = x^3 - x$, then $f'(x) = 3x^2 - 1$ and $f''(x) = 6x$. Thus $f''(0) = 0$ and $(0, f(0)) = (0, 0)$ is a point of inflection.
2. The converse of Theorem 30 is false. Take $f(x) = x^4$, $f'(x) = 4x^3$, $f''(x) = 12x^2$. Then $f''(0) = 0$ but $(0, 0)$ is not a point of inflection. The graph of f is in fact concave upward at $(0, 0)$, since it is above the tangent line there. So we can only use Theorem 30 to confirm a point of inflection.

The converse of Theorem 29 is also true.

Theorem 32. Let I be an open interval. Suppose $f : I \to \mathbf{R}$ is differentiable.
(1) If the derived function f' is strictly increasing on I, then the graph of the function f is concave upward on I.
(2) If the derived function f' is strictly decreasing on I, then the graph of the function f is concave downward on I.

Proof.
Part(1). The proof makes use of a similar construct as in the proof of Theorem 28. Take any point c in I. Define $g(x) = f(x) - f(c) - f'(c)(x - c)$ for any x in I. Then g is differentiable on I and $g'(x) = f'(x) - f'(c)$ for any x in I. Now for $x > c$ in I, $f'(x) > f'(c)$. Therefore, $x > c$ in I implies that $g'(x) = f'(x) - f'(c) > 0$. Therefore, g is strictly increasing on the interval $[c, \infty) \cap I$. Now note that $g(c) = 0$. Hence we can conclude that $x > c$ in I implies that $g(x) > g(c) = 0$. This means

$$f(x) - f(c) - f'(c)(x - c) > 0 \text{ for any } x > c \text{ in } I.$$

Hence, for any $x > c$ in I, we have

$$f(x) > f(c) + f'(c)(x - c).$$

Similarly for any $x < c$ in I, $f'(x) < f'(c)$. Therefore, for any $x < c$ in I,

$$g'(x) = f'(x) - f'(c) < 0.$$

We can now conclude that g is strictly decreasing on $(-\infty, c] \cap I$. Therefore, for any $x < c$ in I, $g(x) > g(c) = 0$. Hence for any $x < c$ in I, $f(x) - f(c) - f'(c)(x - c) > 0$. We then have for any $x < c$ in I,

$$f(x) > f(c) + f'(c)(x - c).$$

In this way we have shown that for all $x \ne c$ in I, $f(x) > f(c) + f'(c)(x - c)$. Therefore, by Definition 24, the graph of f is concave upward at $x = c$. Since this is so for any c in I, the graph of f is concave upward on I. We have proved much more. For any $x > c$ in I, $\frac{f(x) - f(c)}{x - c} > f'(c)$ and for any $x < c$ in I, $\frac{f(x) - f(c)}{x - c} < f'(c)$. The case of part (2) when the derived function f' is strictly decreasing is proven similarly.

Remark.
Theorem 29 and 32 says that concavity of f on an *open interval* is equivalent to the strict monotonicity of derived function f'. This is where the monotonicity of the derived function may be defined as the concavity of the graph of the function. It is possible for a function whose graph is concave upward at a point to have its derived function not increasing in any neighbourhood of the point (see the next example).

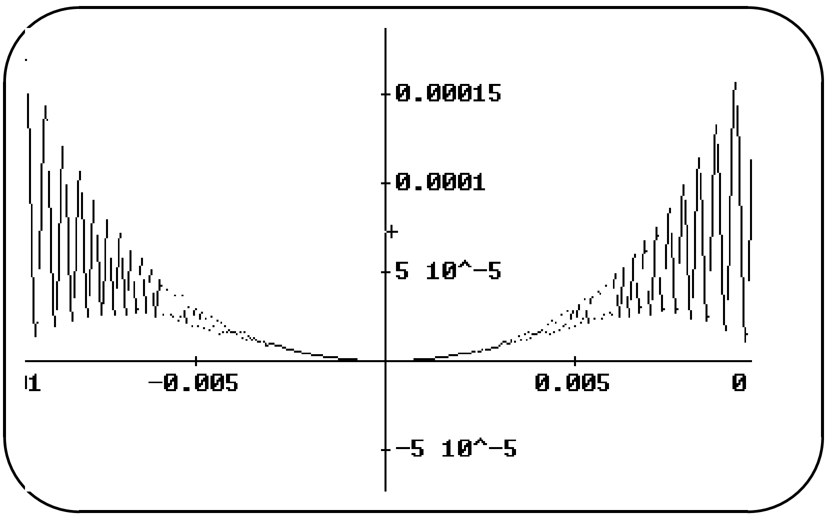
Example 33. Let $f(x)=\begin{cases} x^2+10000x^4\sin(1/x), & x\neq 0 \\ 0, & x=0 \end{cases}$. Then f is differentiable on **R**, the derivative at $x = 0$, $f'(0)$ is equal to 0 by the Squeeze Theorem. The derived function is given by

$$f'(x)=\begin{cases} 2x+40000x^3\sin(1/x)-10000x^2\cos(1/x), & x\neq 0 \\ 0, & x=0 \end{cases}$$

and the second derived function is

$$f''(x)=\begin{cases} 2+120000x^2\sin(1/x)-60000x\cos(1/x)-10000\sin(1/x), & x\neq 0 \\ 2, & x=0 \end{cases}.$$

The large constant is given here for a help to plot the graph of this function to observe the perpectual small oscillation. Note that when $x = 1/((2k+1)\pi/2)$, $\sin(1/x) = 1$ when k is even and -1 when k is odd. Thus for any $\delta > 0$ choose integer k such that $1/((2k+1)\pi/2) < \min(\delta, 1/100)$. Let x_δ to be $1/((2k+1)\pi/2)$. Then obviously, for even k, $f''(x_\delta)=2+120000x_\delta^2-10000 < 14-10000<0$. Since f'' is continuous at x_δ, there exists a small open neighbourhood N_δ of x_δ in $(0, \delta)$ such that $f''(x) < 0$ for all x in this neighbourhood. Therefore f' is strictly decreasing in N_δ. This means for any $\delta > 0$, we can find a neighbourhood (an interval) N_δ such that f' is strictly decreasing in N_δ. Note that $f''(0) = 2 > 0$. Therefore, the graph of f is concave upward at the point $x = 0$. But by the above remark f' cannot be increasing in any neighbourhood containing $x = 0$. The derived function f' fails to be strictly increasing in any open interval containing $x = 0$ simply because we can always find a subinterval on which f' is decreasing. Because we can always find arbitrarily small x_δ such that $f''(x_\delta) < 0$, we can thus find arbitrary small x_δ such that the graph of f is concave downward at $x = x_\delta$. For this function, there is no open interval containing 0 on which the function f is concave upward. Below is a sketch of the function.



4.5 Derivative of inverse function.

Suppose I is an interval and $f : I \to \mathbf{R}$ is a function. If x is in the interior of I, then we have defined what it means for f to be differentiable at x. If I is not open, then we can define the derivative of f at the end point to be the appropriate left or right limit. For instance if $I = [a, b]$, then if the right limit $\lim_{x \to a^+} \frac{f(x) - f(a)}{x - a}$ exists, f is said to be differentiable at a and we write $f'(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a}$. If the left limit $\lim_{x \to b^-} \frac{f(x) - f(b)}{x - b}$ exists, then we say f is differentiable at b and we write $f'(b) = \lim_{x \to b^-} \frac{f(x) - f(b)}{x - b}$. Thus f is said to be differentiable at the end point of I if the appropriate left or right limit exists. In this way we extend the definition of derivative to the end points of an interval. Thus f is said to be differentiable if f is differentiable at every x in I.

Theorem 34. Suppose I is an interval and $f : I \to \mathbf{R}$ is a strictly monotone continuous function. Suppose f is differentiable. Let x_0 be in I and $y_0 = f(x_0)$. Suppose $f'(x_0) \neq 0$. Then the inverse function of f, f^{-1} is differentiable at $y_0 = f(x_0)$ and $(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))} = \frac{1}{f'(x_0)}$.

Proof.

By Theorem 23 Chapter 3, since f is continuous and strictly monotone on I, f^{-1} is continuous and also strictly monotone on the range of f, $J = f(I)$. If $f'(x_0) \neq 0$, we shall show that f^{-1} is differentiable at $y_0 = f(x_0)$. Note that by Theorem 15 Chapter 3, f maps the interior of I onto the interior of J. Suppose x_0 is in *Int I*, then $y_0 = f(x_0)$. is in the *Int J*. Then for this fix y_0, there exists an open interval $(c, d) \subseteq J$ such that $y_0 \in (c, d)$. For each y in (c, d), let

$$g(y) = f^{-1}(y) - f^{-1}(y_0).$$

Since f^{-1} is continuous, g is a continuous function with domain (c, d). We note here that $\lim_{y \to y_0} g(y) = \lim_{h \to 0} f^{-1}(y) - f^{-1}(y_0) = f^{-1}(y_0) - f^{-1}(y_0) = 0$ since f^{-1} is continuous at y_0. Now if we write k for $g(y)$, then $f^{-1}(y) = f^{-1}(y_0) + k = x_0 + k$. So applying f, we get $y = f(x_0 + k)$. Since f^{-1} is strictly monotone and therefore injective, g is injective and so $k = g(y) \neq 0$ unless $y = y_0$. Thus we can write for $y \neq y_0$,

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{k}{f(x_0 + k) - f(x_0)} .$$

In this way we can consider $\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0}$ as a function of k which in turn is a function of y. That means $\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = F \circ g(y)$, where

$$F(t) = \frac{t}{f(x_0 + t) - f(x_0)} .$$

Since it is given that $f'(x_0) \neq 0$ $\lim_{t \to 0} F(t) = \lim_{t \to 0} \frac{t}{f(x_0 + t) - f(x_0)} = \frac{1}{f'(x_0)}$. Note that g is continuous at y_0 and $\lim_{y \to y_0} g(y) = 0$. Therefore, by Theorem 34 part (B) Chapter 3,

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{y \to y_0} F \circ g(y) = \lim_{t \to 0} F(t) = \frac{1}{f'(x_0)}.$$

This proves that f^{-1} is differentiable at $y_0 = f(x_0)$ and

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))} \text{ when } x_0 \text{ is in } Int\ I.$$

If x_0 is an end point of I, then $y_0 = f(x_0)$ is also an end point of J. If y_0 is a left end point of J, then there exists a half open interval $[y_0, d)$ such that $[y_0, d) \subseteq J$. Then the argument above applies with (c, d) replaced by $[y_0, d)$. Similarly if y_0 is a right end point of J, then there exists a half open interval $(c, y_0]$ such that $(c, y_0] \subseteq J$. Then the above argument applies with (c, d) replaced by $(c, y_0]$ to give the same conclusion but with the appropriate one sided limit. This completes the proof of Theorem 34.

Example 35

Let $f : (0, \infty) \to (0, \infty)$ be defined by $f(y) = y^{\frac{1}{n}}$. Then f is the inverse function to $g : (0, \infty) \to (0, \infty)$ defined by $g(x) = x^n$. g is differentiable and $g'(x) = nx^{n-1} > 0$ for all $x > 0$. Therefore g is strictly monotone on $(0, \infty)$. Thus $f = g^{-1}$ is differentiable and

$$f'(y) = (g^{-1})'(y) = \frac{1}{g'(g^{-1}(y))} = \frac{1}{g'(f(y))} = \frac{1}{g'(y^{\frac{1}{n}})} = \frac{1}{n(y^{\frac{1}{n}})^{n-1}} = \frac{1}{n} y^{\frac{1}{n}-1}.$$

Then using this and the chain rule we can easily obtained a similar formula for the differentiation of rational power. The reader is urged to carry out this simple exercise.

4.6 Cauchy Mean Value Theorem, L' Hôpital's Rule

We have seen application of the Mean Value Theorem. It is a very useful tool in analysis. There is a generalization of the Mean Value Theorem that may be applied to a wider context. This is called the Cauchy Mean Value Theorem. Mean Value Theorem and Rolle's Theorem may be viewed as a special case of the Cauchy Mean Value Theorem.

Theorem 36 (Cauchy Mean Value Theorem). Suppose f and g are functions continuous on $[a, b]$, differentiable on (a, b) and suppose that $g'(x) \neq 0$ for all x in (a, b). Then there exists a point c in (a, b) such that $\dfrac{f'(c)}{g'(c)} = \dfrac{f(b) - f(a)}{g(b) - g(a)}$.

Proof. First note that $g'(x) \neq 0$ for all x in (a, b) implies that $g(b) \neq g(a)$ (otherwise if $g(b) = g(a)$, by *Rolle's Theorem* (Theorem 13), there would be a point k in (a, b) with $g'(k) = 0$, contradicting $g'(x) \neq 0$ for all x in (a, b)). Define $F : [a, b] \to \mathbf{R}$ by

$$F(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Then F is continuous on $[a, b]$, and is differentiable on (a, b) since f and g are continuous on $[a, b]$ and differentiable on (a, b). Observe that $F(a) = f(a)$ and

$$F(b) = f(b) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(b) - g(a)) = f(b) - (f(b) - f(a)) = f(a).$$

Therefore, $F(a) = F(b)$ and so by *Rolle's Theorem*, there exists c in (a, b) such that $F'(c) = 0$. Thus, $F'(c) = f'(c) - \dfrac{f(b) - f(a)}{g(b) - g(a)} \cdot g'(c) = 0$. Therefore, since $g'(x) \neq 0$,

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

Remark.
1. If g is the identity function, then Theorem 36 is just the Mean Value Theorem and if further $f(a) = f(b)$, then we get the Rolle's Theorem.

2. As in the case of Mean Value Theorem, most application of Theorem 36 is to differentiable functions defined on closed and bounded interval and involves inequality. Then the following result may be used in its place. Suppose f and g are functions continuous and differentiable on $[a, b]$ and suppose that $g'(x) \neq 0$ for all x in $[a, b]$. Then there are points p and q in $[a, b]$ such that

$$\frac{f'(p)}{g'(p)} \geq \frac{f(b)-f(a)}{g(b)-g(a)} \geq \frac{f'(q)}{g'(q)}.$$

(Reference: "L' Hôpital's Rule and A generalized Version" on my Calculus Web)
For instance, we may use this result to prove Theorem 37.

There are many applications of the Cauchy Mean Value Theorem. Among the more spectacular ones are to the various form of *L' Hôpital's Rule*. We can use it also to prove the Taylor's expansion of a function with the Lagrange form of the remainder.

Theorem 37 (L' Hôpital's Rule).
(A) Suppose f and g are functions differentiable on (a, b) and $g'(x) \neq 0$ for all x in (a, b).
(1) Suppose $\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$. Then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ if the second limit (on the right) exists.
(2) Suppose $\lim_{x\to b^-} f(x) = \lim_{x\to b^-} g(x) = 0$. Then $\lim_{x\to b^-} \frac{f(x)}{g(x)} = \lim_{x\to b^-} \frac{f'(x)}{g'(x)}$ if the second limit (on the right) exists.
(B) Suppose f and g are functions differentiable for all $x > K$ for some real number $K > 0$ and that $g'(x) \neq 0$ for all $x > K$. Suppose $\lim_{x\to+\infty} f(x) = \lim_{x\to+\infty} g(x) = 0$. Then

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = \lim_{x\to+\infty} \frac{f'(x)}{g'(x)}$$

if the second limit (on the right) exists.
(C) Suppose f and g are functions differentiable for all $x < L$ for some real number $L < 0$ and that $g'(x) \neq 0$ for all $x < L$. Suppose $\lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} g(x) = 0$. Then

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

if the second limit (on the right) exists.

Proof . We shall prove only part (A) part (1) and (B). (A) part (2) and (C) is proved similarly.
(A) Suppose $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = l$. Then by the definition of right limit, given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$a < x < a+\delta \Rightarrow \left|\frac{f'(x)}{g'(x)} - l\right| < \varepsilon \quad \text{--------------------------} \quad (1)$$

Extend the function f and g to $[a, b)$ by defining $f(a) = 0$ and $g(a) = 0$. By supposition, $\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$ and so both the extended functions f and g are continuous at $x = a$.

But by applying *Cauchy Mean Value Theorem* (Theorem 36) to f, g on the interval $[a, x]$, we have for any x such that $a < x < a + \delta$, there is a point c in the interval (a, x) such that

$$\frac{f'(c)}{g'(c)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f(x)}{g(x)}$$

since $f(a) = g(a) = 0$.

Hence, $\left|\frac{f(x)}{g(x)} - l\right| = \left|\frac{f'(c)}{g'(c)} - l\right| < \varepsilon$ by (1) since $a < c < x < a + \delta$. Therefore, $\lim_{x\to a^+} \frac{f(x)}{g(x)} = l$.

This completes the proof of part (1).

Similar argument applies to part (2).

(B) We may assume f and g are functions differentiable for all $x \geq K$ for some real number $K > 0$ and that $g'(x) \neq 0$ for all $x > K$. (Just replaced K by $K' > K$ and renamed K' as K.) Let the following transformation H: $[K, \infty) \to (0, 1/K]$ be defined by $H(t) = 1/t$ for t in $[K, \infty)$. This devise would transform our problem to one of the type of (A) part (1).

Define Q : $(0, 1/K) \to \mathbf{R}$ by $Q(x) = \frac{f(\frac{1}{x})}{g(\frac{1}{x})}$ for x in. Note that $.\lim_{t\to\infty} H(t) = 0$. Then by Theorem 54 Chapter 3 if $\lim_{x\to 0^+} Q(x) = \lim_{x\to 0^+} \frac{f(\frac{1}{x})}{g(\frac{1}{x})}$ exist (as a finite number), then $\lim_{t\to\infty} Q(H(t)) = \lim_{x\to 0^+} Q(x) = \lim_{x\to 0^+} \frac{f(\frac{1}{x})}{g(\frac{1}{x})}$. But $\lim_{t\to\infty} Q(H(t)) = \lim_{t\to\infty} \frac{f(t)}{g(t)}$. That means if $\lim_{x\to 0^+} \frac{f(\frac{1}{x})}{g(\frac{1}{x})}$ exists, then $\lim_{t\to\infty} \frac{f(t)}{g(t)}$ exists and the two limits are the same. Conversely, if $\lim_{t\to\infty} \frac{f(t)}{g(t)}$ exists, then $\lim_{x\to 0^+} \frac{f(\frac{1}{x})}{g(\frac{1}{x})} = \lim_{t\to\infty} \frac{f(t)}{g(t)}$. Now by Theorem 53 part(1), $\lim_{x\to 0^+} f(\frac{1}{x}) = \lim_{x\to\infty} f(x) = 0$ and $\lim_{x\to 0^+} g(\frac{1}{x}) = \lim_{x\to\infty} g(x) = 0$. Therefore, by (A) part (1), $\lim_{x\to 0^+} \frac{f(\frac{1}{x})}{g(\frac{1}{x})}$ exists if the limit $\lim_{x\to 0^+} \frac{f'(\frac{1}{x})(-\frac{1}{x^2})}{g'(\frac{1}{x})(-\frac{1}{x^2})} = \lim_{x\to 0^+} \frac{f'(\frac{1}{x})}{g'(\frac{1}{x})}$ exists. But by a similar argument $\lim_{x\to 0^+} \frac{f'(\frac{1}{x})}{g'(\frac{1}{x})}$ exists if and only if $\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$ exists and the two limits are the same. It then follows that if $\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$ exists, then $\lim_{x\to+\infty} \frac{f(x)}{g(x)} = \lim_{x\to+\infty} \frac{f'(x)}{g'(x)}$.

Corollary 38 (L' Hôpital's Rule). Suppose f and g are continuous on $[a, b]$ with x_0 in (a, b). Suppose f and g are differentiable on (a, b) except possibly at x_0 and that $g'(x) \neq 0$ for all x in $(a, b)-\{x_0\}$. If $f(x_0) = g(x_0) = 0$, then $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}$, provided the second limit exists.

Proof. Just put Theorem 37 (A) part(1) and part(2) together.

Example 39.

(1) $\lim_{x\to 0} \frac{e^x - 1 - x - \frac{1}{2}x^2}{x^2} = \lim_{x\to 0} \frac{e^x - 1 - x}{2x} = \lim_{x\to 0} \frac{e^x - 1}{2} = 0$ by applying L'Hôpital's rule twice.

(2) $\lim_{x\to 1} \frac{\ln(x) - x + 1}{(x-1)^2} = \lim_{x\to 1} \frac{\frac{1}{x} - 1}{2(x-1)} = \lim_{x\to 1} \frac{-\frac{1}{x^2}}{2} = -\frac{1}{2}$ by applying L'Hôpital's rule twice.

Remark.

There are other versions of *L' Hôpital's Rule* , including the infinity/infinity versions, i.e., when the limit ($\lim_{x\to a^+} f(x)$ or $\lim_{x\to a^+} g(x)$ *or* $\lim_{x\to a^-} f(x)$ or $\lim_{x\to a^-} g(x)$ *or* $\lim_{x\to \pm\infty} f(x)$ *or* $\lim_{x\to \pm\infty} g(x)$) is either $+\infty$ or $-\infty$ and the infinity versions where the conclusion is the limit equals to either $+\infty$ or $-\infty$. For the various forms and generalization, see my article, "L' Hôpital's Rule and A generalized Version" on my Calculus web.

Next we shall describe an analytic consequence of the Cauchy Mean Value Theorem. Let I be an open interval. Let $f : I \to \mathbf{R}$ be a function. Suppose f is differentiable on I and its derivative $f' : I \to \mathbf{R}$ is again differentiable. Then we say f is twice differentiable or f has two derivatives. The second derivative is the derivative of f' and is denoted by $f'' : I \to \mathbf{R}$. We now use the notation $f^{(1)}$ for the derivative and $f^{(2)}$ for the second derivative. Now we define inductively the meaning of k-th derivative. Suppose for a positive integer k, the k-th derivative of f is defined and denoted by $f^{(k)} : I \to \mathbf{R}$. If $f^{(k)} : I \to \mathbf{R}$ is differentiable, then we say f has k+1 derivatives or is k+1 times differentiable and we define $f^{(k+1)} : I \to \mathbf{R}$ to be the derivative of $f^{(k)}$. We denote $f : I \to \mathbf{R}$ also by $f^{(0)}$.

Theorem 40. Let I be an open interval and let n be a positive integer.
Suppose $f : I \to \mathbf{R}$ has n derivatives. Let x_0 be a point in I. Suppose that

$$f^{(k)}(x_0) = 0 \text{ for } 0 \le k \le n-1.$$

Then for each point $x \ne x_0$, there exists a point ξ strictly between x and x_0 , such that

$$f(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

Proof. Let g : $I \to \mathbf{R}$ be defined by $g(x)=(x - x_0)^n$ for x in I. Then, for each $1 \le k \le n$,

$g^{(k)}(x) = n(n-1)\dots(n-k+1)(x - x_0)^{n-k}$. Thus.

$$g^{(k)}(x_0) = 0 \text{ for } 0 \le k \le n-1 \qquad \text{-------------------------- (1)}$$

and $g^{(n)}(x) = n!$ for any x in I.

We shall now apply the Cauchy Mean Value Theorem repeatedly.

Take any x a point in I not equal to x_0. We may assume that $x > x_0$. Then

$$\frac{f(x)}{g(x)} = \frac{f(x) - f^{(0)}(x_0)}{g(x) - g^{(0)}(x_0)}$$

since $f^{(0)}(x_0) = g^{(0)}(x_0) = 0$. Observe that f and g are both continuous and differentiable on $[x_0, x]$ and $g^{(1)}(x) \ne 0$ for x in (x_0, x). Therefore, applying the Cauchy Mean Value Theorem to the functions f and g restricted to $[x_0, x]$, we conclude that there exists x_1 in (x_0, x) such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f^{(0)}(x_0)}{g(x) - g^{(0)}(x_0)} = \frac{f^{(1)}(x_1)}{g^{(1)}(x_1)}.$$

If $n = 1$, then we are done (and this is just the Mean Value Theorem). If $n > 1$, then $f^{(1)}(x_0) = g^{(1)}(x_0) = 0$ and we can apply the Cauchy Mean Value Theorem to the functions $f^{(1)}$ and $g^{(1)}$ restricted to $[x_0, x_1]$, to give a point x_2 in (x_0, x_1) such that

$$\frac{f^{(1)}(x_1)}{g^{(1)}(x_1)} = \frac{f^{(1)}(x_1) - f^{(1)}(x_0)}{g^{(1)}(x_1) - g^{(1)}(x_0)} = \frac{f^{(2)}(x_2)}{g^{(2)}(x_2)}.$$

By (1), $f^{(k)}(x_0) = g^{(k)}(x_0) = 0$ for $0 \le k \le n-1$. We can then apply the Cauchy Mean Value Theorem repeatedly as above n-1 times, to get x_{n-1} in (x_0, x_{n-2}) such that

$$\frac{f^{(n-2)}(x_{n-2})}{g^{(n-2)}(x_{n-2})} = \frac{f^{(n-2)}(x_{n-2}) - f^{(n-2)}(x_0)}{g^{(n-2)}(x_{n-2}) - g^{(n-2)}(x_0)} = \frac{f^{(n-1)}(x_{n-1})}{g^{(n-1)}(x_{n-1})}.$$

Then since $f^{(n-1)}(x_0) = g^{(n-1)}(x_0) = 0$, we can apply the Cauchy Mean Value Theorem to the functions $f^{(n-1)}$ and $g^{(n-1)}$ restricted to $[x_0, x_{n-1}]$, to give a point x_n in (x_0, x_{n-1}) such that

$$\frac{f^{(n-1)}(x_{n-1})}{g^{(n-1)}(x_{n-1})} = \frac{f^{(n-1)}(x_{n-1}) - f^{(n-1)}(x_0)}{g^{(n-1)}(x_{n-1}) - g^{(n-1)}(x_0)} = \frac{f^{(n)}(x_n)}{g^{(n)}(x_n)} = \frac{f^{(n)}(x_n)}{n!}.$$

Therefore, taking ξ to be x_n , which is obviously strictly between x and x_0 such that

$$\frac{f(x)}{g(x)} = \frac{f^{(n)}(\xi)}{n!}.$$

Hence, $f(x) = \dfrac{f^{(n)}(\xi)}{n!} g(x) = \dfrac{f^{(n)}(\xi)}{n!}(x - x_0)^n$.

If $x < x_0$. the above argument applies similarly to give ξ strictly between x and x_0 such that

$$f(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

This completes the proof.

4.7 Taylor Polynomials, Taylor's Theorem

One of the triumph of Calculus is the approximation of function by polynomials. It allows computation to be done efficiently.

We shall now use Theorem 40 to prove the Taylor's Theorem with remainder.

Order of Contact of Two Functions.

Definition 41. Let I be an open interval containing the point x_0.
Two functions $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$ are said to have *contact of order* 0 at x_0 if $f(x_0) = g(x_0)$. For a positive integer n, the functions f and g are said to have *contact of order n* at x_0 provided $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$ have n derivatives and

$$f^{(k)}(x_0) = g^{(k)}(x_0) \text{ for } 0 \le k \le n.$$

Example 42.
Let $f(x) = \sqrt{2 - x^2}$ and $g(x) = e^{1-x}$ for x in $(0, \sqrt{2})$. Then $f^{(0)}(1) = g^{(0)}(1) = 1$ and $f^{(1)}(1) = g^{(1)}(1) = -1$ but $f^{(2)}(1) = -2 \neq g^{(2)}(1) = 1$. Therefore, $f : (0, \sqrt{2}) \to \mathbf{R}$ and $g : (0, \sqrt{2}) \to \mathbf{R}$ have contact of order 1 at $x = 1$ but do not have contact of order 2 at 1.

Let I be a neighborhood of the point x_0 and $f : I \to \mathbf{R}$ a function. Let n be a non-negative integer. Suppose f has $(n+1)$ derivatives. Then the n-th degree *Taylor's polynomial* of f at x_0 and the function $f : I \to \mathbf{R}$ have contact of order n at x_0.

Let us see how the Taylor polynomial can be assembled. For a positive integer k, let $g(x) = (x - x_0)^k$. Then we have

$$g^{(j)}(x) = k(k-1)\ldots(k-j+1)(x - x_0)^{k-j} \text{ for } 1 \le j \le k.$$

In particular, for all x, $g^{(k)}(x) = k!$ and $g^{(j)}(x) = 0$ for $j > k$ and $g^{(j)}(x_0) = 0$ for $1 \le j < k$. Thus we may write this in the more familiar form

$$\frac{d^j}{dx^j}[(x-x_0)^k] \,\Big|_{x=x_0} = \begin{cases} k! \text{ if } j = k \\ 0 \text{ if } j \ne k \end{cases} . \quad \text{-----------------------} \quad \text{(D)}$$

Proposition 43. Let I be an open interval containing the point x_0 and $f : I \to \mathbf{R}$ a function. Let n be a non-negative integer. Suppose f has n derivatives. Then there is a unique polynomial function of degree at most n that has contact of order n with the function $f : I \to \mathbf{R}$ at x_0. This polynomial is defined by

$$p_n(x) = f(x_0) + \frac{1}{1!}(x-x_0)f'(x_0) + \cdots + \frac{1}{k!}(x-x_0)^k f^{(k)}(x_0)$$
$$+ \cdots + \frac{1}{(n-1)!}(x-x_0)^{n-1}f^{(n-1)}(x_0) + \tfrac{1}{n!}(x-x_0)^n f^{(n)}(x_0) \quad \text{-----------------} \quad \text{(T)}$$

Proof. By (D)

$$\frac{d^j}{dx^j}[p_n(x)] \,\Big|_{x=x_0} = f^{(j)}(x_0) \text{ for } 1 \le j \le n.$$

Obviously, $p_n(x_o) = f(x_0)$. Therefore, $p_n : I \to \mathbf{R}$ and $f : I \to \mathbf{R}$ have contact of order n at x_0.

Now we shall show that the polynomial p_n is unique.

Write a general polynomial $p{:}I \to \mathbf{R}$ of degree n in terms of powers of $(x - x_0)$ as follows.

$$p(x) = c_0 + c_1(x-x_0) + \cdots + c_k(x-x_0)^k \cdots + c_{n-1}(x-x_0)^{n-1} + c_n(x-x_0)^n.$$

As before by (D), we get

$$\frac{d^j}{dx^j}[p(x)] \,\Big|_{x=x_0} = j!\, c_j \text{ for } 1 \le j \le n,$$

and $p(x_o) = c_0$. Then if p and f have contact of order n at x_0,

$$j!\, c_j = f^{(j)}(x_0) \text{ for } 1 \le j \le n \text{ and } c_0 = f(x_0).$$

Hence, $c_j = \frac{1}{j!}f^{(j)}(x_0)$ for $1 \le j \le n$. Therefore, $p = p_n$

The polynomial p_n given by (T) is called the *n-th degree Taylor polynomial of f at x_0*.

How good is the Taylor polynomial for approximation? The next theorem gives the error as a remainder term and the error term may be used as a gauge for the approximation.

Theorem 44. Taylor's Theorem (with remainder). Let I be an open interval containing the point x_0 and n be a non-negative integer. Suppose $f : I \to \mathbf{R}$ has n+1 derivatives. Then for any x in I,

$$f(x) = f(x_0) + \tfrac{1}{1!}(x-x_0)\,f'(x_0) + \cdots + \tfrac{1}{k!}(x-x_0)^k f^{(k)}(x_0)$$
$$+\cdots + \tfrac{1}{n!}(x-x_0)^n f^{(n)}(x_0) + R_n(x) \qquad \text{-----------------------------------} \qquad \text{(TR)}$$

where the error term $R_n(x)$ satisfies $R_n(x) = \frac{1}{(n+1)!}(x-x_0)^{n+1} f^{(n+1)}(\eta)$ for some η between x and x_0. (14) is known as the Taylor's expansion around x_0 and $R_n(x)$ is called the Lagrange form of the remainder.

Proof. Let g: $I \to \mathbf{R}$ be define by

$$g(x) = f(x) - f(x_0) - \tfrac{1}{1!}(x-x_0)\,f'(x_0)$$
$$-\cdots - \tfrac{1}{k!}(x-x_0)^k f^{(k)}(x_0) \cdots - \tfrac{1}{n!}(x-x_0)^n f^{(n)}(x_0).$$

Then since $f : I \to \mathbf{R}$ and $f(x_0) + \frac{1}{1!}(x-x_0)\,f'(x_0) + \cdots + \frac{1}{n!}(x-x_0)^n f^{(n)}(x_0)$ have contact of order n at x_0 , by Proposition 43, $g^{(k)}(x_0) = 0$ for $0 \le k \le n$.
Therefore, since g has n+1 derivatives, by Theorem 40, for any $x \neq x_0$, there exists η strictly between x and x_0 such that

$$g(x) = \frac{g^{(n+1)}(\eta)}{(n+1)!}(x-x_0)^{n+1}.$$

But $g^{(n+1)}(x) = f^{(n-+1)}(x)$ and so $g^{(n+1)}(\eta) = f^{(n+1)}(\eta)$. Consequently,

$$g(x) = \frac{f^{(n+1)}(\eta)}{(n+1)!}(x-x_0)^{n+1}.$$

Therefore, the remainder term $R_n(x) = g(x) = \dfrac{f^{(n+1)}(\eta)}{(n+1)!}(x-x_0)^{n+1}$.

This completes the proof.

Remark .
(1) There are other ways to make the same statement. For example: If $I = (a, b)$ and x, $x + h$ are in the interval (a, b), then there exists a θ with 0 < θ < 1 such that

$$f(x+h) = f(x) + h\,f'(x) + \tfrac{1}{2!}h^2 f''(x) + \cdots + \tfrac{1}{n!}h^n\, f^{(n)}(x) + \tfrac{1}{(n+1)!}h^{n+1}\, f^{(n+1)}(x+\theta h).$$

(2) The theorem says that the function can be considered as a real polynomial of degree n upto a remainder term. Thus to look upon f in this way the remainder term $R_n(x)$ counts.
We will not in general have that this remainder is small. Here the remainder after (n + 1) terms is
(i) of order $(x - x_0)^{n+1}$, which is very small if x and x_0 are close, and
(ii) depends on $f^{(n+1)}$.
Thus if we know $f^{(n+1)}$ is bounded, say $|\,f^{(n+1)}(x)| < M$, then the modulus of the remainder $|R_n(x)| \le \frac{M}{(n+1)!}|x-x_0|^{n+1}$.

Example 45.
Let $f(x) = e^x$.
Consider the expansion of $f(x)$ around $x_0 = 0$. The Taylor's expansion upto the power x^n is given by

$$f(x) = f(0) + \tfrac{1}{1!}(x-0)\,f'(0) + \tfrac{1}{2!}(x-0)^2\, f''(0) + \cdots + \tfrac{1}{n!}(x-0)^n\, f^{(n)}(0) + R_n(x)$$
$$= 1 + \tfrac{x}{1!} + \tfrac{1}{2!}x^2 + \cdots + \tfrac{1}{n!}x^n + R_n(x)\,,$$

where $R_n(x) = \frac{1}{(n+1)!}x^{n+1}e^{\eta}$ for some η between 0 and x, since $f^{(j)}(x) = e^x$ so that $f^{(j)}(0) = e^0 = 1$ for all non-negative integer j.

When $x_0 = 0$, then the Taylor's polynomial of f at 0 becomes

$$f(0) + x f'(0) + \cdots + \tfrac{1}{n!} x^n f^{(n)}(0)$$

and is called the n-th degree *Maclaurin polynomial*.

Example 46.
For $-1 < x < 1$, define $g(x) = \ln\left(\frac{1+x}{1-x}\right)$. We shall use this to compute ln(1.2).
First note that $g(\frac{1}{11}) = \ln(1.2)$. We shall obtain a MacLaurin polynomial of degree 4 for g.

Notice that $g(x) = \ln(1+x) - \ln(1-x)$, $g'(x) = (1+x)^{-1} + (1-x)^{-1}$,
$g''(x) = -(1+x)^{-2} + (1-x)^{-2}$, $g^{(3)}(x) = 2(1+x)^{-3} + 2(1-x)^{-3}$,
$g^{(4)}(x) = -6(1+x)^{-4} + 6(1-x)^{-4}$, $g^{(5)}(x) = 24(1+x)^{-5} + 24(1-x)^{-5}$.
Hence, $g(0) = 0$, $g'(0) = 2$, $g''(0) = 0$, $g^{(3)}(0) = 4$ and $g^{(4)}(0)=0$. Thus the MacLaurin polynomial of degree 4 is given by $p_4(x) = g'(0)x + \frac{1}{3!} g^{(3)}(0)x^3 = 2x + \frac{2}{3}x^3$. Now the remainder $|R_4(x)| = \left|\frac{g^{(5)}(\xi)}{5!} x^5\right|$ for some ξ between 0 and x. Therefore for $x = \frac{1}{11}$,

$$\left|\frac{g^{(5)}(\xi)}{5!}\right| = \tfrac{1}{5}[(1+\xi)^{-5} + (1-\xi)^{-5}] < \frac{1}{5}\left(1 + \frac{1}{(\frac{10}{11})^5}\right).$$

Thus $\left|R_4(\frac{1}{11})\right| < \frac{1}{5}\left[1 + \frac{1}{(\frac{10}{11})^5}\right] \cdot \left(\frac{1}{11}\right)^5 = \frac{1}{5}\left(\frac{1}{(11)^5} + \frac{1}{(10)^5}\right) < 0.000004$. Therefore, to compute ln(1.2) upto to four decimal places, we can use the MacLaurin's polynomial of degree 4. Hence, $\ln(1.2) \approx p_4(\frac{1}{11}) = 2 \cdot \frac{1}{11} + \frac{2}{3}(\frac{1}{11})^3 \approx 0.1823$ upto 4 decimal places.

4.8 Intermediate Value Theorem for Derivative

We close this chapter with a property of the derived function of a differentiable function $f:[a, b] \to \mathbf{R}$,namely the intermediate value property of f'.

Theorem 47 (Darboux Theorem) Let I be an interval and suppose $f : I \to \mathbf{R}$ is a differentiable function. Let a, b be two points in I such that $a < b$. Suppose $f'(a) \neq f'(b)$. Then for any value γ strictly between $f'(a)$ and $f'(b)$, there is a point c in (a, b) such that $f'(c) = \gamma$.

Proof. Let us define the following function g: $I \to \mathbf{R}$ by $g(x) = f(x) - \gamma x$ for x in I. Then g is differentiable and $g'(x) = f'(x) - \gamma$. If we can show that g has either a relative maximum or a relative minimum at a point c in (a, b), then we are done. Consider the function $h : [a, b] \to \mathbf{R}$, the restriction of g to the closed interval, $[a, b]$. Then since g is differentiable, g is also continuous on $[a, b]$. Therefore, by the Extreme Value Theorem, the restriction of g, h attains both its maximum and minimum in $[a, b]$. We shall show that at least one of the maximum or minimum of h occurs in the interior of $[a, b]$. Suppose $h(a)$ is the maximum and $h(b)$ is the minimum. Then for all x in $[a, b]$, $h(x) \le h(a)$ and $h(x) \ge h(b)$. Hence for all x in $[a, b]$, $f(x) - \gamma x \le f(a) - \gamma a$, *that is,* $f(x) - f(a) \le \gamma x - \gamma a$. Therefore, for all x in $[a, b]$ and $x \neq a$, $(f(x) - f(a))/(x - a) \le \gamma$. Since f is differentiable at a,

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

Since for all x in $[a, b]$, $h(x) \geq h(b)$, $f(x) - \gamma x \geq f(b) - \gamma b$ for all x in $[a, b]$, that is, $f(b) - f(x) \leq \gamma b - \gamma x$. Consequently, for all x in $[a, b)$, $(f(b) - f(x))/(b - x) \leq \gamma$ *since* $b - x >$ *for* $x < b$. Similarly since f is differentiable at b,

$$f'(b) = \lim_{x \to b^-} \frac{f(b) - f(x)}{b - x} \leq \gamma.$$

Therefore, we can conclude that $f'(a)$ and $f'(b)$ are both less than or equal to γ, contradicting that γ is strictly between $f'(a)$ and $f'(b)$. Similarly, if $h(a)$ is the minimum and $h(b)$ is the maximum, we can show in like manner that $f'(a)$ and $f'(b)$ are both greater than or equal to γ, giving a contradiction to that γ is strictly between $f'(a)$ and $f'(b)$. We have thus shown that one of the maximum or minimum of h must occur at *a point* c in (a, b). Since $h(c)$ is also a relative extremum and h is differentiable, $h'(c) = f'(c) - \gamma = 0$ and so $f'(c) = \gamma$. (See Theorem 11) This completes the proof.

Corollary 48. Let I be an interval and suppose $f : I \to \mathbf{R}$ is a differentiable function. Then the image of the derived function of f, $f'(I)$ is also an interval.

Proof. The proof is similar to the proof that the continuous image of an interval is an interval. Let $f'(a) \neq f'(b)$ be in $f'(I)$. We may assume that $f'(a) < f'(b)$. Theorem 47 says that for any γ such that $f'(a) < \gamma < f'(b)$, $\gamma \in f'(I)$. Hence the interval $[f'(a), f'(b)] \subseteq f'(I)$. Therefore, by the usual characterization of an interval, $f'(I)$ is an interval.

Here is a useful application:

Theorem 49. Suppose f is differentiable on an interval I (not necessarily bounded). If the derived function f' is non-zero on I, then f' is of constant sign, i.e., for all x in I, $f'(x) > 0$ *or* for all x in I, $f'(x) < 0$.

Proof. Suppose f' is not of constant sign. Then there exist x and y in I such that $f'(x) > 0$ and $f'(y) < 0$. Thus 0 is an intermediate value between $f'(x)$ and $f'(y)$. Therefore, by Darboux's Theorem (Theorem 47), there exists a point c between x and y such that $f'(c) = 0$. This contradicts that f' is non-zero on I and so f' must be of constant sign.

Remark.
Theorem 49 is a useful tool. Often we need to know that there are no sign changes in a neighbourhood of a point where the derivative is not zero as in the proof of a weaker form of the Cauchy Mean Value Theorem in my article " L' Hôpital's Rule and A generalized Version" .