Derivation of the Chain Rule

Suppose $y = f \circ g(x)$. Assuming f and g have derivatives where appropriate, the Chain Rule says that $(f \circ g)' = (f' \circ g) \cdot g'$. In more practical language, if we write $y = f(u)$ and $u = g(x)$, it comes out as $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$.

To see why, go back to the definition of a derivative and write

$\frac{dy}{dx} = (f \circ g)'(x) = \lim_{h\to 0} \frac{f \circ g(x+h) - f \circ g(x)}{h} = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h}$.

We'd like to write this as

$\lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \cdot \frac{g(x+h) - g(x)}{h}$, but it's possible that $g(x+h) = g(x)$, and hence the denominator $g(x+h) - g(x) = 0$, for some $h \neq 0$, so we consider two separate cases.

If there are arbitrarily small values of h for which $g(x+h) = g(x)$, then both $(f \circ g)'(x)$ and $g'(x)$ will have to equal 0 and the Chain Rule certainly holds in a trivial fashion.

So we need only verify the Chain Rule in the remaining case where $g(x+h) \neq g(x)$ is h is close to 0. For this case, we write

$u = g(x)$, $g(x+h) = g(x) + k$, $k = g(x+h) - g(x)$.

We can then write

$\frac{dy}{dx} = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \cdot \frac{g(x+h) - g(x)}{h} = \lim_{h\to 0} \frac{f(u+k) - f(u)}{k} \cdot \frac{g(x+h) - g(x)}{h}$.

Since the limit of a product is equal to the product of limits, we have

$\frac{dy}{dx} = \lim_{h\to 0} \frac{f(u+k) - f(u)}{k} \cdot \lim_{h\to 0} \frac{g(x+h) - g(x)}{h}$.

Since $k \approx 0$ when $h \approx 0$, the limit $\lim_{h\to 0} \frac{f(u+k) - f(u)}{k}$ will equal the limit $\lim_{k\to 0} \frac{f(u+k) - f(u)}{k}$, which is equal to $f'(u)$ or $\frac{dy}{du}$, while $\lim_{h\to 0} \frac{g(x+h) - g(x)}{h}$ is, by definition, equal to $g'(x) = \frac{du}{dx}$.

This demonstrates that $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$.