Theorem 1 (Chain Rule)**.** *Let f and g be two differentiable functions (and hence, you can find the derivatives of f and g respectively at any point). Suppose f is differentiable at a point x, and suppose that g is differentiable at $f(x)$, then the* ***chain rule*** *gives*

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

Proof. We will use the definition of the derivative to show that

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

First, since f is differentiable at x, then by the definition of the derivative

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

Then rewrite the equation

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

However, we are only interested in the expression $\frac{f(x+h) - f(x)}{h} - f'(x)$, so we will define a function $i(h)$ as

$$i(h) = \begin{cases} \dfrac{f(x+h) - f(x)}{h} - f'(x) & \text{if } h \neq 0 \\ 0 & \text{if } h = 0 \end{cases}$$

Then note that when $h = 0$, i is continuous at 0. Otherwise, for $h \neq 0$, we have

$$i(h) = \frac{f(x+h) - f(x)}{h} - f'(x)$$

Then by rewriting the equation so we solve for $f(x+h)$,

$$f(x+h) = h[i(h) + f'(x)] + f(x) \tag{1}$$

Similarly, since g is differentiable at $f(x)$, then by the definition of the derivative,

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

Then rewrite

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

As we are only interested in the expression $\frac{g(f(x) + k) - g(f(x))}{k} - g'(f(x))$, we will define a function $j(k)$ as

$$j(k) = \begin{cases} \dfrac{g(f(x) + k) - g(f(x))}{k} - g'(f(x)) & \text{if } k \neq 0 \\ 0 & \text{if } k = 0 \end{cases}$$

Then note that when $k = 0$, j is continuous at 0. Otherwise, for $k \neq 0$,

$$j(k) = \frac{g(f(x) + k) - g(f(x))}{k} - g'(f(x))$$

Then by rewriting the equation so we solve for $g(f(x)+k)$,

$$g(f(x)+k) = k[j(k) + g'(f(x))] + g(f(x)) \tag{2}$$

Now that we have the two required pieces of information to prove the chain rule, we have by the definition of the derivative,

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

Then working with the numerator, since we have an expression for $f(x+h)$ from (1),

$$g(f(x+h)) - g(f(x)) = g(h[i(h) + f'(x)] + f(x)) - g(f(x))$$

And since from (2), if we let $k = h[i(h) + f'(x)]$, we have that

$$\begin{aligned} g(f(x) + h[i(h) + f'(x)]) - g(f(x)) &= h[i(h) + f'(x)][j(h[i(h) + f'(x)]) + g'(f(x))] + g(f(x)) - g(f(x)) \\ &= h[i(h) + f'(x)][j(h[i(h) + f'(x)]) + g'(f(x))] \end{aligned}$$

Then substituting back to the definition of the derivative above,

$$\begin{aligned} (g \circ f)'(x) &= \lim_{h \to 0} = \frac{h[i(h) + f'(x)][j(h[i(h) + f'(x)]) + g'(f(x))]}{h} \\ &= \lim_{h \to 0} [i(h) + f'(x)][j(h[i(h) + f'(x)]) + g'(f(x))] \\ &= [0 + f'(x)][0 + g'(f(x))] \\ &= f'(x)g'(f(x)) \\ &= g'(f(x))f'(x) \end{aligned}$$

which is exactly what we wanted to show. Therefore, the chain rule is proven. $\square$