Proof of the Chain Rule

• Given two functions f and g where g is differentiable at the point x and f is differentiable at the point $g(x) = y$, we want to compute the derivative of the composite function $f(g(x))$ at the point x. In other words, we want to compute

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

The answer involves the derivative of the *outer* function f and the derivative of the *inner* function g. Specifically,

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

• To prove this rule we first set up some notation. We are assuming that the function $g(x)$ is differentiable at the point x. This means that the number $g'(x)$ exists and is equal to our limit definition of the derivative, and so

$$\frac{g(x+h) - g(x)}{h} - g'(x) \to 0 \text{ as } h \to 0.$$

We define a new variable v by

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

Notice that v depends on the number h and that $v \to 0$ as $h \to 0$. Similarly, because we are assuming the function f is differentiable at the point $y = g(x)$ we have

$$\frac{f(y+k) - f(y)}{k} - f'(y) \to 0 \text{ as } k \to 0,$$

where here we are using the variable k to stand for the h we normally use. Again, we define another variable w by

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

Again, notice that w depends on k and that $w \to 0$ as $k \to 0$.

- From our definitions of v and w it follows that

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

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

We now use these equations to rewrite $f(g(x+h))$. In particular, use the first equation to obtain

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

and use the second equation applied to the right-hand-side with $k = [g'(x) + v]h$ and $y = g(x)$. Note that using this quantity for k tells us that $k \to 0$ as $h \to 0$, and so $w \to 0$ as $h \to 0$. Applying this equation gives

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

Using this last equation we now simplify the fraction

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

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

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

• We are now ready to compute the derivative:

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

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

$$= f'(g(x)) \cdot g'(x),$$

since $v \to 0$ as $h \to 0$ and $w \to 0$ as $h \to 0$.□

• **The Flawed Proof** Many calculus texts present an incorrect proof of the chain rule that goes as follows:

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

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

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

Therefore $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$.

For 3 bonus points, explain why this proof is technically incorrect.