Chain Rule

The product rule tells us how to find the derivative of a product of functions like $f(x) \cdot g(x)$. The composition or "chain" rule tells us how to find the derivative of a composition of functions like $f(g(x))$. Composition of functions is about substitution – you substitute a value for x into the formula for g, then you substitute the result into the formula for f. An example of a composition of two functions is $y = (\sin t)^{10}$ (which is usually written as $y = \sin^{10} t$).

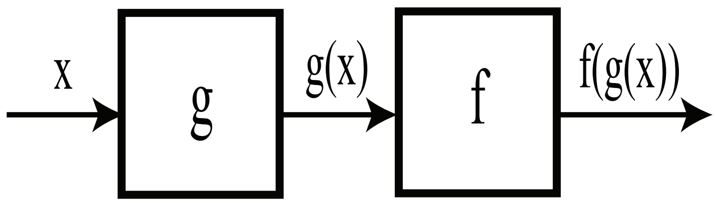

Figure 1: Composition of functions: $(f \circ g)(x) = f(g(x))$

One way to think about composition of functions is to use new variable names. For example, for the function $y = \sin^{10} t$ we can say $x = \sin t$ and then $y = x^{10}$. Notice that if you plug $x = \sin t$ in to the formula for y you get back to $y = \sin^{10} t$. It's good practice to introduce new variables when they're convenient, and this is one place where it's very convenient.

So, how do we find the derivative of a composition of functions? We're trying to find the slope of a tangent line; to do this we take a limit of slopes $\frac{\Delta y}{\Delta t}$ of secant lines. Here y is a function of x, x is a function of t, and we want to know how y changes with respect to the original variable t. Here again using that intermediate variable x is a big help:

$$\frac{\Delta y}{\Delta t} = \frac{\Delta y}{\Delta x} \cdot \frac{\Delta x}{\Delta t}$$

because when we perform the multiplication, the small change Δx cancels.

The derivative of y with respect to t is $\lim_{\Delta t \to 0} \dfrac{\Delta y}{\Delta t}$; what happens when Δt gets small? Because $x = \sin t$ is a continuous function, as Δt approaches 0, Δx also approaches zero. It turns out that:

$$\lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} = \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} \qquad \longleftarrow \textbf{The Chain Rule!}$$

The derivative of a composition of functions is a product. In the example $y = (\sin t)^{10}$, we have the "inside function" $x = \sin t$ and the "outside function" $y = x^{10}$. The chain rule tells us to take the derivative of y with respect to x and multiply it by the derivative of x with respect to t.

The derivative of $y = x^{10}$ is $\dfrac{dy}{dx} = 10x^9$. The derivative of $x = \sin t$ is $\dfrac{dx}{dt} = \cos t$. The chain rule tells us that $\frac{d}{dt} \sin^{10} t = 10x^9 \cdot \cos t$. This is correct, but if a friend asked you for the derivative of $\sin^{10} t$ and you answered $10x^9 \cdot \cos t$ your friend wouldn't know what x stood for. The last step in this process is to rewrite x in terms of t:

$$\frac{d}{dt}\sin^{10} t = 10(\sin t)^9 \cdot \cos t = 10\sin^9 t \cdot \cos t.$$

Here is another way of writing the chain rule:

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