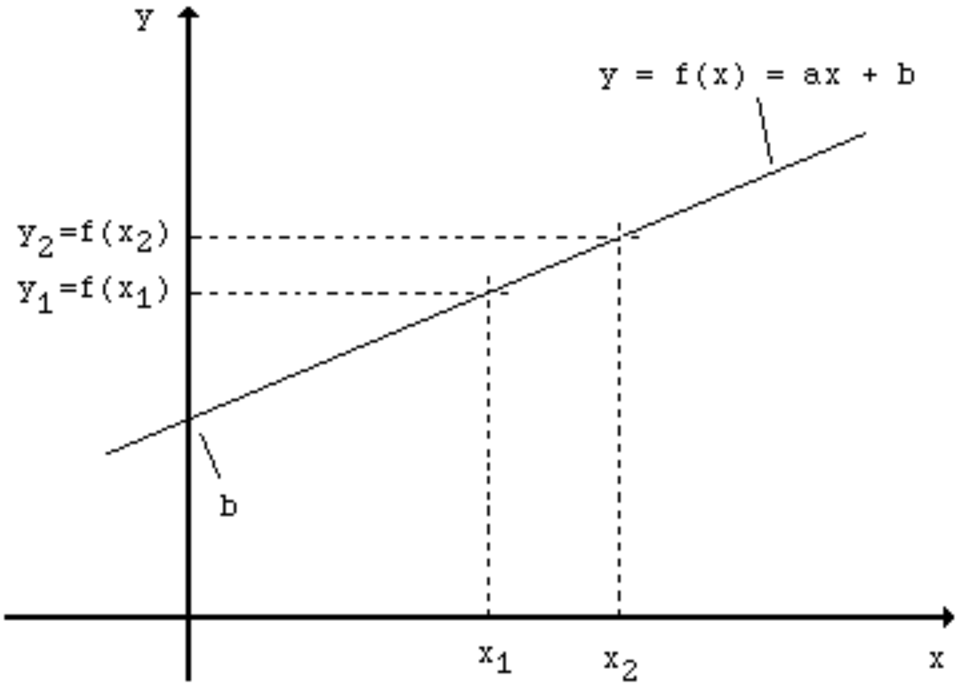
A One Day Tour of Calculus

The rule $y = f(x) = ax + b$ has a graph called a **straight line** as shown in the figure below:



where

$$a = \text{slope of the line} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

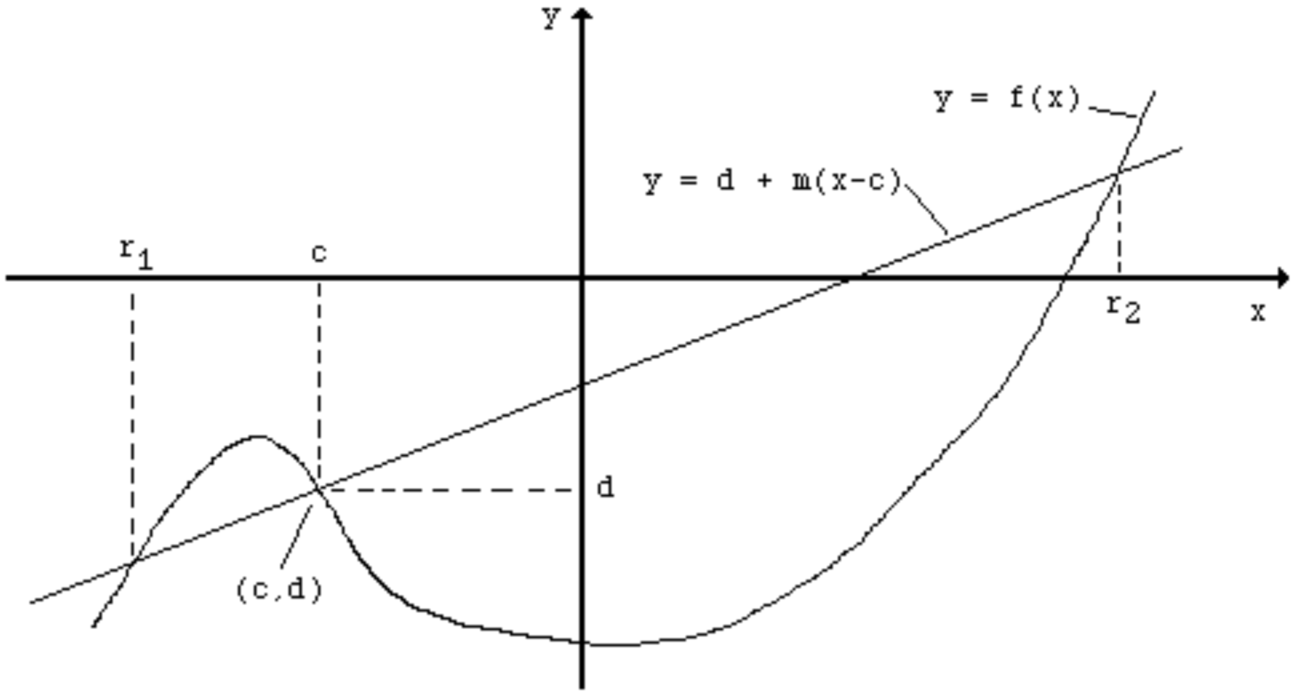
and

$$b = \text{intercept on the y-axis} = f(0)$$

Now, the problem of finding the equation of that line which is tangent to a function $y = f(x)$ at the point (c,d), where $d = f(c)$ is central to CALCULUS. The most general straight line passing through the point (c,d) is given by $y - d = m(x - c)$.

Proof: This certainly represents a straight line since we can rewrite it as $y = mx + (d - mc)$, which is the standard form of the equation of a straight line with slope = m and intercept $(d - mc)$. The line passes through the point (c,d) since when we choose $x = c$, we have $y = mc + (d - mc) = d$. This completes the proof. It represents the most general straight line through (c,d) since its slope m is arbitrary.

Suppose we plot two graphs as shown in the figure below:



where we have indicated three intersections labelled by their x –values, namely, r_1 , c and r_2 . More than three intersections are possible, but that case only represent an unnecessary complication for the present discussion.

The intersections r_1 , c and r_2 represent the zeroes of the function $p(x) = f(x) - d - m(x - c)$. Now, very close to the point $x = c$, $p(x)$ must have the form $p(x) = (x - c)g(x)$ since it equals 0 at $x = c$. Similarly, near $x = r_1$, $p(x)$ must have the form $p(x) = (x - r_1)g(x)$.

If we rotate (change the slope) of the line about the point (c, d) until the straight line becomes tangent to the curve at (c, d) , then, since this means that r_1 approaches c , we must have

$$\begin{aligned} p(x) &= (x - c)(x - r_1)d(x) \quad \text{near } x = c = r_1 \\ &= (x - c)(x - c)d(x) = (x - c)g(x) \end{aligned}$$

In other words, when the line is tangent to the curve at (c, d) we must have $g(c) = 0$. From the definition of $g(x)$, we then have

$$g(x) = \frac{p(x)}{x - c} = \frac{f(x) - d}{x - c} - m = q(x) - m$$

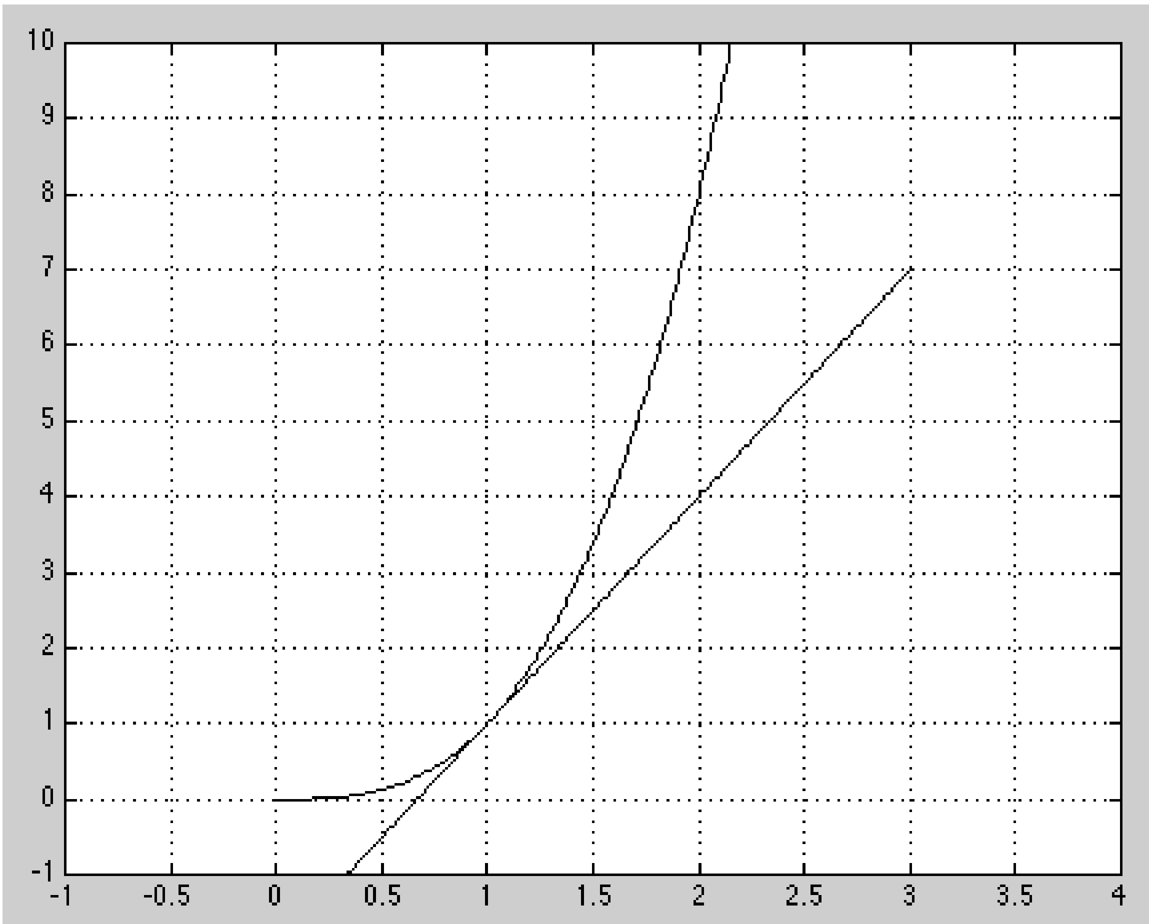
When $x = c$, this implies that $g(c) = 0 = q(c) - m$ or $m = q(c) =$ slope of the tangent line to $f(x)$ at (c, d), where

$$q(x) = \frac{f(x) - d}{x - c} = \frac{f(x) - f(c)}{x - c}$$

For simple functions this rule is easily applied. Consider the case $y = f(x) = x^3$. We have

$$q(x) = \frac{f(x) - f(c)}{x - c} = \frac{x^3 - c^3}{x - c} = \frac{(x - c)(x^2 + cx + c^2)}{x - c} = x^2 + cx + c^2$$

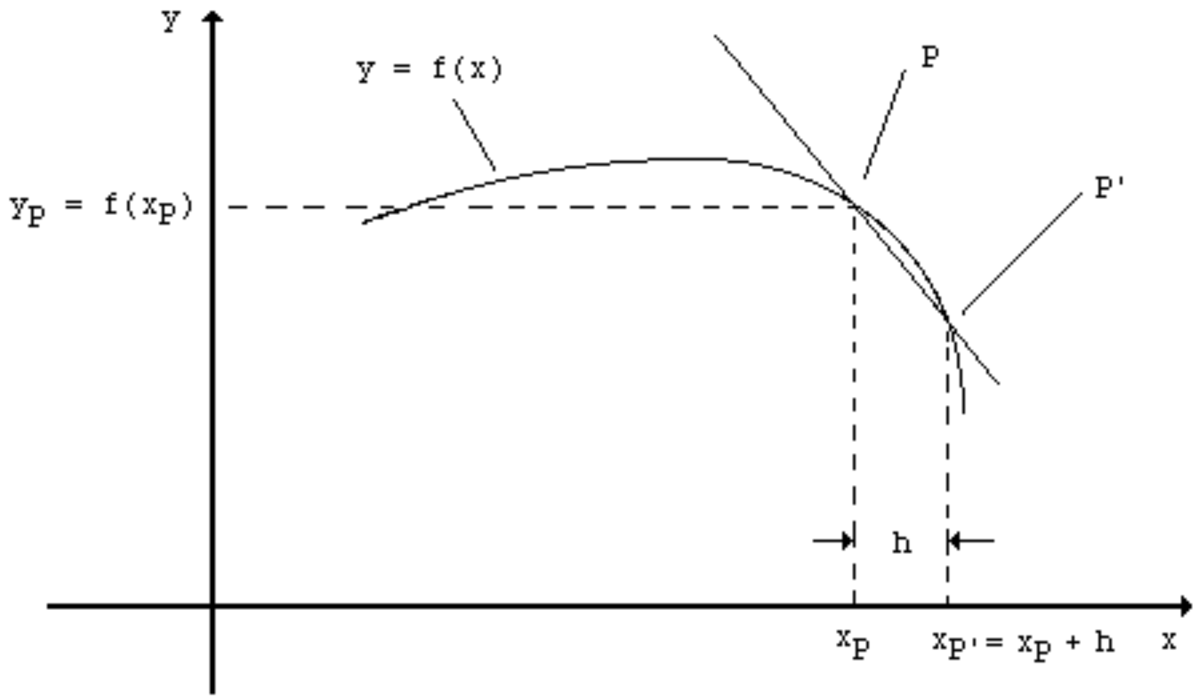
This implies that $m =$ slope of line tangent to $f(x) = x^3$ at the point $(c, c^3) = q(c) = 3c^2$. Then, the equation of the tangent line is $y - c^3 = 3c^2(x - c)$ or $y = 3c^2 x - 2c^3$(the tangent line). The case $c = 1$ is plotted below:



In this case, the tangent line at the point (1,1) is given by the line $y = 3x - 2$.

The Derivative

This procedure, while transparent, is hard to apply for more complicated functions. We now develop an alternative approach (called the **derivative**) which enables us to find the slope of the tangent line for arbitrary functions. Consider the figure below:



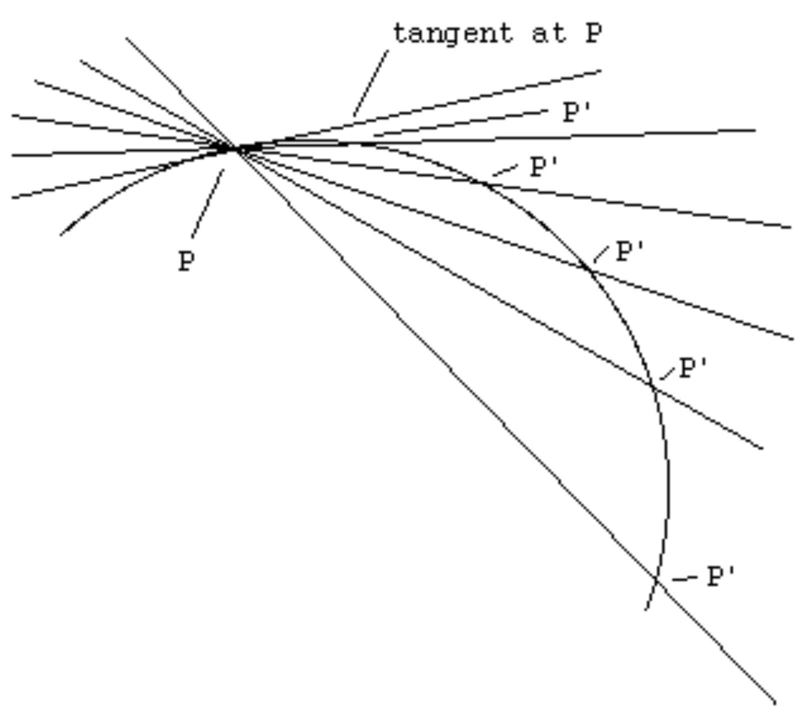
Now follow this procedure:

(1) Choose a point P corresponding to $(x_P, f(x_P))$

(2) Choose a second point P' such that $x_{P'} = x_P + h$

(3) Find the equation of the straight line through P and P'

(4) As P' approaches P, the slope of the line PP' approaches a limiting value equal to the slope of the tangent line at point P. This is shown schematically in the diagram below:



Now the slope of PP' is

$$m_h = \frac{f(x_{P'}) - f(x_P)}{x_{P'} - x_P} = \frac{f(x_P + h) - f(x_P)}{h}$$

and the slope of the tangent line at P is

$$m_P = \lim_{h\to 0} m_h = \lim_{h\to 0} \frac{f(x_P + h) - f(x_P)}{h}$$

To illustrate these ideas, let us return to our previous example $f(x) = x^3$. We then have

$$m_p = \lim_{h\to 0} \frac{(x_P + h)^3 - x_P^3}{h} = \lim_{h\to 0} \frac{x_P^3 + 3hx_P^2 + 3h^2 x_P + h^3 - x_P^3}{h}$$

$$= \lim_{h\to 0} \frac{3hx_P^2 + 3h^2 x_P + h^3}{h} = \lim_{h\to 0}(3x_P^2 + 3hx_P + h^2) = 3x_P^2$$

For the point $x_P = c$, we have $m = 3c^2$ as before.

Before proceeding, we state (without proof) some useful rules involving limits.

(1) For $c =$ constant $\lim_{h\to a} cF(h) = c\lim_{h\to a} F(h) = cF(a)$

(2) $\lim_{h\to a}(F(h) \pm G(h)) = \lim_{h\to a} F(h) \pm \lim_{h\to a} G(h) = F(a) \pm G(a)$

(3) $\lim_{h\to a}(F(h)G(h)) = (\lim_{h\to a} F(h))(\lim_{h\to a} G(h)) = F(a)G(a)$

(4) $\lim_{h\to a}\left(\frac{F(h)}{G(h)}\right) = \frac{\lim_{h\to a} F(h)}{\lim_{h\to a} G(h)} = \frac{F(a)}{G(a)}$ if $G(a) \neq 0$

In general, the derivative of the function $f(x)$ at the arbitrary point x is defined by

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

The derivative is also called the **rate of change** , i.e., $\frac{df}{dx} =$ rate of change of $f(x)$ with respect to x . From our previous discussion we have $f(q) =$ slope of the tangent line to the graph $y = f(x)$ at the point $(q, f(q))$.

Simple Example

$$f(x) = cx^2$$

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{c(x+h)^2 - cx^2}{h} = \lim_{h\to 0} \frac{2cxh + ch^2}{h} = \lim_{h\to 0}(2cx + h) = 2cx$$

This is the slope of the line tangent to $f(x) = cx^2$ at $(x, f(x))$.

Although this procedure is straightforward, we do not wish to do all of this work for every $f(x)$ that arises in this course. We need to develop some general rules that we can use.

Consider the function $f(x) = x^n$, which includesthe previous example. We have

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(x+\Delta x)^n - x^n}{\Delta x}$$

where we have made a change to standard notation, i.e., h to Δx. In order to evaluate this we need the following **algebraic** result:

$$(x+\Delta x)^n = x^n + nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}(\Delta x)^2 + \frac{n(n-1)(n-2)}{6}x^{n-3}(\Delta x)^3 + ...$$

This is the so-called **Binomial expansion**. It can be proved by simple multiplication. Using it we then have

$$f'(x) = \lim_{\Delta x \to 0} \frac{(x+\Delta x)^n - x^n}{\Delta x} = \lim_{\Delta x \to 0} \frac{\left(x^n + nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}(\Delta x)^2 + ...\right) - x^n}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}(\Delta x)^2 + ...}{\Delta x} = \lim_{\Delta x \to 0}\left(nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}\Delta x + ...\right) = nx^{n-1}$$

or $$\frac{d\left(x^n\right)}{dx} = nx^{n-1}$$

This result agrees with all the special case , that we did explicitly earlier.

Given the functions $u(x)$ and $v(x)$ we can now state several important derivative rules (we will prove a couple of them to illustrate the methods involved).

(1) $$\frac{d}{dx}(cu(x)) = c\frac{du}{dx}$$

(2) $$\frac{d}{dx}(u(x) \pm v(x)) = \frac{du}{dx} \pm \frac{dv}{dx}$$

(3) $$\frac{d}{dx}(u(x)v(x)) = v(x)\frac{du}{dx} + u(x)\frac{dv}{dx}$$

(4) $$\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) = \frac{v(x)\frac{du}{dx} - u(x)\frac{dv}{dx}}{(v(x))^2}$$

(5) $$\frac{d}{dx}\left(u^p\right) = pu^{p-1}\frac{du}{dx}$$

Another very important rule that is used all the time is called the **chain rule**. If $y = f(u)$ and $u = g(x)$, then we can define the **composite function** $y = f(g(x)) = H(x)$. We can then write

$$\frac{dH}{dx} = \frac{df(g(x))}{dx} = \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{\Delta x}$$
$$= \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{g(x+\Delta x) - g(x)} \frac{g(x+\Delta x) - g(x)}{\Delta x}$$
$$= \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{g(x+\Delta x) - g(x)} \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x}$$
$$= \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{g(x+\Delta x) - g(x)} \frac{dg}{dx}$$

Now as $\Delta x \to 0$, $g(x+\Delta x) \to g(x)$. Therefore we have $\Delta g = g(x+\Delta x) - g(x) \to 0$ as $\Delta x \to 0$. We can then write

$$\frac{dH}{dx} = \frac{df(g(x))}{dx} = \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{g(x+\Delta x) - g(x)} \frac{dg}{dx}$$
$$= \lim_{\Delta g \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{\Delta g} \frac{dg}{dx} = \frac{df}{dg} \frac{dg}{dx}$$

which is a very powerful and useful rule.

Example: Find the derivative of $y = g^3 + g - 5$ where $g = x^2 + 6x$. We have

$$\frac{dy}{dx} = \frac{dy}{dg} \frac{dg}{dx} = \left(3u^2 + 1\right)(2x+6) = \left(3\left(x^2 + 6x\right)^2 + 1\right)(2x+6)$$

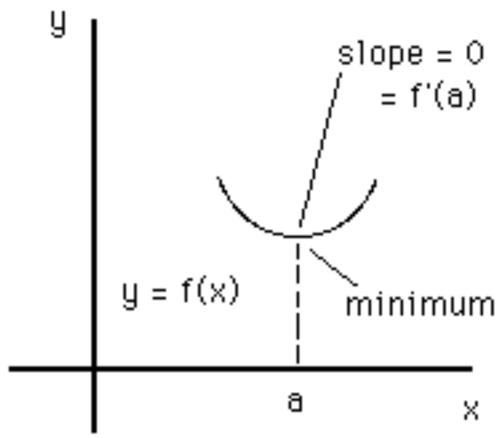
Properties Derivable from Derivatives

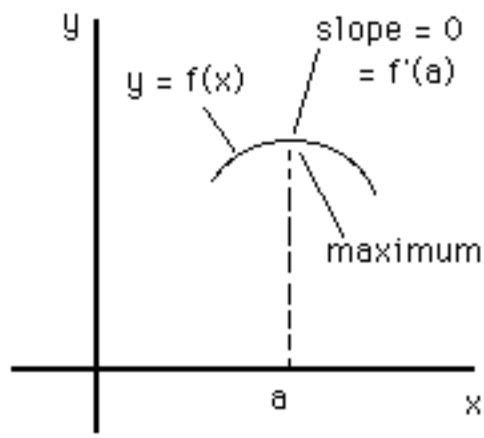
First, we must define **higher derivatives** . $f'(x) = \frac{df}{dx}$ is called the **first** derivative of $f(x)$ with respect to x . We define the **second** derivative by

$$f''(x) = \frac{df'(x)}{dx} = \frac{d}{dx}\left(\frac{df(x)}{dx}\right) = \frac{d^2 f(x)}{dx^2}$$

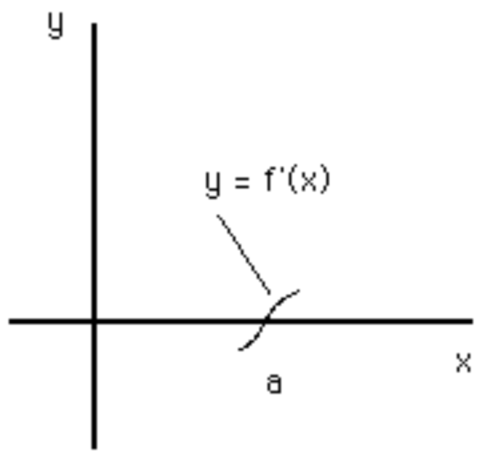
and so on for third, fourth, ... derivatives, etc.

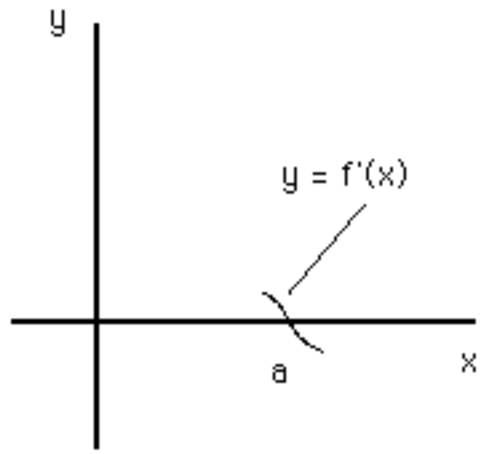
Now, if the graph of a function $f(x)$ has a minimum (maximum) at some point $x = a$, then $f'(a) = 0$ as can be seen in the figures below.



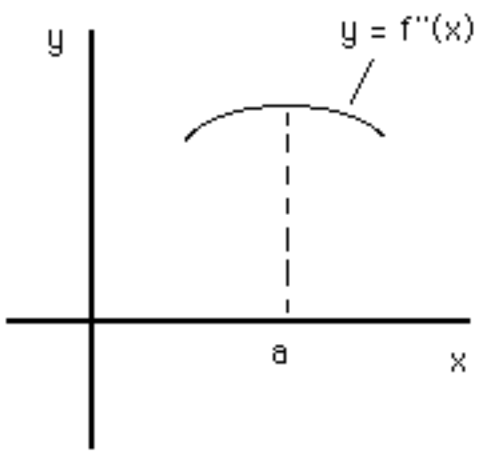


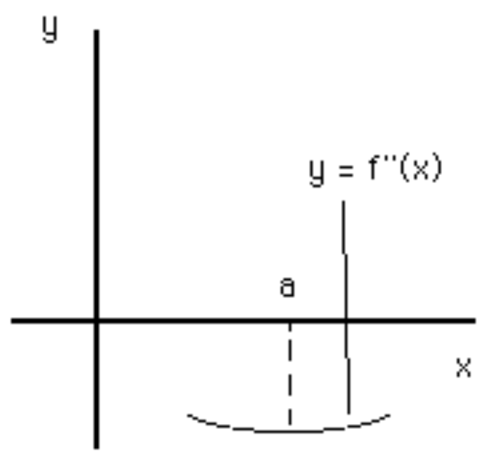
The slope of these graphs are clearly zero at the minimum and the maximum. If, in each case, we plot $f'(x)$ versus x, then the graphs would look like the figures below:





The graphs pass through zero at the minimum and the maximum. If, in each case we plot $f''(x)$ (which is just the slope of the $f'(x)$ graph) versus x, then the graphs would look like the figures below:





The second derivative is positive for a minimum and negative for a maximum. Summarizing, in words, **near a minimum**

$$f'(x) = 0 \text{ and } f''(x) > 0$$

and **near a maximum**

$$f'(x) = 0 \text{ and } f''(x) < 0$$

Some useful derivatives are given below:

$$\frac{d}{dx}\left(e^{ax}\right) = ae^{ax}, \frac{d}{dx}\left(e^{u(x)}\right) = e^{u(x)}\frac{du}{dx}, \frac{d}{dx}(\sin ax) = a\cos ax, \frac{d}{dx}(\cos ax) = -a\sin ax, \frac{d}{dx}(\log x) = \frac{1}{x}$$

This set of rules along with the rule derived earlier : $\frac{d}{dx}\left(x^n\right) = nx^{n-1}$ are all of the derivatives we need in this course.

Let us prove one of the rules to illustrate the method used. To start, we need to know how the factorial function involved is defined. We define

$$n! = n - factorial = (1)(2)(3)(4).......(n-1)(n)$$
$$0! = 1$$

If we also define the notation

$$\sum_{m=0}^{M} a_m y^m = a_0 y^0 + a_1 y^1 + a_2 y^2 + ... + a_M y^M$$

then the definition of the exponential function is

$$e^{ax} = 1 + ax + \frac{(ax)^2}{2} + \frac{(ax)^3}{6} + \frac{(ax)^4}{24} + ... = \sum_{n=0}^{\infty} \frac{(ax)^n}{n!}$$

We can then write (using the rules)

$$\frac{d}{dx}\left(e^{ax}\right) = \frac{d}{dx}\left(\sum_{n=0}^{\infty}\frac{(ax)^n}{n!}\right) = \sum_{n=0}^{\infty}\frac{a^n}{n!}\frac{d(x^n)}{dx} = \sum_{n=0}^{\infty}\frac{a^n}{n!}nx^{n-1}$$

$$= 0 + a + a\frac{ax}{2} + a\frac{(ax)^2}{6} + a\frac{(ax)^3}{24} + ... = a\sum_{n=0}^{\infty}\frac{(ax)^n}{n!}ae^{ax}$$

which completes the proof.

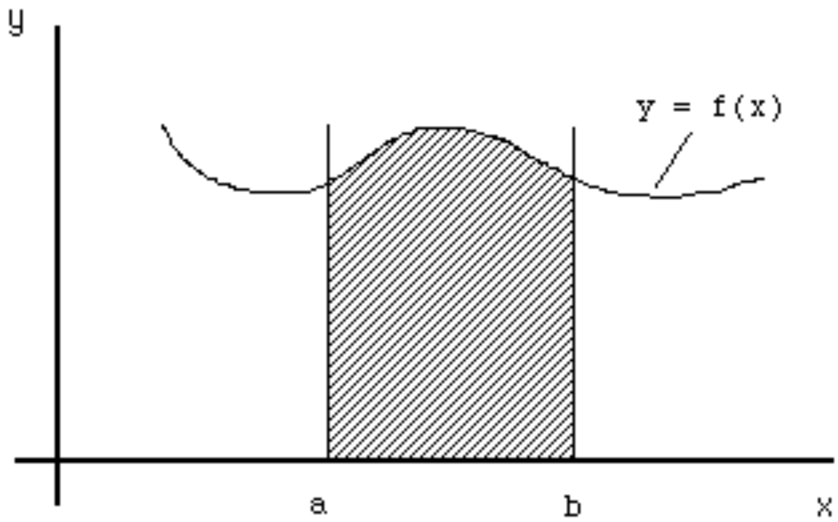
We will need one further definition. If we can write

$$h(x) = \frac{d}{dx}(g(x))$$

then $g(x) + c$, where $c =$ an arbitrary constant is called the **antiderivative** of $h(x)$, i.e., $g(x) + c$ is the function whose derivative is $h(x)$.

Integration

Suppose we now ask the following question: what is the area(**shaded region** in the figure below) under the curve $y = f(x)$ between $x = a$ and $x = b$?



A good **approximation** to this area is given by the following procedure:

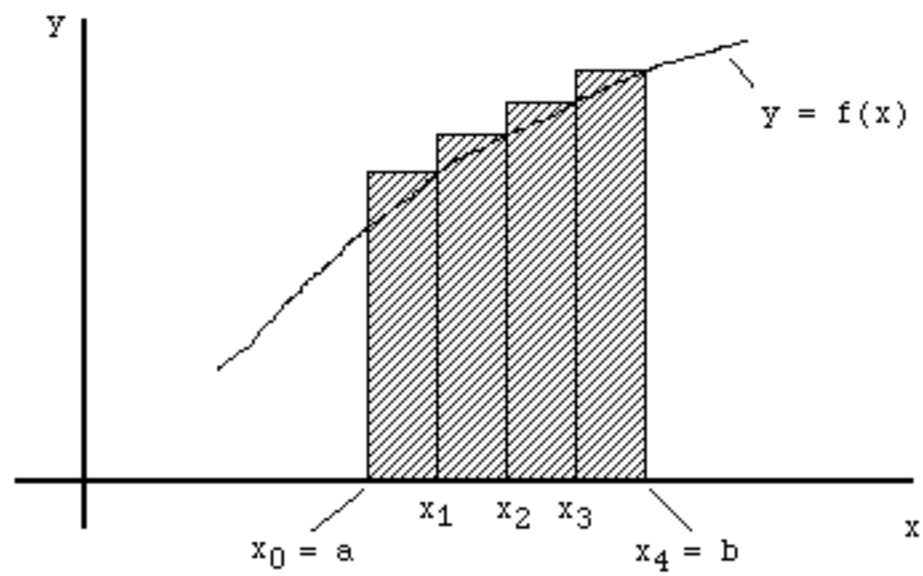
(1) divide the interval $a \le x \le b$ into N equal segments each of length $\Delta = \dfrac{b-a}{N}$

(2) define $x_k = a + k\Delta$ for $k = 1,2,3,4,......,N$

(3) calculate the corresponding values of $f(x)$, $f(x_k) = f(a+k\Delta)$ for $k = 1,2,3,4,......,N$

(4) then an approximation to the area is given by $AREA = \sum_{k=1}^{N} f(x_k)\Delta$

as shown in the figure below:



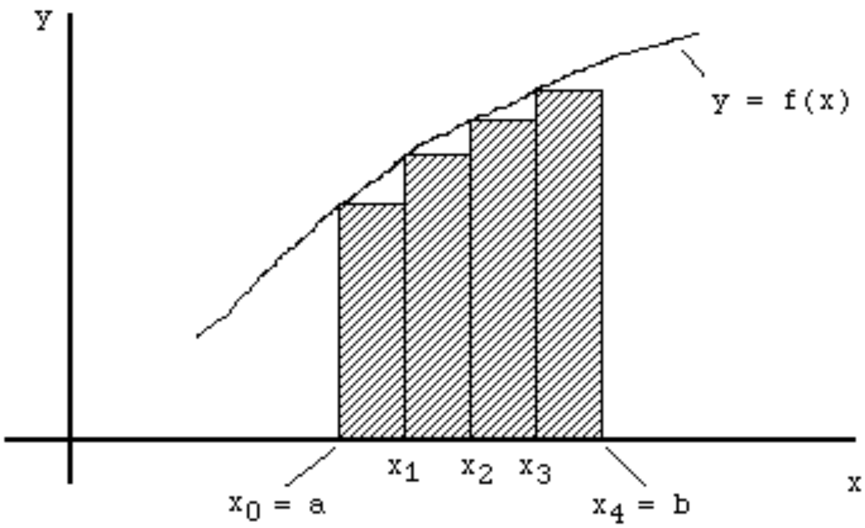
where for the case $N = 4$ we have

$$\Delta = \frac{b-a}{4}$$

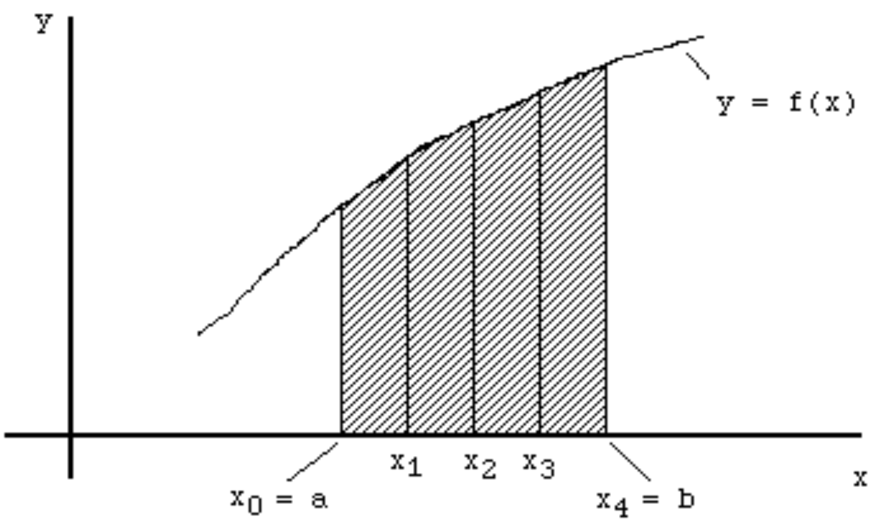
$$x_0 = a \,,\; x_1 = a + \Delta \,,\; x_2 = a + 2\Delta \,,\; x_3 = a + 3\Delta \,,\; x_4 = a + 4\Delta = a + (b-a) = b$$

As can be seen from the figure, our approximation for the area equals the sum of the shaded rectangles. In the case shown, the calculated area is **greater than** the actual area.

Alternatively, we could have used $AREA = \sum_{k=1}^{N} f(x_{k-1})\Delta$, which underestimates the area as shown in the figure below:



For an even better result, we could have used $AREA = \frac{1}{2}\sum_{k=1}^{N}\left(f(x_{k+1}) + f(x_{k-1})\right)\Delta$, which is called the **trapezoid rule**. The is rule calculates the area shown in the figure below:



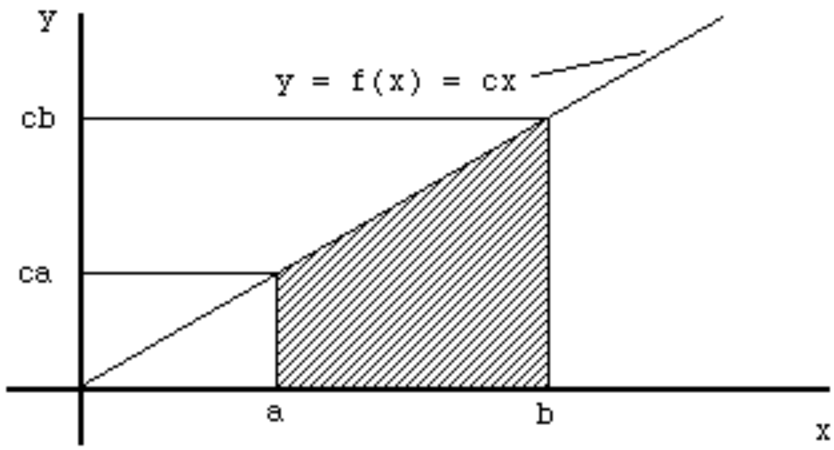
It uses straight lines between points on the actual curve rather than horizontal lines. It is clearly the best approximation of these three.

In the limit $N \to \infty$, all of these approximations give identical results (= to the actual area under the curve). The limit $N \to \infty$ is usually written as

$$AREA = \int_a^b f(x)dx = \text{integral of } f(x) \text{ from } a \text{ to } b = \lim_{N\to\infty}\sum_{k=1}^{N}\Delta f(x_k)$$

and a and b are called **limits of integration**.

Simple Integral: $f(x) = cx$ (a straight line) as shown in the figure below:



$$\int_a^b f(x)dx = \int_a^b cxdx = \lim_{N\to\infty}\sum_{k=1}^{N} f(x_k)\Delta = \lim_{N\to\infty}\sum_{k=1}^{N} cx_k\Delta = c\lim_{N\to\infty}\frac{b-a}{N}\sum_{k=1}^{N}(a+k\Delta)$$

$$= c\lim_{N\to\infty}\frac{b-a}{N}\sum_{k=1}^{N} a + c\lim_{N\to\infty}\frac{b-a}{N}\frac{b-a}{N}\sum_{k=1}^{N} k$$

$$= c\lim_{N\to\infty}\frac{b-a}{N}aN + c\lim_{N\to\infty}\frac{b-a}{N}\frac{b-a}{N}\frac{N(N+1)}{2}$$

$$= \lim_{N\to\infty} c(b-a) + \lim_{N\to\infty}\frac{1}{2}c(b-a)^2\left(1+\frac{1}{N}\right)$$

$$= c(b-a) + \frac{1}{2}c(b-a)^2 = \frac{1}{2}cb^2 - \frac{1}{2}ca^2 = \text{shaded area}$$

In this manner, we could evaluate any integral (find the area under the corresponding curve). The procedure quickly becomes very cumbersome and tedious, however. A better method is to realize that there is a connection between integrals and derivatives.

The Fundamental Theorem of Calculus

If $\frac{dF}{dx} = f(x)$,i.e., $F(x)+c$ the antiderivative of $f(x)$, then

$$\int_a^b f(x)dx = F(b) - F(a) = F(x)\Big|_{x=a}^{x=b} = \text{definite integral (a number)}$$

Alternatively, another way of saying the same thing is to use the definition

$$\int f(x)dx = F(x) + c = \text{indefinite integral (function of } x\text{)}$$

The indefinite integral represents the most general antiderivative of $f(x)$.

Examples:

(1) Since $\cos x = \frac{d}{dx}(\sin x)$, we have $\int_a^b \cos x dx = \sin b - \sin a$ or $\int \cos x dx = \sin x + c$.

(2) Since $x^n = \frac{d}{dx}\left(\frac{x^{n+1}}{n+1}\right)$, $N \neq -1$, we have $\int_a^b x^n dx = \left.\frac{x^{n+1}}{n+1}\right|_{x=a}^{x=b} = \frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}$ or $\int x^n dx = \frac{x^{n+1}}{n+1} + c$.

We now introduce a convenient notation and define the **differential**. Define

$$du = u'(x)dx = \left(\frac{du}{dx}\right)dx \qquad \text{(this is NOTcancellation)}$$

where the quantity dx is called the **differential of x** . Using this notation our last example can be written as follows: let $y = qx$, which implies that $dy = \left(\frac{dy}{dx}\right)dx = qdx$ or

$$\int_a^b \cos(qx)dx = \frac{1}{q}\int_{qa}^{qb} \cos y dy = \frac{1}{q}\int_{qa}^{qb} \frac{d}{dy}(\sin y)dy = \frac{1}{q}(\sin(qb) - \sin(qa))$$

Note the change in the integration limits at one point.

Some Useful Properties of the Differential

$$d(u+v) = du + dv$$
$$d(cu) = cdu$$
$$d(f(u)) = \left(\frac{df}{du}\right)du$$
$$d(uv) = udv + vdu$$

The trick to most integrals is to evaluate the antiderivative, i.e., Let us call $x = \sin\theta$. We can write $dx = \cos\theta d\theta = d\theta\sqrt{1-\sin^2\theta} = d\theta\sqrt{1-x^2}$ so that $\frac{dx}{\sqrt{1-x^2}} = d\theta$ and

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

Similarly, $x = \tan\theta$ We can write $dx = \left(1+\tan^2\theta\right)d\theta = d\theta\left(1+x^2\right)$

$$\int \frac{dx}{1+x^2} = \theta = \tan^{-1} x$$

Algebraic manipulation works in many cases,

$$\int \frac{dx}{1-x^2} = \int \frac{dx}{(1+x)(1-x)} = \int \frac{dx}{2}\left(\frac{1}{1+x} + \frac{1}{1-x}\right) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)$$

A very useful result is a procedure called **integration by parts**. It goes as follows: since

$$\frac{d}{dx}\big(u(x)v(x)\big) = v(x)\frac{du}{dx} + u(x)\frac{dv}{dx}$$

we have

$$\int_a^b u(x)\frac{dv}{dx}dx = \int_a^b \frac{d}{dx}\big(u(x)v(x)\big)dx - \int_a^b v(x)\frac{du}{dx}dx$$

$$= u(b)v(b) - u(a)v(a) - \int_a^b v(x)\frac{du}{dx}dx$$

An example is shown below:

$$\int_a^b x\cos x dx = \int_a^b \frac{d}{dx}(x\sin x) - \int_a^b \sin x dx = b\sin b - a\sin a + \cos b - \cos a$$

Another useful result is as follows: since $\frac{d}{dx}\big(F(g(x))\big) = \frac{dF}{dg}\frac{dg}{dx}$, we have

$$\int_a^b \frac{dF}{dg}\frac{dg}{dx}dx = \int_a^b \frac{d}{dx}\big(F(g(x))\big)dx = F(g(b)) - F(g(a))$$

Finally, we discuss a different kind of derivative, which is based on the standard derivative. It is called a **partial derivative**. Suppose we have a function z of two independent variables x and y that is given by the relationship $z = f(x,y)$. We then define the partial derivative of z with respect to x by

$$\frac{\partial z}{\partial x} = \frac{\partial f(x,y)}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}$$

and similarly for the partial derivative of z with respect to y , This corresponds to taking the normal derivative of $z = f(x,y)$ with respect to x while holding the variable y constant (if there are more than two independent variables, then we would hold all of the other variables constant). We can easily see this result from a few examples:

(1)

$$z = f(x,y) = xy$$

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \to 0} \frac{(x+\Delta x)y - xy}{\Delta x} = \lim_{\Delta x \to 0} \frac{y\Delta x}{\Delta x} = \lim_{\Delta x \to 0} y = y$$

This is the same result we would get if we assumed that $y =$ constant and take the normal derivative with respect to x , i.e.,

$$\frac{d}{dx}(xy) = y\frac{d}{dx}(x) = y$$

(2)

$$z = f(x,y) = e^{xy}$$

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \to 0} \frac{e^{(x+\Delta x)y} - e^{xy}}{\Delta x} = \lim_{\Delta x \to 0} \frac{(e^{y\Delta x} - 1)e^{xy}}{\Delta x} = \lim_{\Delta x \to 0} \frac{e^{xy}(y\Delta x + \frac{1}{2}(y\Delta x)^2 + ...)}{\Delta x} = ye^{xy}$$

Again, this is the same result we would get if we assumed that $y =$ constant and take the normal derivative with respect to x .