Section 2
The Derivative

One of the most important areas in undergraduate mathematics is that of calculus. In this section we'll introduce the derivative of a real function, the differential operator, and how to apply this operator to some simple functions.

What in the derivative?

Consider the following construction:

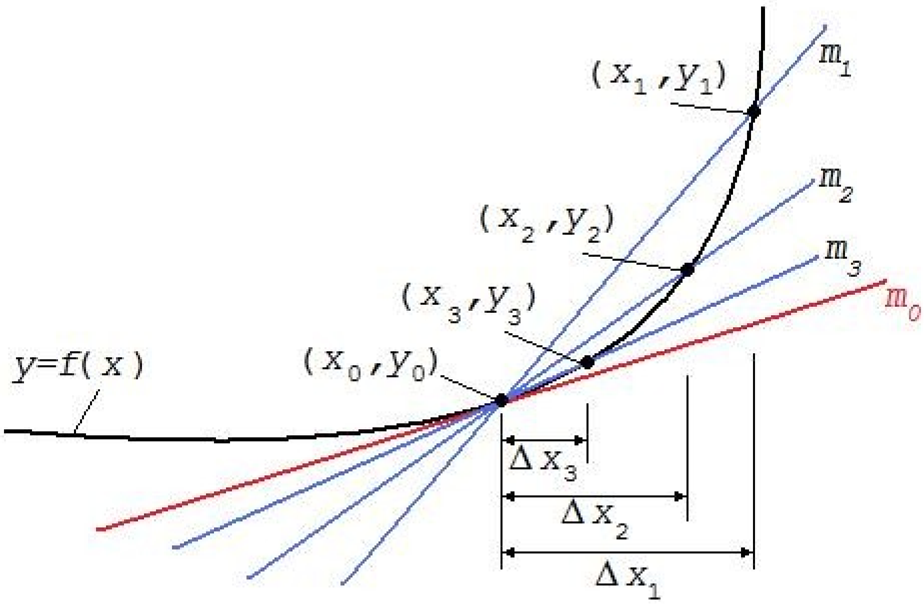

Figure: limiting tangent to the curve $y=f(x)$

Let $f:U\rightarrow V$ be continuous at some $x_0\in U$ and let $y=f(x)$ be the graph of $f(x)$ on U. Let $(x_1, y_1), (x_2, y_2)$ and (x_3, y_3) be three neighbouring points such that the line segments joining them to (x_0, y_0) have slopes $m_1, m_2,$ and m_3 respectively. Then

$$m_1=\frac{y_1-y_0}{x_1-x_0}=\frac{f(x_1)-f(x_0)}{x_1-x_0}=\frac{f(x_0+\Delta x_1)-f(x_0)}{\Delta x_1}$$
$$m_2=\frac{y_2-y_0}{x_2-x_0}=\frac{f(x_2)-f(x_0)}{x_2-x_0}=\frac{f(x_0+\Delta x_2)-f(x_0)}{\Delta x_2}$$
$$m_3=\frac{y_3-y_0}{x_3-x_0}=\frac{f(x_3)-f(x_0)}{x_3-x_0}=\frac{f(x_0+\Delta x_3)-f(x_0)}{\Delta x_3}$$

We define the differences in slopes between the above three and the red slope m_0 by

$$\Delta m_1 = m_1 - m_0, \Delta m_2 = m_2 - m_0, \Delta m_3 = m_3 - m_0$$

and it is apparent from the above construction that

$$\Delta m_1 > \Delta m_2 > \Delta m_3$$

and as x approaches x_0 this slope difference, defined by

$$\Delta m = m - m_0 \text{ where } m = \frac{f(x) - f(x_0)}{x - x_0},$$

is tending to zero: i.e.

$$\lim_{x \to x_0} \Delta m = 0$$

In other words, as x approaches x_0 the slope of the line segment joining (x, y) to (x_0, y_0) is tending to m_0, the slope of the tangent to the curve $y = f(x)$ at the point (x_0, y_0).

Then

$$\lim_{x \to x_0} \Delta m = 0$$

$$\Rightarrow \lim_{x \to x_0} m = m_0$$

$$\Rightarrow \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = m_0$$

$$\Rightarrow \lim_{x \to x_0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = m_0$$

So after all this, we have a method of calculating the slope of the tangent to the curve $y = f(x)$ at the point (x_0, y_0) subject to $f(x)$ being continuous and smooth.

N.B.
By smooth we mean a continuous function at a point (x_0, y_0)

1. has a unique tangent slope m_0
2. the slope of the line segment between (x_0, y_0) and some arbitrarily close point (x, y) tends to m_0 as x approaches x_0 from either direction.

What does all this have to do with the derivative?

Definition

Let $f: U \to V$ be continuous at some $x_0 \in U$. Then we define the derivative of $f(x)$ at $x = x_0$ to be

$$\left[\frac{d}{dx} f(x)\right]_{x_0} \underset{\text{DEF}^{\text{N}}}{\overset{\uparrow}{:=}} \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

If $\Delta x = x - x_0$ then we have

$$\left[\frac{d}{dx} f(x)\right]_{x_0} \underset{\text{DEF}^{\text{N}}}{\overset{\uparrow}{:=}} \lim_{x \to x_0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

In practice, we often consider the derivative over an interval of values and so we drop the x_0 dependence:

$$\frac{d}{dx} f(x) \underset{\text{DEF}^{\text{N}}}{\overset{\uparrow}{:=}} \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

where the limit has been changed to the equivalent but more appropriate $\Delta x \to 0$.

N.B.

1. The differential operator is defined $\dfrac{d}{dx}$ and is not a quotient. It is read "*d dx*". Thus
$$\frac{d}{dx} f(x)$$
is read "*d dx*" of "*f* of *x*".
2. The operation of this differential operator on a function $f(x)$ is more properly written as that above; i.e.
$$\frac{d}{dx} f(x)$$
It can be written as
$$\frac{d f(x)}{dx}$$
but care must be taken to remember that it is not a quotient; i.e. $d\,f(x)$ is not being divided by $d\,x$.
3. A function that has a derivative at some $x_0 \in U$ is deemed differentiable at $x_0 \in U$. A differentiable function is smooth and a smooth function is differentiable.

How do we perform this operation for a given function?

To answer this we'll consider some functions.

Example 1: $f(x)=ax+b \quad \forall a,b\in\mathbb{R};\ a\neq 0$
Then

$$f(x+\Delta x)=a(x+\Delta x)+b$$

and

$$\frac{d}{dx}f(x)=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{a(x+\Delta x)+b-(ax+b)}{\Delta x}$$

Simplifying the above we get

$$\begin{aligned}\frac{d}{dx}f(x)=\frac{d}{dx}(ax+b) &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{a(x+\Delta x)+b-(ax+b)}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{a\Delta x}{\Delta x}\\ &= \lim_{\Delta x\to 0} a\end{aligned}$$

Then, as *a* is not dependent of Δx

$$\frac{d}{dx}f(x)=\frac{d}{dx}(ax+b)=a$$

What is *a* with respect to a linear function?

Example 2: $f(x)=ax^2 \quad \forall a\in\mathbb{R}\setminus\{0\}$
Then

$$f(x+\Delta x)=a(x+\Delta x)^2$$

and

$$\frac{d}{dx}f(x)=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{a(x+\Delta x)^2-ax^2}{\Delta x}$$

Simplifying[1] the above we get

$$\begin{aligned}\frac{d}{dx}f(x)=\frac{d}{dx}(ax^2) &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{a(x+\Delta x)^2-ax^2}{\Delta x}\\ &= \lim_{\Delta x\to 0} a\left(\frac{x^2+2x\Delta x+(\Delta x)^2-x^2}{\Delta x}\right)\end{aligned}$$

1 Using the identity $(x+y)^2=x^2+2xy+y^2$

$$
\begin{aligned}
\frac{d}{dx} f(x) = \frac{d}{dx}\left(a x^2\right) \quad &= \quad \lim_{\Delta x \to 0} a\left(\frac{2x\Delta x + (\Delta x)^2}{\Delta x}\right) \\
&= \quad \lim_{\Delta x \to 0} a\left(2x + \Delta x\right) \\
&= \quad 2ax
\end{aligned}
$$

Therefore

$$\frac{d}{dx} f(x) = \frac{d}{dx} a x^2 = 2ax$$

N.B.

In the last example we expanded $(x+y)^2 = x^2 + 2xy + y^2$ in order to simplify the numerator terms. For higher powers of $(x+y)$ this calculation becomes unwieldy and so we require other means to determine the terms in the expansion: i.e. we want to substitute calculations such as

$$(x+y)^n = \underbrace{(x+y)\times(x+y)\times\ldots\times(x+y)}_{n\text{-times}}$$

with more manageable methods.

For relatively small numbers, n < 10, we can use Pascal's Triangle. For larger n we resort to the bionomial expansion itself.

Pascal's Triangle

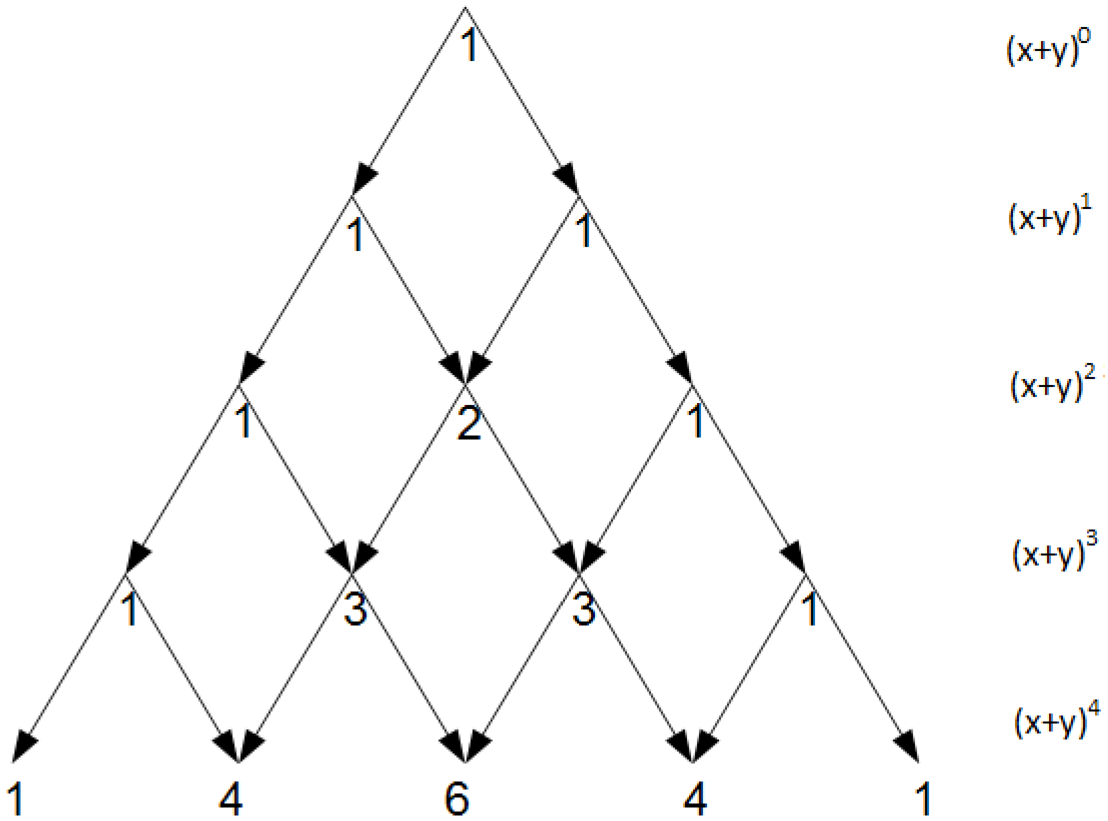

Figure: Pascal's Triangle up to n = 4.

The triangle works as follows:

- The apex is the coefficient for the expansion $(x+y)^0=1$
- The next layer is calculated as follows:
  1. The extreme left and right nodes are labelled 1.
  2. Any interior node at this level is calculated from the sum of all nodes on the previous level that are directed to it.
  3. This is repeated until all nodes at this level are labelled.
- Move onto the next level until the *n* you require is reached.
- The (*n* + 1) labels are read left to right as the coefficients of the terms $x^n, x^{n-1}y, \ldots, xy^{n-1}, y^n$ respectively.

Therefore we have, from the triangle, for $(x+y)^3$ the labels 1, 3, 3, 1. The first is the coefficient for x^3, the second for x^2y, the third for xy^2, and the last for y^3 : i.e

$$\underbrace{1,3,3,1}_{\text{Pascal's}\,\Delta} \equiv 1x^3+3x^2y+3xy^2+1y^3$$

Exercise:

Expand Pascal's Triangle for $(x+y)^6$ and verify the coefficients obtained are correct by explicitly calculating $(x+y)^6$.

Binomial Expansion:

This is a more complete method and the preferred approach by mathematicians when determining the coefficients in expansions of the form $(x+y)^n \ \forall n \in \mathbb{N}$.

The full explanation of all the terms will be attempted at a later stage. For now we'll just state the formula:

$$(x+y)^n=\sum_{k=0}^{n} {}^nC_k\, x^{n-k} y^k$$

where the coefficients are determined from the ${}^nC_k=\dfrac{n!}{k!(n-k)!}$ terms, called combinatorial terms, and

$$k!=k\times(k-1)\times\ldots\times 2\times 1 \quad \forall k\in\mathbb{N}$$

is the factorial of *k* or *k* factorial.

Exercise:

Use either method above to determine $\dfrac{d}{dx}f(x)$ if $f(x)=x^3$