Chapter 3. Derivatives

3.6. The Chain Rule

Note. In this section we state our last rule of differentiation, the Chain Rule. The **C**hain Rule allows us to differentiate **C**ompositions of functions. We still have some other functions we will differentiate (such as logarithms and inverse trigonometric functions in Sections 3.8 and 3.9), but we will not have any more *rules* of differentiation.

Note. *Thomas' Calculus* offers an intuitive argument for the Chain Rule by considering the following system of gears:

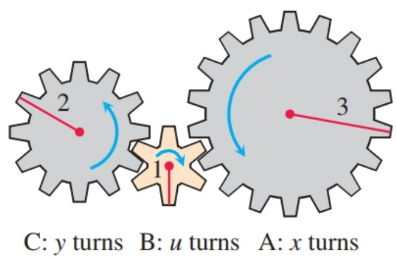

Figure 3.26

When gear A makes x turns, gear B makes u turns and gear C makes y turns. By comparing circumferences (or counting teeth), we see that $y = u/2$ (C turns one-half turn for each complete turn of B) and $u = 3x$ (B turns three times for each one complete turn of A), so $y = 3x/2$. Thus, $dy/dx = 3/2 = (1/2)(3) = (dy/du)(du/dx)$. This relationship between the rates of rotation (the "derivatives") suggests a relationship between rates of change of functions when one function is inside another. Here, we have that y (a function ultimately of x) can be written as the composition of y as a function of u, and u as a function of x; in the gear figure, the intermediate function u is represented by the center gear. So as "variable" x varies (i.e., as the right gear rotates), the center "function" u reacts to this variation of x (i.e., the center gear rotates in response to the rotation of the right gear), and "function" y reacts to the variation of u (i.e., the left gear rotates in response to the rotation of the center gear).

Note. We have already dealt with the composition of continuous functions in Theorem 2.9 (see Section 2.5. Continuity and Figure 2.42 of that section). We now consider a derivative (as a rate of change) of a composition $(f \circ g)(x) = f(g(x))$. We claim, as with the gears above, that the rate of change of $f \circ g$ with respect to x is the rate of change of f with respect to $u = g(x)$ times the rate of change of $u = g$ with respect to x. Consider the following figure:

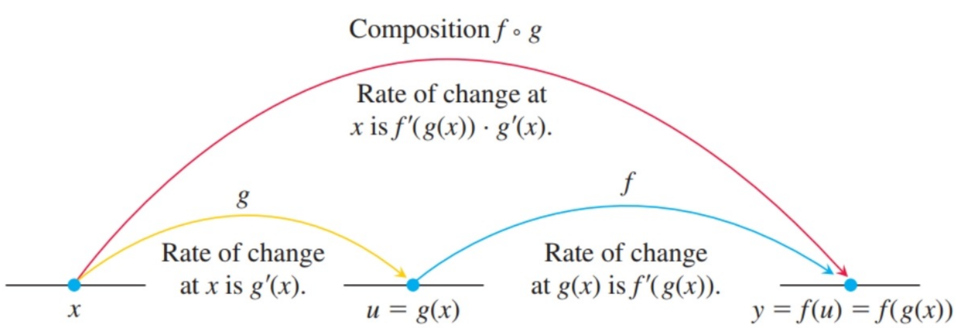

Figure 3.27

In this figure, think of x as the right gear, $u = g(x)$ as the center gear, and $y = f(u) = f(g) = f \circ g$ as the left gear in analogy with the gear idea of Figure 3.26 above.

Note. We are now ready to state the Chain Rule, which gives the derivative of the composition $(f \circ g)(x) = f(g(x))$, where f and g are each differentiable. We give a general proof of the Chain Rule in Section 3.11. Linearization and Differentials (see also Theorem 5-4 of my online notes for Analysis 1 [MATH 4217/5217] on 5.1. The Derivative of a Function).

Theorem 3.2. The Chain Rule.
If $f(u)$ is differentiable at the point $u = g(x)$ and $g(x)$ is differentiable at x, then the composite function $(f \circ g)(x) = f(g(x))$ is differentiable at x, and

$$(f \circ g)'(x) = f'(g(x))[g'(x)].$$

In Leibniz's notation, if $y = f(u)$ and $u = g(x)$, then

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx},$$

where dy/du is evaluated at $u = g(x)$.

Note. We give a general proof of the Chain Rule later. For now, we give a proof under the special condition that there is some $\varepsilon > 0$ such that $\Delta u = g(x + \Delta x) - g(x) \neq 0$ for all x in the domain of g and for all $\Delta x < \varepsilon$.

Note. The text book describes applying the Chain Rule as an "outside-inside" rule. Instead, we modify the square bracket notation slightly and when we use the Chain Rule, we will insert a little arrow indicating that the Chain Rule "spits out" the derivative of the inner function in the composition:

$$(f \circ g)'(x) = f'(g(x))\overset{\curvearrowright}{}[g'(x)] \text{ and } \frac{dy}{dx} = \frac{dy}{du}\overset{\curvearrowright}{}\frac{du}{dx}.$$

Note. If $f(u) = u^n$ where n is an integer, then

$$\frac{d}{dx}[f(g(x))] = \frac{d}{dx}[(g(x))^n] = ng(x)^{n-1}\overset{\curvearrowright}{}[g'(x)].$$

The text calls this the *Power Chain Rule.*

Examples. Exercise 3.6.8, Exercise 3.6.48, Exercise 3.6.64, Exercise 3.6.88, and Exercise 3.6.96.

Note. If $f(u) = e^u$ where $u = g(x)$ is a function of x, then

$$\frac{d}{dx}[e^u] = \frac{d}{dx}[e^{g(x)}] = e^{g(x)}\overset{\curvearrowright}{}[g'(x)].$$

Examples. Exercise 3.6.58 and Exercise 3.6.104.

Note. You now know a lot about differentiation since you know so many rules of differentiation. Also, you have the square bracket notation which lets you streamline the differentiation process (though you may want to simplify derivatives, especially later when applying differentiation to specific problems). To illustrate a combination of several differentiation rules at once, we consider the following differentiation problem.

Example 3.6.A. Differentiate $f(x) = \dfrac{(5x^3 - 4x + 2)^5}{(\tan^4 x)(e^{3x})}$.

Solution. We treat this as a quotient with a composition of functions in the numerator and a product in the denominator (each factor of which is a composition). We have

$$
\begin{aligned}
f'(x) &= \frac{d}{dx}\left[\frac{(5x^3 - 4x + 2)^5}{(\tan^4 x)(e^{3x})}\right] \\
&= \frac{\frac{d}{dx}[(5x^3 - 4x + 2)^5]\left((\tan^4 x)(e^{3x})\right) - \left((5x^3 - 4x + 2)^5\right)\frac{d}{dx}\left[(\tan^4 x)(e^{3x})\right]}{\left((\tan^4 x)(e^{3x})\right)^2} \\
&= \frac{\frac{d}{dx}[(5x^3 - 4x + 2)^5]\left((\tan^4 x)(e^{3x})\right)}{\left((\tan^4 x)(e^{3x})\right)^2} - \frac{\left((5x^3 - 4x + 2)^5\right)\frac{d}{dx}\left[(\tan^4 x)(e^{3x})\right]}{\left((\tan^4 x)(e^{3x})\right)^2} \\
&= \frac{[5(5x^3 - 4x + 2)^{\overset{\frown}{4}}[15x^2 - 4]]((\tan^4 x)(e^{3x}))}{((\tan^4 x)(e^{3x}))^{\mathbf{2}}} \\
&\quad - \frac{(5x^3 - 4x + 2)^5)[[(4\tan^3 \overset{\frown}{x})[\sec^2 x]](e^{3x}) + (\tan^4 x)[(\overset{\frown}{e^{3x}})[3])]}{((\tan^4 x)(e^{3x}))^{\mathbf{2}}}
\end{aligned}
$$

Here we have used red brackets to illustrate the use of the product rule, used blue brackets to illustrate the use of the quotient rule (though it is split up over two lines), and little arrows to indicate the uses of the Chain Rule.