Section 4.1: The Chain Rule

You may have noticed that on one or two occasions this semester we've gone to extra lengths to avoid having to find the derivatives of certain functions directly.

Examples.

(1) On one of the recent homework sets, you had to differentiate $f(x) = \sin(2x)$; in order to do this you had to use a trig identity to rewrite $\sin(2x)$ as $2\sin(x)\cos(x)$, so we could use the Product Rule.

(2) Once or twice we've come across powers of polynomials, like $g(x) = (x^2 - x + 3)^3$. We've had to multiply this out before finding the derivative...and this multiplying out can be a heinous chore!

This is to say nothing of nasty functions like e^{x^2-1} and $\sin\left(\frac{x^2+1}{\cos(x)}\right)$, both of which are formed not by products, quotients, sums, or differences alone, but by ________________ .

Definition. Given two functions f and g, the ________________ of f with g (written $f \circ g$) is defined by

$$(f \circ g)(x) = f\big(g(x)\big).$$

Note that in order for this definition to make sense, we need x to be in the domain of g and $g(x)$ to be in the domain of f.

Examples. Let $f(x) = x^2 + 1$, $g(x) = \sqrt{x}$, and $h(x) = \sin(x)$. Find formulas for each of the indicated compositions.

(1) $(f \circ h)(x)$

(2) $(f \circ g)(x)$

(3) $(h \circ h)(x)$

(4) $(g \circ f)(x)$

Notice that in general $f \circ g \neq g \circ f$: the order in which we write the two matters.

To get a sense as to what the derivative of a composition should be, let's think about the following scenario:

Thought experiment. Suppose that $u = g(x)$ and $y = f(u)$. That is, $y = f(u) = f(g(x)) = (f \circ g)(x)$. What we'd like to know is the rate of change of this composition with respect to x: that is, what's $(f \circ g)'(x)$?

If we change x by a little bit, say Δx, we obtain a slight change in u, say Δu. How much does u change? You can draw a picture below reminding us that $\Delta u \approx g'(x)\Delta x$ when Δx is small:

Okay, so now we've got a small change in x leading to a small change in u. But we're really after is the change in y, Δy. Well, essentially the same picture as that above shows us that for small Δu, $\Delta y \approx f'(u)\Delta u \approx f'(u)g'(x)\Delta x$:

Recall that the derivative we're after is $(f \circ g)'(x)$, which, using Leibniz's notation, can be written $\frac{dy}{dx}$ since $y = f(u) = f\big(g(x)\big)$. But if Δx (and therefore Δy) is small, we know

$$\frac{dy}{dx} \approx \frac{\Delta y}{\Delta x}.$$

The formulas above now come to the rescue! As Δx goes to 0, we get the formula we're after:

The Chain Rule. Let $u = g(x)$ and $y = f(u)$, so that $y = f(u) = f\big(g(x)\big) = (f \circ g)(x)$. Then $(f \circ g)'(x) = f'(u)g'(x) = f'\big(g(x)\big)g'(x)$. In Leibniz's notation, the formula is even nicer: $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$.

Caveat. It is **INCORRECT** to write the Chain Rule as $(f \circ g)'(x) = f'(x)g'(x)$: notice that you have to find out what f' is at $g(x)$, **NOT** at x itself!

Examples. Find the derivative of each of the compositions given below. (*Hint:* when you first start out, it's never a bad idea to explicitly break down your composition into its constituent functions!)

(1) $\sin(x^3 + 2x)$

(2) $e^{\sqrt{x}}$

(3) $\sqrt{e^x}$

(4) $\frac{1}{\left(\cos(\theta)-\sin(\theta)\right)^3}$

(5) $\cos(\sin(t^4 + t))$

(6) $\sqrt[3]{e^{x^2} - x^2}$

Homework. The following exercises are due on *Friday, April 24th.*

(1) Let $f(x) = \cos(x)$, $g(x) = e^x$, and $h(x) = (x-1)^2$. Find each of the compositions given below, and then find the derivative of the composition using the Chain Rule.
   (a) $(f \circ g)(x)$
   (b) $(g \circ f)(x)$
   (c) $(f \circ f)(x)$
   (d) $(h \circ f)(x)$
   (e) $(h \circ h)(x)$

(2) For each function below, write it explicitly as a composition of the form $f \circ g$ and use the Chain Rule to find the derivative of the composition.
   (a) $(e^x + x^2)^{10}$
   (b) $\sin\big(\sin(\sin(x))\big)$
   (c) $\frac{1}{e^x+1}$

(3) Come up with two functions, f and g, such that $f \circ g = g \circ f$ *does* hold.

(4) Let $F(x) = \sin(2x)$. First find the derivative directly using the Chain Rule, and then find the derivative using the Product Rule after first applying the trig identity $\sin(2x) = 2\sin(x)\cos(x)$. What new trig identity to you get from knowing that the two answers you obtained are equal?