Differentiation From First Principles

> "First principles, Clarice. Simplicity. Read Marcus Aurelius. Of each particular thing ask: what is it in itself? What is its nature? What does he do, this man you seek?" — Hannibal Lecter, The Silence Of The Lambs

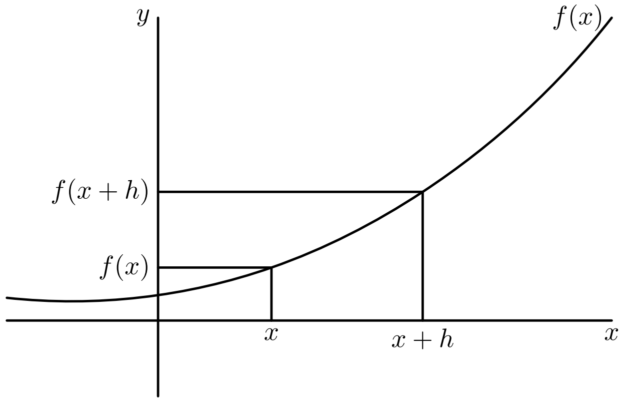
After studying differentiation for the first time we know the following:

$$\text{Differential of } y = x^2 \quad \Rightarrow \quad \frac{dy}{dx} = 2x.$$

$$\text{Differential of } y = x^3 \quad \Rightarrow \quad \frac{dy}{dx} = 3x^2.$$

$$\text{Gradient of } y = x^3 \text{ when } x = 4 \quad \Rightarrow \quad \left.\frac{dy}{dx}\right|_{x=4} = 3 \times 4^2 = 48.$$

We will derive these results *from first principles*. Consider the following graph of a function $y = f(x)$. ($y = f(x)$ could be *any* function. For example $y = x^2$, $y = x^3$, $y = x^n$, $y = \sin x$, $y = e^x \dots$)



Consider the graph at x and $x + h$; the y values that these points take respectively are $f(x)$ and $f(x + h)$. Now let us consider the gradient of the line joining the two points $(x, f(x))$ and $(x+h, f(x+h))$. From our previous work on coordinate geometry we know that the gradient is

$$\boxed{\text{Gradient} = \frac{\text{difference in } y}{\text{difference in } x} = \frac{f(x+h) - f(x)}{x + h - x} = \frac{f(x+h) - f(x)}{h}.}$$

From our original graph of $f(x)$ we can see that as we make h smaller we get an increasingly accurate measure of the gradient of the curve at x. Indeed if we allow h to equal 0, then the measure of the gradient *should* be perfect. However, if we glance at our expression for the gradient we can see that we *cannot* let $h = 0$. So we have to do the next best thing and let h to tend to zero ($h \to 0$).

Example of $y = x^2$

$$\text{Gradient} = \frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = \frac{x^2 + 2xh + h^2 - x^2}{h} = \frac{2xh + h^2}{h} = 2x + h.$$

As $h \to 0$ we can see that the gradient becomes $2x$, as required.

Example of $y = x^3$

$$\text{Gradient} = \frac{f(x+h) - f(x)}{h} = \frac{(x+h)^3 - x^3}{h} = \frac{x^3 + 3x^2h + 3xh^2 + h^3 - x^3}{h}$$

$$= \frac{3x^2h + 3xh^2 + h^3}{h} = 3x^2 + 3xh + h^2.$$

As $h \to 0$ we can see that the gradient becomes $3x^2$, as required.

Example of $y = x^3$ when $x = 4$

$$\text{Gradient} = \frac{f(4+h) - f(4)}{h} = \frac{(4+h)^3 - 4^3}{h} = \frac{4^3 + 3 \times 4^2 h + 3 \times 4h^2 + h^3 - 4^3}{h}$$
$$= \frac{3 \times 4^2 h + 3 \times 4h^2 + h^3}{h} = 3 \times 4^2 + 3 \times 4h + h^2.$$

As $h \to 0$ we can see that the gradient becomes $3 \times 4^2 = 48$, as required.

The General Case of $y = x^n$ (for integer n)

$$\begin{aligned}
\text{Gradient} &= \frac{f(x+h) - f(x)}{h} \\
&= \frac{(x+h)^n - x^n}{h} \\
&= \frac{\binom{n}{0}x^n + \binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \binom{n}{3}x^{n-3}h^3 + \cdots + \binom{n}{n-1}xh^{n-1} + \binom{n}{n}h^n - x^n}{h} \\
&= \frac{x^n + nx^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \binom{n}{3}x^{n-3}h^3 + \cdots + nxh^{n-1} + h^n - x^n}{h} \\
&= \frac{nx^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \binom{n}{3}x^{n-3}h^3 + \cdots + nxh^{n-1} + h^n}{h} \\
&= nx^{n-1} + \binom{n}{2}x^{n-2}h + \binom{n}{3}x^{n-3}h^2 + \cdots + nxh^{n-2} + h^{n-1}.
\end{aligned}$$

As $h \to 0$ we can see that the gradient becomes nx^{n-1}, as required.