Notes for Section 2.5

In this section, we learn how to differentiate composite functions.

Theorem (The Chain Rule)**.** *Let f and g be functions of real numbers and $a \in \mathbb{R}$. If g is differentiable at a and f is differentiable at $g(a)$ then $f \circ g$ is also differentiable at a and*

$$(f \circ g)'(a) = f'(g(a))g'(a).$$

You can also express Chain Rule as

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

If you want to use the $\frac{d}{dx}$ notation to express the Chain Rule, then let $u = g(x)$ think of $f \circ g(x) = f(g(x))$ as $f(u)$:

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx}$$

I suggest you skip the discussion that follows the formula on p. 115 for now. Move on to the examples instead. The Chain Rule is not particularly difficult to use, it just takes some practice. Study Examples 1 and 2, then try a few examples yourself. You can try to make up your own, or pick from among exercises 2.5.1-16 (if you pick odd numbered exercises, you can check your solutions against the answers in the back of the book). I can certainly come up with examples myself, but wouldn't it be more interesting if you brought one to class?

The Power Rule combined with the Chain Rule is just a straightforward application of the Chain Rule to the special case when the outer function is a power function. I do not know why our textbook makes a big deal out of it and states it as a separate rule. If you know how to use the Chain Rule and the Power Rule, you also know how to do this. Do not waste your precious brain cells trying to memorize something that easily follows from the other two rules. Look at Examples 3 and 4 and notice that there is nothing special to them.

Once you are comfortable using the Chain Rule to find the derivative of simple compositions of two functions, look at Examples 5 and 6 to see how you can use the Chain Rule in combination with the some of the other rules we have already learned to differentiate a more complicated function. Make up an example for yourself by composing functions and also including products and/or quotients, or find one among the exercises. Finally, look at Examples 7 and 8 on how to deal with several nested layers of functions. Again make up an example for yourself by composing at least three functions, or find one among the exercises.

Do not stress about learning the proof of the Chain Rule on pp. 119-120. It uses a fairly sophisticated new idea and I would say that puts it a little above the grade of a course like MCS 119. Most importantly, I doubt that it would really help you understand why the Chain Rule says what it says. However, do turn your attention back to the discussion that follows the statement of the Chain Rule on p. 115. As the book points out, this relatively simple argument is not a correct way to prove the Chain Rule. But it is still worth studying it because it says more about why the Chain Rule is what it is than the actual correct proof at the end of the section. Here is my version of that argument.

Let f, g, and a be as in the statement of the Chain Rule. By the definition of the derivative,

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

The idea is to multiply the numerator and the denominator of the difference quotient by $g(x)-g(a)$. Then

$$\begin{aligned}
(f \circ g)'(a) &= \lim_{x\to a}\left[\frac{f(g(x))-f(g(a))}{x-a}\,\frac{g(x)-g(a)}{g(x)-g(a)}\right] \\
&= \lim_{x\to a}\left[\frac{f(g(x))-f(g(a))}{g(x)-g(a)}\,\frac{g(x)-g(a)}{x-a}\right] \\
&= \left[\lim_{x\to a}\frac{f(g(x))-f(g(a))}{g(x)-g(a)}\right]\left[\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\right] \qquad \text{by LL5.}
\end{aligned}$$

It is easy to recognize the second limit as $g'(a)$. To make the first limit easier to recognize, set $y = g(x)$ and $b = g(a)$. Then

$$\frac{f(g(x))-f(g(a))}{g(x)-g(a)} = \frac{f(y)-f(b)}{y-b}$$

which looks like it could be the difference quotient in the derivative of f at b. But the limit is as $x \to a$. Since g is differentiable at a, it must also be continuous at a. Therefore as $x \to a$, $g(x) \to g(a)$. That is as x approaches a, $y = g(x)$ approaches $b = g(a)$. Hence

$$\lim_{x\to a}\frac{f(g(x))-f(g(a))}{g(x)-g(a)} = \lim_{y\to b}\frac{f(y)-f(b)}{y-b} = f'(b) = f'(g(a)).$$

Putting the pieces together, we get

$$(f \circ g)'(a) = f'(g(a))g'(a),$$

which is exactly the Chain Rule. Unfortunately, this argument suffers from a significant flaw. When we multiplied and divided by $g(x) - g(a)$, we did not know if $g(x) - g(a) \neq 0$. In fact, it could be that $g(x) - g(a) = 0$. For example, this would be the case if g is a constant function. But dividing by 0 is not defined. So this kind of argument works well in same cases, but not in every case. Therefore it is not a legitimate proof of the Chain Rule. The problem can be avoided by some clever, but rather technical means. That is exactly what is done on pp. 119-120 in the textbook. Unfortunately, the ideas there are not well motivated and therefore difficult to follow. My take on this is that the flawed argument above gives you more intuition about why the Chain Rule says what it says than the correct argument. So I am happy if you can make sense of it, learn it, and understand why it is not a completely valid argument.

Now that we are equipped with the Chain Rule, we can give a proof of the Quotient Rule, which is much easier than the one given in Section 2.4. In fact, not only does it prove the Quotient Rule, it also shows you how you can avoid having to use (and having to memorize) the Quotient Rule. So suppose f and g are functions that are differentiable at x and $g(x) \neq 0$. We will prove

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}.$$

The idea is to write $f(x)/g(x) = f(x)[g(x)]^{-1}$ and use the Product Rule to differentiate the product and the Chain Rule to differentiate the composite function $[g(x)]^{-1}$.

$$
\begin{aligned}
\frac{d}{dx}\frac{f(x)}{g(x)} &= \frac{d}{dx}\big(f(x)[g(x)]^{-1}\big) \\
&= f'(x)[g(x)]^{-1} + f(x)\frac{d}{dx}[g(x)]^{-1} && \text{by the Product Rule} \\
&= f'(x)[g(x)]^{-1} + f(x)(-1)[g(x)]^{-2}g'(x) && \text{by the Chain Rule} \\
&= \frac{f'(x)}{g(x)} - \frac{f(x)g'(x)}{g^2(x)} \\
&= \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}
\end{aligned}
$$

Here is an example of how you can use this on the fly:

$$
\begin{aligned}
\frac{d}{dx}\frac{x^2+1}{\sin(x)} &= \frac{d}{dx}\big((x^2+1)[\sin(x)]^{-1}\big) \\
&= 2x[\sin(x)]^{-1} + (x^2+1)\frac{d}{dx}[\sin(x)]^{-1} && \text{by the Product Rule} \\
&= 2x[\sin(x)]^{-1} + (x^2+1)(-1)[\sin(x)]^{-2}\cos(x) && \text{by the Chain Rule} \\
&= \frac{2x}{\sin(x)} - \frac{(x^2+1)\cos(x)}{\sin^2(x)} \\
&= \frac{2x\sin(x) - (x^2+1)\cos(x)}{\sin^2(x)}
\end{aligned}
$$