Section 2.5 Chain Rule

First we will start with couple of motivational examples:

**Example** Find \( \frac{d}{dx}[(x^2 + 1)^2] = \frac{d}{dx}[(x^2 + 1)(x^2 + 1)] \).

The second part of the problem is suggesting the use of the Product Rule.

\[
\frac{d}{dx}[(x^2 + 1)^2] = \frac{d}{dx}[(x^2 + 1)(x^2 + 1)] \\
= (2x)(x^2 + 1) + (x^2 + 1)(2x) \\
= 2 \cdot (x^2 + 1) \cdot 2x
\]

**Example** Find \( \frac{d}{dx}[(x^2 + 1)^3] = \frac{d}{dx}[(x^2 + 1)(x^2 + 1)^2] \).

Another call for Product Rule. We will also need the help of the previous example:

\[
\frac{d}{dx}[(x^2 + 1)^3] = \frac{d}{dx}[(x^2 + 1)(x^2 + 1)^2] \\
= 2x(x^2 + 1)^2 + (x^2 + 1)[2 \cdot (x^2 + 1) \cdot 2x] \\
= 3 \cdot (x^2 + 1)^2 \cdot 2x
\]

By similar argument as above,

\[
\frac{d}{dx}[(x^2 + 1)^4] = 4 \cdot (x^2 + 1)^3 \cdot 2x \\
\frac{d}{dx}[(x^2 + 1)^5] = 5 \cdot (x^2 + 1)^4 \cdot 2x \\
\frac{d}{dx}[(x^2 + 1)^6] = 6 \cdot (x^2 + 1)^5 \cdot 2x
\]

Notice a pattern here? \( \frac{d}{dx}[(x^2 + 1)^n] = n \cdot (x^2 + 1)^{n-1} \cdot 2x \).

Here our original function \( (x^2 + 1)^n = f \circ g(x) \) is composition of two functions where \( f(u) = u^n \) is the "outer" function and \( u(x) = x^2 + 1 \) is the inner function. So we can re-interpret the pattern above as:

\[
\frac{d}{dx}[(x^2 + 1)^n] = \underbrace{n \cdot (x^2 + 1)^{n-1}}_{\text{derivative of outer function at } u(x)} \cdot \underbrace{2x}_{\text{derivative of inner function}}
\]
Let’s move towards a more general formula. Assume that \( f(x) \) is differentiable. Let’s calculate \( \frac{d}{dx} (f(x))^3 \)

\[
\frac{d}{dx} (f(x))^3 = \frac{d}{dx} [f(x)f(x)f(x)]
= f'(x)f(x)f(x) + f(x)f'(x)f(x) + f(x)f(x)f'(x)
= 3 \cdot (f(x))^2 \cdot f'(x)
\]

**Theorem (The Generalized Power Rule)** If \( f(x) \) is differentiable and \( n \) is a positive integer, then

\[
\frac{d}{dx} (f(x))^n = n(f(x))^{n-1}f'(x)
\]

For proof use the idea that I have used for \((f(x))^3\) above and prove it yourself.

**Example** \( \frac{d}{dx} ((x^3 - \pi)^7) \)

Here the ”outer” function is \( f(u) = u^7 \) and the ”inner” function is \( u(x) = x^3 - \pi \)

\[
\frac{d}{dx} [(x^3 - \pi)^7] = (\text{derivative of the outer function at } u(x)) \cdot (\text{derivative of the inner function})
= 7(u)^6 \cdot (3x^2)
= 7(x^3 - \pi)^6 \cdot (3x^2)
\]

Above investigation and the rule that followed only helps us to find the derivative of composition of two functions one of which is a power function such as \( f(u) = u^7 \) above. What about a general composition of two arbitrary functions?
**Theorem (The Chain Rule)** If $g$ is differentiable at $x$ and $f$ is differentiable at $g(x)$, then

$$
\frac{d}{dx}[f(g(x))] = f'(g(x)) \cdot g'(x)
$$

Sometimes it is easier to remember the Chain Rule in Leibnitz notation. If $y = f(g(x))$ then $\frac{dy}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$

**Flawed Proof** We must evaluate

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

in order to do this we need to find a way to separate the contribution of $f$ from the contribution of $g$ to this limit. The way I will do this is to multiply and divide by the quantity $[g(x) - g(a)]$:

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

Now I will evaluate limits (A) and (B) separately. Limit (B) is obviously $g'(a)$; nothing more needs to be said about that. As for limit (A), notice that as $x \to a$, $g(x) \to g(a)$, because $g$ is continuous at $a$ (it is differentiable remember!!). This means that

$$
(A) = \lim_{x \to a, g(x) \to g(a)} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} = \lim_{u \to g(a)} \frac{f(u) - f(g(a))}{u - g(a)}
$$

(In (C) I let $u$ stand for $g(x)$.) Now as $u = g(x) \to g(a)$, by the definition of the derivative, (C) must be approaching $f'(g(a))$. Thus the limit (A) equals $f'(g(a))$ and the product of the two limits (A) and (B) is: $f'(g(a)) \cdot g'(a)$
The Flaw of the Proof above The problem with this proof is this: I can’t be sure that I did not multiply and divide by zero. The definition of limit guarantees that \((x \to a)\) \(x\) will not equal \(a\); but it cannot guarantee that \(g(x) \neq g(a)\). This, after all, depends on the what function \(g\) does. Let me emphasize that this is the only flaw in the proof; if \(g\) happens to be a function for which \(g(x) \neq g(a)\) whenever \(x \neq a\), then the proof above is perfectly valid: by making \(g(x)\) sufficiently close to \(g(a)\) (which can be done by choosing \(x\) close to \(a\)), you can make \(\frac{f(g(x))-f(g(a))}{x-a}\) as close as you like to \(f'(g(a))\).

For those interested readers I am attaching the ”How to correct the flaw” discussion at the end of these notes. I know the ”correction proof” requires an acquired taste for proof and I do not expect everyone to look at it.

Example Find \(F'(x)\) when \(F(x) = \sqrt[3]{x^2 - 1}\)

De-compose your functions \(F(x) = f(g(x))\); where ”outer” function \(f(g) = g^{1/3}\) and ”inner” function \(g(x) = x^2 - 1\). Find the derivatives of each of these functions as you’ll need it for the Chain Rule.

\[
\begin{align*}
\frac{df}{dg} &= \frac{1}{3}g^{-2/3} \\
\frac{dg}{dx} &= 2x
\end{align*}
\]

Use the Chain Rule now to find

\[
F'(x) = f'(g(x)) \cdot g'(x) = \frac{1}{3}(x^2 - 1)^{-2/3} \cdot 2x = \frac{2x}{3(x^2 - 1)^{2/3}}
\]

Example Let \(y = (x^3 - 1)^{100}\) and find \(\frac{dy}{dx}\)

Let \(y = u^{100}\) and \(u(x) = x^3 - 1\). By Chain Rule (Leibnitz notation)

\[
\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}
\]

So again find each of the derivatives on the right hand side above and you are done.

\[
\frac{du}{dx} = 100u^{99} \quad \text{and} \quad \frac{du}{dx} = 3x^2 \Rightarrow \frac{dy}{dx} = 100(x^3 - 1)^{99} \cdot 3x^2 = 300x^2(x^3 - 1)
\]

Example Let \(f(x) = \frac{1}{\sqrt{x^2 + x + 1}}\) \(= (x^2 + x + 1)^{-1/2}\). Find \(f'(x)\) ?

\(f(x) = u(v(x))\) where \(u(v) = v^{-1/5}\) and \(v(x) = x^2 + x + 1\).

\[
\begin{align*}
\frac{df}{dx} &= \frac{du}{dv} \cdot \frac{dv}{dx} = -\frac{1}{5}(x^2 + x + 1)^{-6/5}(2x + 1) = \frac{-(2x + 1)}{5(x^2 + x + 1)^{6/5}}
\end{align*}
\]
Example Let \( g(t) = \left( \frac{t-2}{2t+1} \right)^9 \). Find \( g'(t) \).

\[
y = g(t) \text{ decomposes into } y = u^g \text{ and } u(t) = \frac{t-2}{2t+1}
\]

\[
\frac{dy}{dt} = \frac{dy}{du} \cdot \frac{du}{dx} = 9 \left( \frac{t-2}{2t+1} \right)^8 \cdot \frac{d}{dt} \left( \frac{t-2}{2t+1} \right)
\]

needs Quotient Rule

\[
\frac{d}{dt} \left( \frac{t-2}{2t+1} \right) = \frac{d}{dt} \left( \frac{f}{g} \right) = \frac{f'g - fg'}{g^2} \text{; since } f'(t) = 1 \text{ and } g'(t) = 2 \Rightarrow
\]

\[
\frac{d}{dt} \left( \frac{t-2}{2t+1} \right) = \frac{1 \cdot (2t+1) - (t-2) \cdot 2}{(2t+1)^2}
\]

\[
\frac{dy}{dt} = 9 \left( \frac{t-2}{2t+1} \right)^8 \cdot \left[ \frac{1 \cdot (2t+1) - (t-2) \cdot 2}{(2t+1)^2} \right] = \frac{45(t-2)^8}{(2t+1)^6}
\]

Example Let \( y = (2x + 1)^5(x^3 - x + 1) \) and find \( y' \).

\[
y' = \frac{d}{dx} \left[(2x + 1)^5 \cdot (x^3 - x + 1) + (2x + 1)^5 \cdot \frac{d}{dx} [(x^3 - x + 1)] \right] \text{ by Product Rule}
\]

\[
\text{Chain Rule}
\]

\[
= 5(2x + 1)^4 \cdot (2x) \cdot (x^3 - x + 1) + (2x + 1)^5(3x^2 - 1)
\]

**Theorem (Derivative of inverse functions)** If \( f(x) \) is differentiable at all \( x \), and has an inverse \( g(x) = f^{-1}(x) \), then

\[
g'(x) = -\frac{1}{f'(g(x))} \quad \text{provided } f'(g(x)) \neq 0
\]

**Proof** Recall that if \( f \) and \( g \) are inverses, then \( f(g(x)) = x \) for all \( x \) in the domain of \( g \). Differentiate both sides of this equality with respect to \( x \):

\[
\frac{d}{dx} [f(g(x))] = \frac{d}{dx} [x]
\]

by Chain Rule

\[
f'(g(x)) \cdot g'(x) = 1
\]

Hence \( g'(x) = \frac{1}{f'(g(x))} \)
Example Let \( f(x) = x^2 \) for \( x \geq 0 \), then \( g(x) = f^{-1}(x) = \sqrt{x} \) for \( x \geq 0 \). We want to find \( g'(x) \) and by theorem we need \( f'(x) = 2x \). So

\[
(f^{-1}(x))' = g'(x) = \frac{1}{f'(g(x))} = \frac{1}{2g(x)} = \frac{1}{2\sqrt{x}}
\]

For this problem you can also compute the answer by finding the derivative of \( g(x) \) using Power Rule as follows:

\[
g'(x) = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}
\]

Example Let \( x = \sqrt{y} + 5 \). Find \( \frac{dy}{dx} \)

Since our problem is given \( x \) in terms of \( y \) I’ll find \( \frac{dx}{dy} \) first and use the theorem above to find \( \frac{dy}{dx} \)

\[
\frac{dx}{dy} = \frac{1}{2}y^{-1/2} = \frac{1}{2\sqrt{y}}
\]

Since \( \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} \) by the theorem, we have:

\[
\frac{dy}{dx} = 2\sqrt{y} = 2(x - 5)
\]

Example Air is being pumped into a spherical balloon at the constant rate of \( 200\pi \, \text{cm}^3/\text{s} \). What is the rate of increase of the radius \( r \) when \( r = 5 \, \text{cm} \). (Hint: The volume of a sphere \( V = \frac{4\pi r^3}{3} \))

In this problem, both the volume and the radius are changing in time \( t \), i.e. \( V(t) = \frac{4\pi(r(t))^3}{3} \)

\[
\frac{dV}{dt} = \frac{d}{dt} \left[ \frac{4\pi(r(t))^3}{3} \right] = 4\pi r^2 \cdot \frac{dr}{dt} = 4\pi r^2 \cdot \frac{dr}{dt}
\]

In the problem we are given \( \frac{dV}{dt} = 200\pi \, \text{cm}^3/\text{s} \), and at this moment of time, \( r = 5 \, \text{cm} \). Using this info and solving the above expression for \( \frac{dr}{dt} \):

\[
200\pi \, \text{cm}^3/\text{s} = 4\pi(5 \, \text{cm})^2 \cdot \frac{dr}{dt} \Rightarrow \frac{dr}{dt} = \frac{200\pi \, \text{cm}^3/\text{s}}{100\pi \, \text{cm}^2} = 2 \, \text{cm/s}
\]
How to correct the flaw in the Proof of Chain Rule

We will start by asking: So what if \( g \) is a function for which 
\[ g(x) = g(a) \]
for lots of numbers \( x \neq a \)? The Chain Rule should hold in this case as well. How can we fix the flaw and make a proof that works for all differentiable functions \( g \)?

One way to proceed is to make a new function that is like:
\[ f(g(x)) - f(g(a)) \]
\[ g(x) - g(a) \]
(the source of the problem), but which won’t care whether \( g(x) \) equals \( g(a) \) or not. This can be done as follows. Start with \(^1\)

\[
F(u) = \begin{cases} 
  \frac{f(u) - f(g(a))}{u - g(a)} & \text{if } u \neq g(a) \\
  f'(g(a)) & \text{if } u = g(a)
\end{cases}
\]

So that for any \( x \),

\[
F(g(x)) = \begin{cases} 
  \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} & \text{if } g(x) \neq g(a) \\
  f'(g(a)) & \text{if } g(x) = g(a)
\end{cases}
\]

Clearly the definition of \( f'(g(a)) \) and (1) guarantee that
\[
\lim_{u \to g(a)} F(u) = F(g(a)) \quad (\text{i.e. } F \text{ is continuous at } "g(a)"
\]
; therefore, since \( \lim g(x) = g(a) \), the Continuity Theorem (2nd Theorem in the lectures of Section 1.4) implies that

\[
\lim_{x \to a} F(g(x)) = F(\lim_{x \to a} g(x)) = F(g(a)) \overset{\text{by (2)}}{=} f'(g(a))
\]

Moreover, (2) implies the following lemma, which says that \( F(g(x)) \) and
\[
\frac{f(g(x)) - f(g(a))}{g(x) - g(a)}
\]
are similar enough to afford a route around the flaw in the crucial limit computation.

Lemma For any \( x \neq a \)

\[
\frac{f(g(x)) - f(g(a))}{x - a} = F(g(x)) \cdot \frac{g(x) - g(a)}{x - a} \quad (\text{E})
\]

\[
\frac{g(x) - g(a)}{x - a} \quad (\text{D})
\]

\(^1\)Here, "u" is a pure variable rather than a stand-in for \( g(x) \).
Proof There are two cases. If $g(x) \neq g(a)$, then (by (2))

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

so that

$$F(g(x)) \cdot \frac{g(x) - g(a)}{x - a} = \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \cdot \frac{g(x) - g(a)}{x - a} = \frac{f(g(x)) - f(g(a))}{x - a}$$

On the other hand, if $g(x) = g(a)$, then in (4), both (E) and (D) equal zero, while $F(g(x)) = f'(g(a))$; thus in this case (4) just says ”0 = 0”

Now using the Lemma and the discussion above it we can have the ”flawless” proof of Chain Rule as follows:

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

$$= \lim_{x \to a} F(g(x)) \cdot \frac{g(x) - g(a)}{x - a}$$

by (4)

$$= \lim_{x \to a} F(g(x)) \cdot \lim_{x \to a} \frac{g(x) - g(a)}{x - a}$$

by (3)

$$= f'(g(a)) \cdot g'(a)$$