The Chain Rule

Three brothers, Kevin, Mark, and Brian like to hold an annual race to start off each new year. In the race the three brothers like to compete to see who is the fastest, and who will come in last, and have to buy the others breadsticks (these are three crazy brothers!). Last year Mark won the race, and he is feeling very confident in winning once again. The brothers have set up a camera to watch the race track, so there can be no dispute as to who won. However, since these three brothers never studied calculus, they do not know how to find their velocity from their position. Instead, they use the previous year's winner as a pace setter, and measure the velocity of the brothers relatively. Since Mark won last year's competition, he will be the pace setter.

Let K, M, B be functions that represent the position of each of the brothers at any given time. Since the brother's don't know how to calculuate velocity, Mark's velocity dM/dt is an unknown function. Nevertheless, the brother's aren't interested in absolute velocity, but only knowing who will come in first. If Brian is running 2/3 as fast as Mark, and Kevin runs twice as fast as Brian, how fast is Kevin running compared to Mark? If the brothers maintain these paces, who will win the race?

First, let's consider the dependence of these functions. Since Mark is the pace setter, his position only depends on time, so we write $M(t)$. Now Brian's velocity (and thus position) depend on Mark's, so we write $B(M(t))$. Finally, Kevin's position depends on Brian's (and thus indirectly on Mark's) so we write $K(B(M(t)))$. Now since Brian is running 2/3 as fast as Mark, the rate of change of Brian's position with respect to Mark's position should be

$$\frac{dB}{dM} = \frac{2}{3}$$

meaning that Brian's position changes 2/3 as fast as Mark's position. If we knew Mark's velocity (or instantaneous rate of change of position with time) we should be able to find Brian's velocity, just by multiplying by 2/3, the rate at which Brian's position is changing relative to Mark's. If Mark is running at 10mph, it follows Brian is running at $20/3 \approx 6.67$ mph. Written symbolically, we observe

$$\frac{dB}{dt} = \frac{dB}{dM} \cdot \frac{dM}{dt}$$

The above statement is called the chain rule, and is perhaps the most powerful of the differentiation laws. We will return to this soon, but first we'd like to figure out who wins the race! Since Kevin is running twice as fast as Brian, his position should be changing twice as quickly, or in other words

$$\frac{dK}{dB} = 2$$

which is a useful piece of information, but we really want to know dK/dM, to see who will win the race (of course $dM/dM = 1$, as Mark's position is changing at the same rate as Mark's position changes). If Kevin is running twice as fast as Brian, who is running 2/3 as fast as Mark, it stands to reason that Kevin should be running 4/3 as fast as Mark. Thus,

$$\frac{dK}{dM} = \frac{dK}{dB} \cdot \frac{dB}{dM}$$

Since Kevin turns out to be running the fastest (whoever is running the fast relative to Mark must also be running the fastest absolutely), he wins the race, and it looks like Brian will be buying breadsticks. We can go even one step further and find Kevin's velocity, not relative to Mark. If Mark is running 10mph, it follows Brian is running about 6.67mph, and finally Kevin is running twice that fast at about 13.33mph. Once again, written symbolically

$$\frac{dK}{dt} = \frac{dK}{dB} \cdot \frac{dB}{dM} \cdot \frac{dM}{dt}$$

This expression is the result of applying the chain rule twice.

The Chain Rule

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

for all such x that g is differentiable at x and f is differentiable at $g(x)$. In Leibniz notation, we write

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

where $\frac{df}{dg}$ is evaluated at $g(x)$.

In the composition $f(g(x))$ we refer to f as the outer function, and g as the inner function. We can describe the basic mechanism of the chain rule as follows: differentiate the outer function holding the inner function as a constant. Then, multiply the result by the derivative of the inner function. If there is a composition of more than two functions, the above process is simply repeated as many times as necessary. Consider the following examples.

Example 1 Find the derivative of $e^{\alpha t}$ (with respect to t), $\alpha \in \mathbb{R}$.
Solution The above function is a composition of two functions, e^u and $u = \alpha t$. Thus, we can apply the chain rule. We take the derivative of the outer function (which is e^u), evaluate the result at the inner function ($u = \alpha t$), differentiate the inner function (yielding α), and then multiply the results.

$$\frac{d}{dt} e^{\alpha t} = e^{\alpha t} \cdot \alpha = \alpha e^{\alpha t}$$

Thus, to find the derivative of an exponential function where the argument is multiplied by a constant, simply multiply the exponential function by that constant. This is consistent with the derivative of the exponential function, where that constant is simply a 1.

Example 2 Find the derivative of e^{x^2}.
Solution Once again we can apply the chain rule, to the composition of e^u and $u = x^2$. We take the derivative of the outer function (which is e^u), evaluate the result at the inner function ($u = x^2$), differentiate the inner function (yielding $2x$), and then multiply the results.

$$\frac{d}{dx} e^{x^2} = e^{x^2} \cdot 2x$$

Example 3 Find the derivative of a^t (with respect to t), where $a \in \mathbb{R}^+$.
Solution Using the chain rule we can find the derivative of a base a exponential using the derivative of the base e exponential. Write $a = e^{\ln(a)}$, which can be done as the exponential function and natural logarithm are inverses (we restricted a to positive real numbers, because otherwise we could not use this representation with the natural logarithm). Now take the derivative using the chain rule

$$\frac{d}{dt}a^t = \frac{d}{dt}e^{\ln(a)t} = e^{\ln(a)t} \cdot \ln(a)$$

Example 4 Find the derivative of $\frac{1}{1+y^2}$ (with respect to y).
Solution In order to solve this problem it is helpful to first rewrite the function as $(1+y^2)^{-1}$. Then we can see there is a composition of two functions u^{-1} and $u = 1+y^2$. Applying the chain rule we find

$$\frac{d}{dy}(1+y^2)^{-1} = -(1+y^2)^{-2} \cdot 2y = \frac{-2y}{(1+y^2)^2}$$

Example 5 Find the derivative of $\sin(1/x)$.
Solution Here we have $\sin(u)$ and $u = 1/x$. Applying the chain rule

$$\frac{d}{dx}\sin(1/x) = \cos(1/x) \cdot (-x^{-2}) = -\frac{\cos(1/x)}{x^2}$$

Example 6 Find the derivative of $x\sin(1/x)$
Solution We have already used the chain rule to find the derivative of $\sin(1/x)$, so now we just need to use the product rule

$$\frac{d}{dx}x\sin(1/x) = \sin(1/x) + x \cdot (-\frac{\cos(1/x)}{x^2}) = \sin(1/x) - \frac{\cos(1/x)}{x}$$

Example 7 Find the derivative of $\cos(\cos(\cos x))$.
Solution In this case we're actually looking at a composition of three functions we know how to differentiate. How can we apply the chain rule in this case? Let us choose our two functions as $\cos(u)$ and $u = \cos(\cos x)$. The chain rule tells us

$$\frac{d}{dx}\cos(\cos(\cos x)) = \frac{d}{du}\cos(u) \cdot \frac{d}{dx}\cos(\cos x) = -\sin(\cos(\cos x)) \cdot \frac{d}{dx}\cos(\cos x)$$

Now we need to find the derivative of $\cos(\cos x)$ to complete the problem. Choosing $\cos(u)$ and $u = \cos(x)$ we find

$$\frac{d}{dx}\cos(\cos x) = -\sin(\cos x) \cdot (-\sin x) = \sin(\cos x)\sin x$$

Now we can substitute this into the first equation to find

$$\frac{d}{dx}\cos(\cos(\cos x)) = -\sin(\cos(\cos x))\sin(\cos x)\sin x$$

This example leads us to an extremely useful observation. If we have a composition of three functions, to find the derivative we simply the multiply the derivatives of the three component functions together. Thus, we find that

$$\frac{d}{dx}f(g(h(x))) = \frac{df}{dg}\frac{dg}{dh}\frac{dh}{dx}$$

where we evalute df/dg at $g(h(x))$ and dg/dh at $h(x)$. We can easily extend the formula above to a composition of as many functions as we like.

Example 8 Find the derivative of $\tan(x^2+1)$.

Solution There is nothing specific about the chain rule that tells us *how* we must choose our inner and outer functions; the only requirement is we choose functions that we know how to differentiate (even if we need to use the chain rule again to do so). Let us choose $\tan(u)$ and $u = x^2 + 1$ as our functions, because we know how to differentiate both of them

$$\frac{d}{dx}\tan(x^2+1) = \sec^2(x^2+1)\cdot 2x$$

Choosing those two functions as our *chain* of functions worked well because we knew how to differentiate both of them. Nevertheless, another *chain* of functions would work just as well. Let us choose $\tan(u)$, $u = v + 1$ and $v = x^2$. Then we find

$$\frac{d}{dx}\tan(x^2+1) = \sec^2(x^2+1)\cdot 1\cdot 2x = \sec^2(x^2+1)\cdot 2x$$

where $du/dv = 1$ and $dv/dx = 2x$. The result is exactly the same as our previous result. The lesson to be learned here is that the most important thing to consider with the chain rule is choosing functions that are easy to differentiate. As long as the composition or *chain* of functions chosen to represent the original function is equivalent, it doesn't matter how many functions are chosen.

Example 9 Find the derivative of $y = (1 + \cot 2t)^{-4}$ with respect to t.
Solution Letting our representative functions be u^{-4}, $u = 1 + \cot v$ and $v = 2t$ we find

$$\frac{dy}{dt} = -4(1+\cot 2t)^{-5}\cdot(-\csc^2(2t))\cdot 2 = \frac{8\csc^2(2t)}{(1+\cot 2t)^5}$$

There is a final application of the chain rule which is extremely useful for increasing the amount of functions we can differentiate. If we have a function f and its inverse f^{-1}, by definition

$$(f^{-1}\circ f)(x) = (f\circ f^{-1})(x) = x$$

Now if two sides of an expression are equal, it follows that they must remain equal after differentiation (afterall, how can the same function have two different derivatives?). Using this observation, along with the chain rule, we find that

$$(f\circ f^{-1})'(x) = f'(f^{-1}(x))(f^{-1}(x))' = 1$$

where the right hand side comes from the fact $dx/dx = 1$. Solving this equation we find

$$(f^{-1}(x))' = \frac{1}{f'(f^{-1}(x))}$$

This amazing result is enough to allow us to find the derivative of a function we don't know, if we can find the derivative of its inverse function. Through this technique we can find the derivatives of all of the inverse trigonometric functions, as well as the natural logarithm.

Example 10 Find the derivative of $\ln(x)$
Solution If $f^{-1}(x) = \ln(x)$ then $f(x) = f'(x) = e^x$, so using the above result we find

$$\frac{d}{dx}\ln(x) = \frac{1}{e^{\ln(x)}} = \frac{1}{x}$$

because the exponential undoes the action of the logarithm (leaving us with x) in the above equation.