Chain rule: proof

Given two functions f and g we are interested in the composite function $f(g(x))$, often written as $f \circ g$, and in particular, we wish to derive an expression for the derivative

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

Naturally, we insist that g is differentiable at x and f is differentiable at $g(x)$.

The chain rule is the formula:

$$\frac{d}{dx}(f \circ g) = f'(g(x)) \cdot g'(x)$$

which is more easily remembered as

$$\frac{df}{dx} = \frac{df}{dg}\,\frac{dg}{dx}$$

As an example, if we write

$$y = g(x) = x^2$$

$$z = f(y) = y^2$$

$$z' = 2y \cdot 2x = 2(x^2)2x = 4x^3$$

which is easily checked by recognizing that $z = x^4$.

proof

The proof is a little tedious, but here goes. Again, we want to compute:

$$\lim_{h \to 0} \frac{f(g(x+h)) - f(g(x))}{h}$$

Since $g(x)$ is differentiable at x

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

Rearrange and define a new variable v

$$v = \frac{g(x+h) - g(x)}{h} - g'(x)$$

v depends on h and $v \to 0$ as $h \to 0$. We can set up a similar expression involving f, namely:

$$w = \frac{f(y+k) - f(y)}{k} - f'(y)$$

w depends on k and $w \to 0$ as $k \to 0$.

Rearrange some more:

$$g(x+h) = g(x) + [\ g'(x) + v\ ]\ h$$

$$f(y+k) = f(y) + [\ f'(y) + w\ ]\ k$$

So now we want to rewrite $f(g(x+h))$. Using the first equation

$$f(g(x+h)) = f(g(x) + h\ [\ g'(x) + v\ ]\ )$$

Now, pick a particular $k = [\ g'(x) + v\ ]\ h$, So that is (using the second equation and $g(y) = g(x+h)$):

$$f(g(x+h)) = f(g(x)) + [\ f'(g(x)) + w\ ]\ k$$

and substituting for k:

$$f(g(x+h)) =$$
$$= f(g(x)) + \ [\ f'(g(x)) + w\ ]\ \ [\ g'(x) + v\ ]\ h$$

Go back to the difference quotient:

$$\frac{f(g(x+h)) - f(g(x))}{h}$$

Now that we have extracted $f(g(x))$ from the first term we can put everything together and see some cancellations:

$$\frac{f(g(x+h)) - f(g(x))}{h} =$$
$$= \frac{f(g(x)) + \ [\ f'(g(x)) + w\ ]\ \ [\ g'(x) + v\ ]\ h\ ]\ \ - f(g(x))}{h}$$

Cancel the first and last term in the numerator

$$= \frac{[\ f'(g(x)) + w\ ]\ \ [\ g'(x) + v\ ]\ h\ ]}{h}$$

Cancel the h

$$= [\ f'(g(x)) + w\ ]\ \ [\ g'(x) + v\ ]$$

So now, we just need to put in the limit:

$$= \lim_{h \to 0} [\ f'(g(x)) + w\ ]\ \ [\ g'(x) + v\ ]$$

$$= \ [\ \lim_{h \to 0} f'(g(x)) + \lim_{h \to 0} w\ ]\ [\ \lim_{h \to 0} g'(x) + \lim_{h \to 0} v\ ]$$

But, as $h \to 0$, so $k \to 0$, and as $k \to 0$, so $v \to 0$ and $w \to 0$, and we just have:

$$= \ [\ \lim_{h \to 0} f'(g(x))]\ [\ \lim_{h \to 0} g'(x)\ ]$$

which is the chain rule.

According to my source, the often-seen proof involves multiplying by the inverse of $g'(x)$:

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

$$(f \circ g)'(x)\ (\frac{1}{g'(x)} = \lim_{h \to 0} \frac{f(g(x+h)) - f(g(x))}{h}\ \frac{h}{g(x+h) - g(x)}$$

whereupon we cancel the solitary h

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

$$= f'(g(x))$$

Hence

$$(f \circ g)'(x) = f'(g(x))g'(x)$$

But this proof is "technically incorrect".