A special case: if $g(x) = c$, constant, then $g'(x) = 0$, so

$$\frac{d}{dx}\left(cf(x)\right) = c\frac{d}{dx}\left(f(x)\right) = cf'(x).$$

Therefore the sum rule can be generalised as: If $f_1, f_2, \ldots, f_n$ are differentiable and $a_1, a_2, \ldots, a_n$ are constants, then

$$\frac{d}{dx}\left[a_1 f_1(x) + a_2 f_2(x) + \cdots + a_n f_n(x)\right] = a_1 f_1'(x) + a_2 f_2'(x) + \cdots + a_n f_n'(x)$$

Example 2.10. Consider a polynomial of degree n with constant coefficients, i.e.

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n,$$

then

$$p'(x) = a_1 + 2a_2 x + 3a_3 x^2 \cdots + n a_n x^{n-1},$$

a polynomial of degree $n - 1$, with constant coefficients.

The chain rule:

The chain rule tells us how to differentiate "compositions" of functions. If we have two functions f and g, the composition, denoted by $f \circ g$ (name of new function), is the function given by

$$f \circ g(x) = f(g(x)), \qquad \text{(Do } g \text{ then } f\text{).} \tag{2.6}$$

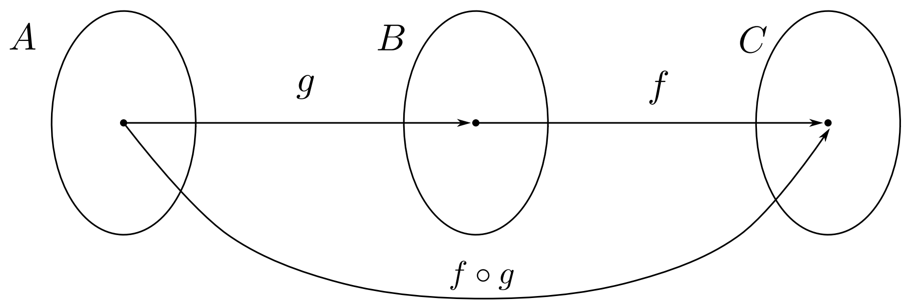

Figure 2.7: *The composition $f \circ g$ first employs g from A to B, then f from B to C.*

Example 2.11. If $f(w) = w^2 + 1$ and $g(u) = \sqrt{u}$, then

$$f \circ g(x) = f(g(x)) = f(\sqrt{x}) = (\sqrt{x})^2 + 1 = x + 1,$$

$$g \circ f(x) = g(f(x)) = g(x^2 + 1) = \sqrt{x^2 + 1}.$$

Here, $f : \mathbb{R} \to \mathbb{R}+$, $g : \mathbb{R}^+ \to \mathbb{R}^+$ and $f \circ g : \mathbb{R}^+ \to \mathbb{R}^+$, $g \circ f : \mathbb{R} \to \mathbb{R}^+$.

But actually the function $h(x) = x + 1$ can be defined as $h : \mathbb{R} \to \mathbb{R}$, so be careful when generating a function by composition, take note of the difference between $h(x)$ and $f \circ g$ in this case.

In general, $f \circ g \neq g \circ f$.

Composition can be generalised further for more functions, for example suppose we have three functions f, g and h, then

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

If f and g are differentiable, then

$$\frac{d}{dx}\left(f(g(x))\right) = f'(g(x)) \cdot g'(x). \tag{2.7}$$

Example 2.12. Consider the function $y(x) = (x^3+2x)^{10}$. Here we will choose $f(w) = w^{10}$ and $g(x) = x^3 + 2x$, (so $f'(w) = 10w^9$ and $g'(x) = 3x^2 + 2$). Then

$$\begin{aligned} \frac{d}{dx}\left(y(x)\right) &= \frac{d}{dx}\left(f(g(x))\right) \\ &= \frac{d}{dx}\left((x^3+2x)^{10}\right) \\ &= f'(g(x))g'(x) \\ &= 10(x^3+2x)^9 \cdot (3x^2+2). \end{aligned}$$

Essentially, what we have done is to substitute $g(x) = x^3 + 2x$ in our function for $y(x)$, to make the differentiation easier.

Example 2.13. Consider the function $y(x) = 1/x^3$. We know how to differentiate $1/x$. So let us choose $g(x) = x^3$ and $f(w) = 1/w$. Therefore we have $f \circ g(x) = y(x)$. We know the derivatives of f and g are $f'(w) = -1/w^2$ and $g'(x) = 3x^2$. So

$$\begin{aligned} \frac{d}{dx}\left(\frac{1}{x^3}\right) &= f'(g(x))g'(x) \\ &= -\frac{1}{(x^3)^2} \cdot 3x^2 \\ &= -\frac{3}{x^4}, \end{aligned}$$

i.e. we have

$$\frac{d}{dx}(x^{-3}) = -3x^{-4}.$$

RECALL: we have seen that if n is a positive integer, then

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

e.g.

$$\frac{d}{dx}(x^{365}) = 365x^{364}.$$

If n is a negative integer, then $m = -n$ is a positive integer. So

$$\begin{aligned}\frac{d}{dx}(x^n) &= \frac{d}{dx}\left(\frac{1}{x^m}\right)\\ &= \frac{d}{dx}(f(g(x)))\\ &= f'(g(x))g'(x)\\ &= -\frac{1}{(x^m)^2}\cdot mx^{m-1}\\ &= -mx^{m-1-2n}\\ &= (-m)x^{(-m)-1}\\ &= nx^{n-1}.\end{aligned}$$

Here we have simply chosen $g(x) = x^m$ and $f(w) = 1/w$, where $g'(x) = mx^{m-1}$ and $f'(w) = -1/w^2$.

Therefore we now know that
$$\frac{d}{dx}(x^n) = nx^{n-1},$$
is true for any whole number $n \in \mathbb{Z}$.

What about when $n = 1/2$ i.e. $f(x) = x^{\frac{1}{2}}$. How can we differentiate this? First let us think about what we know about $x^{\frac{1}{2}}$.

We know that $(x^{\frac{1}{2}})^2 = x$!

Let us consider the function $g(y) = y^2$ and take the composition of f and g, that is

$$g \circ f(x) = g(f(x)) = (x^{\frac{1}{2}})^2 = x.$$

In this case f and g are inverse of one another. What does it mean for g to be the inverse of f or f to be the inverse of g?

Aside (NFE):

If we take a point $x = a$ in the domain of f say, and it takes the value $b = f(a)$ in the range (or *image*). Then the inverse function takes the image point b and sends it back to the point $x = a$. In other words we return ourselves back to where we started. There is a well defined rule that goes from a to b and a well defined rule that takes b to a.

The inverse is usually written as f^{-1}, this is just notation.

$$f^{-1}(f(x)) = x \quad \text{then} \quad \begin{cases} f : A \to B \\ f^{-1} : B \to A \end{cases}$$

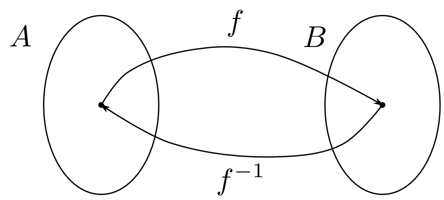

Figure 2.8: *The composition $f \circ f^{-1} = f^{-1} \circ f$ takes you back to where you started!*

Definition 2.2. *The function f^{-1} is called the inverse function for a well defined function f then*

$$f \circ f^{-1}(x) = f^{-1} \circ f(x) = x. \tag{2.8}$$

Not all functions possess inverses. For example the function $f(x) = c$, constant.

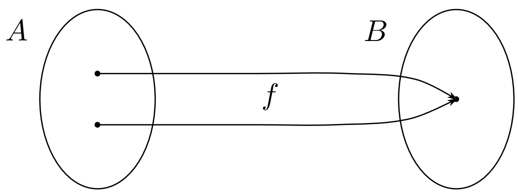

Figure 2.9: *Multi-valued functions do not have inverses (obvious from picture).*

A function is called a *one-to-one* function if it never takes the same value twice, that is

$$f(x_1) \neq f(x_2) \quad \text{whenever} \quad x_1 \neq x_2.$$

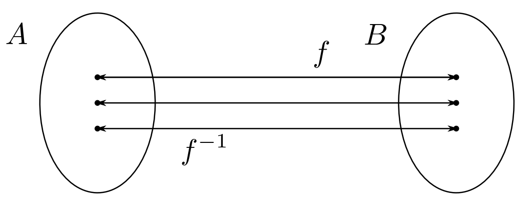

Figure 2.10: *One-to-one functions have inverses (obvious from picture).*

Only for a one-to-one function f, then f^{-1} exists.

Example 2.14. $f(x) = x^2$, if $f : (-\infty, \infty) \to [0, \infty)$ then f^{-1} does not exist since for each $f(x)$, there are two possible x values corresponding to it.

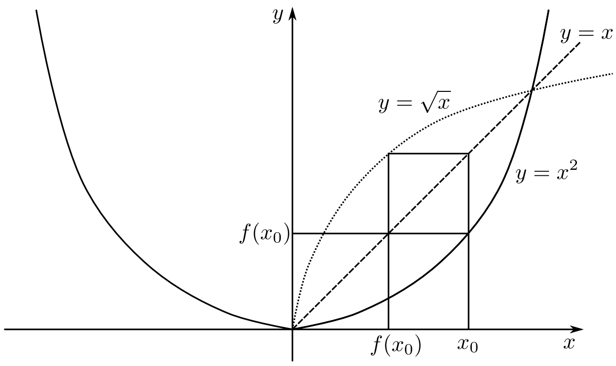

Figure 2.11: *Graph showing the relationship between $y = x^2$ and it's inverse.*

(End NFE)

If it was that $f : [0, \infty) \to [0, \infty)$ i.e. considering the positive x-axis only, then f^{-1} exists. Which in this case we call $f^{-1}(x) = g(x) = \sqrt{x}$, since

$$f(g(x)) = f(\sqrt{x}) = (\sqrt{x})^2 = x.$$

And we also know that

$$g(f(x)) = g(x^2) = \sqrt{x^2} = x,$$

i.e. f and g are inverse of one another.

If you draw a function and its inverse on the same coordinate plane, they must be symmetrical about the line $y = x$. Why? Rotate the xy-plane 90° anticlockwise and then flip across the vertical axis. This is because we want the inverse function f^{-1} who's range is the domain of f and vice-versa. Also, equivalent to switch $x \leftrightarrow y$ in $y = f(x)$, rearranging the equation for y to give the inverse.

Now let us return to take the derivative of the function $g(f(x)) = (x^{\frac{1}{2}})^2 = x$, thus

$$\begin{aligned} \frac{d}{dx}\left[g(f(x))\right] &= \frac{d}{dx}(x) \\ \therefore \quad g'(f(x))f'(x) &= 1. \end{aligned}$$

Now since $g'(y) = 2y$, we have

$$2 \cdot f(x)f'(x) = 1.$$

Finally rearranging for $f'(x)$ we see that

$$f'(x) = \frac{1}{2f(x)} = \frac{1}{2x^{\frac{1}{2}}} = \frac{1}{2}x^{-\frac{1}{2}},$$

i.e. the method is the same for integer n (as shown on pg. 23), in words, “bring the power down, reduce the power by one”.[3]

[3]End Lecture 7.