3.4. **The chain rule.** (Sec. 3.4 in the textbook)

Theorem 3.7. *If g is differentiable at x and f is differentiable at $g(x)$, then $f \circ g$ is differentiable at x and $\frac{d}{dx}(f \circ g)(x) = f'(g(x)) \cdot g'(x)$.*

Remark 3.2. In Leibniz notation, if $y = f(u)$ and $u = g(x)$, then $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$ or more precisely, $\frac{dy}{dx} = \frac{dy}{du}|_{u=g(x)} \cdot \frac{du}{dx}$.

Proof. Fix x and set $\Delta u = g(x+h) - g(x)$, $L = g'(x)$ and $M = f'(g(x))$. Consider the following two cases.

Case 1: $L \neq 0$. In this case, $\Delta u \neq 0$ for h small enough. Write

$$\frac{f(g(x+h)) - f(g(x))}{h} = \frac{f(g(x)+\Delta u) - f(g(x))}{\Delta u} \cdot \frac{g(x+h) - g(x)}{h}.$$

Since g is differentiable at x, g is continuous at x. This implies

$$h \to 0 \quad \Rightarrow \quad \Delta u \to 0 \quad \Rightarrow \quad \frac{f(g(x)+\Delta u) - f(g(x))}{\Delta u} \to f'(g(x))$$

and also

$$h \to 0 \quad \Rightarrow \quad \frac{g(x+h) - g(x)}{h} \to g'(x).$$

By the limit laws, we have

$$\lim_{h \to 0} \frac{f(g(x+h)) - f(g(x))}{h} = f'(g(x)) \cdot g'(x).$$

Case 2: $L = 0$. Let $\epsilon > 0$ and choose $\eta > 0$ such that

$$|f(g(x)+k) - f(g(x)) - Mk| < \epsilon|k|, \quad \forall 0 < |k| < \eta.$$

For such ϵ, we may select $\delta > 0$ such that

$$|g(x+h) - g(x)| < \frac{\epsilon}{|M| + \epsilon}|h|, \quad \forall 0 < |h| < \delta.$$

Set $\delta_0 = \min\{\delta, \eta\}$. If $0 < |h| < \delta_0$ and $g(x+h) \neq g(x)$, then $0 < |g(x+h) - g(x)| < \eta$. By the triangle inequality, we have

$$|f(g(x+h)) - f(g(x))| \leq (|M| + \epsilon)|g(x+h) - g(x)| < \epsilon|h|.$$

It is clear that the above inequality also holds if $0 < |h| < \delta_0$ and $g(x+h) = g(x)$. This proves that $(f \circ g)'(x) = 0 = f'(g(x))g'(x)$. ☐

Example 3.7. Let g be a differentiable function and $f(x) = g(ax)$ with $a \in \mathbb{R}$. By letting $u = ax$ and $y = g(u)$, we have

$$\frac{dy}{du} = g'(u), \quad \frac{du}{dx} = a.$$

This implies

$$f'(x) = \left.\frac{dy}{du}\right|_{u=ax} \cdot \frac{du}{dx} = ag'(ax).$$

In particular,

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

Example 3.8. Let g be a differentiable function and $f(x) = e^{g(x)}$. To compute f', we set $u = g(x)$ and $y = e^u$. Note that

$$\frac{dy}{du} = e^u, \quad \frac{du}{dx} = g'(x).$$

By the chain rule, this implies

$$f'(x) = \frac{dy}{du}\bigg|_{u=g(x)} \cdot \frac{du}{dx} = g'(x)f(x).$$

When $g(x) = (\ln a)x$ with $a > 0$, we have

$$f(x) = e^{(\ln a)x} = a^x, \quad \frac{d}{dx}(a^x) = (\ln a)a^x.$$

Example 3.9. Let $g(x)$ be a positive differentiable function defined on an open interval I and set $f(x) = [g(x)]^r$ with $r \in \mathbb{R}$. By setting $u = g(x)$ and $y = u^r$, we have

$$\frac{dy}{du} = ru^{r-1}, \quad \frac{du}{dx} = g'(x) \quad \Rightarrow \quad f'(x) = r[g(x)]^{r-1}g'(x), \quad \forall x \in I.$$

In the case that $g(x) = x^2 - 1$ and $r = 1/2$, the last identity yields

$$\frac{d}{dx}\sqrt{x^2-1} = \frac{x}{\sqrt{x^2-1}}, \quad \forall |x| > 1.$$