3

DIFFERENTIATION RULES

3.4
The Chain Rule

In this section, we will learn about:

Differentiating composite functions using the Chain Rule.

CHAIN RULE

Suppose you are asked to differentiate the function

$$F(x) = \sqrt{x^2 + 1}$$

- The differentiation formulas you learned in the previous sections of this chapter do not enable you to calculate $F'(x)$.

CHAIN RULE

Observe that *F* is a composite function.

In fact, if we let $y = f(u) = \sqrt{u}$ and let $u = g(x) = x^2 + 1$, then we can write

$y = F(x) = f(g(x))$.

That is, $F = f \quad g$.

CHAIN RULE

We know how to differentiate both *f* and *g*.

So, it would be useful to have a rule that shows us how to find the derivative of $F = f \circ g$ in terms of the derivatives of *f* and *g*.

CHAIN RULE

It turns out that the derivative of the composite function $f \circ g$ is the product of the derivatives of f and g.

This fact is one of the most important of the differentiation rules. It is called the Chain Rule.

CHAIN RULE

It seems plausible if we interpret derivatives as rates of change.

Regard:

- du/dx as the rate of change of u with respect to x
- dy/du as the rate of change of y with respect to u
- dy/dx as the rate of change of y with respect to x

CHAIN RULE

If *u* changes twice as fast as *x* and *y* changes three times as fast as *u*, it seems reasonable that *y* changes six times as fast as *x*.

So, we expect that: $$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

THE CHAIN RULE

If g is differentiable at x and f is differentiable at $g(x)$, the composite function $F = f \circ g$ defined by $F(x) = f(g(x))$ is differentiable at x and F' is given by the product:

$$F'(x) = f'(g(x)) \bullet g'(x)$$

- In Leibniz notation, if $y = f(u)$ and $u = g(x)$ are both differentiable functions, then: $\dfrac{dy}{dx} = \dfrac{dy}{du}\dfrac{du}{dx}$

COMMENTS ON THE PROOF

Let Δu be the change in corresponding to a change of Δx in x, that is,

$$\Delta u = g(x + \Delta x) - g(x)$$

Then, the corresponding change in y is:

$$\Delta y = f(u + \Delta u) - f(u)$$

COMMENTS ON THE PROOF

Equation 1

It is tempting to write:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

$$\boxed{= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}}$$

$$= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta u} \cdot \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x}$$

$$= \lim_{\Delta u \to 0} \frac{\Delta y}{\Delta u} \cdot \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} = \frac{dy}{du} \frac{du}{dx}$$

COMMENTS ON THE PROOF

The only flaw in this reasoning is that, in Equation 1, it might happen that $\Delta u = 0$ (even when $\Delta x \neq 0$) and, of course, we can't divide by 0.

COMMENTS ON THE PROOF

Nonetheless, this reasoning does at least suggest that the Chain Rule is true.

- A full proof of the Chain Rule is given at the end of the section.

The Chain Rule can be written either in the prime notation

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$$

or, if $y = f(u)$ and $u = g(x)$, in Leibniz notation:

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

CHAIN RULE

Equation 3 is easy to remember because, if *dy*/*du* and *du*/*dx* were quotients, then we could cancel *du*.

However, remember:

- *du* has not been defined
- *du*/*dx* should not be thought of as an actual quotient

Find $F'(x)$ if $F(x)=\sqrt{x^2+1}$

- One way of solving this is by using Equation 2.
- At the beginning of this section, we expressed F as $F(x) = (f \circ g)(x) = f(g(x))$ where $f(u)=\sqrt{u}$ and $g(x) = x^2 + 1$.

E. g. 1—Solution 1

- Since $f'(u) = \frac{1}{2}u^{-1/2} = \dfrac{1}{2\sqrt{u}}$ and $g'(x) = 2x$

we have $F'(x) = f'(g(x)) \cdot g'(x)$

$$= \frac{1}{2\sqrt{x^2+1}} \cdot 2x$$

$$= \frac{x}{\sqrt{x^2+1}}$$

We can also solve by using Equation 3.

- If we let $u = x^2 + 1$ and $y = \sqrt{u}$, then:

$$F'(x) = \frac{dy}{dx}\frac{du}{dx} = \frac{1}{2\sqrt{u}}(2x)$$

$$= \frac{1}{2\sqrt{x^2+1}}(2x) = \frac{x}{\sqrt{x^2+1}}$$

When using Equation 3, we should bear in mind that:

- *dy*/*dx* refers to the derivative of *y* when *y* is considered as a function of *x* (called the derivative of *y* with respect to *x*)
- *dy*/*du* refers to the derivative of *y* when considered as a function of *u* (the derivative of *y* with respect to *u*)

CHAIN RULE

For instance, in Example 1, *y* can be considered as a function of *x* $(y = \sqrt{x^2 + 1})$ and also as a function of *u* $(y = \sqrt{u})$.

- Note that:

$$\frac{dy}{dx} = F'(x) = \frac{x}{\sqrt{x^2 + 1}} \quad \text{whereas} \quad \frac{dy}{du} = f'(u) = \frac{1}{2\sqrt{u}}$$

NOTE

In using the Chain Rule, we work from the outside to the inside.

- Equation 2 states that we differentiate the outer function *f* [at the inner function *g*(*x*)] and then we multiply by the derivative of the inner function.

$$\frac{d}{dx}\underbrace{f}_{\text{outer function}}\underbrace{(g(x))}_{\text{evaluated at inner function}} = \underbrace{f'}_{\text{derivative of outer function}}\underbrace{(g(x))}_{\text{evaluated at inner function}} \cdot \underbrace{g'(x)}_{\text{derivative of inner function}}$$

Example 2

Differentiate:

a. $y = \sin(x^2)$

b. $y = \sin^2 x$

Example 2 a

If $y = \sin(x^2)$, the outer function is the sine function and the inner function is the squaring function.

So, the Chain Rule gives:

$$\frac{dy}{dx} = \frac{d}{dx} \underbrace{\sin}_{\text{outer function}} \underbrace{(x^2)}_{\text{evaluated at inner function}} = \underbrace{\cos}_{\text{derivative of outer function}} \underbrace{(x^2)}_{\text{evaluated at inner function}} \cdot \underbrace{2x}_{\text{derivative of inner function}}$$

$$= 2x\cos(x^2)$$

Example 2 b

Note that $\sin^2 x = (\sin x)^2$. Here, the outer function is the squaring function and the inner function is the sine function.

Therefore,

$$\frac{dy}{dx} = \frac{d}{dx}\underbrace{(\sin x)^2}_{\text{inner function}} = \underbrace{2\,.}_{\text{derivative of outer function}} \underbrace{\sin x}_{\text{evaluated at inner function}} \cdot \underbrace{\cos x}_{\text{derivative of inner function}}$$

The answer can be left as $2 \sin x \cos x$ or written as $\sin 2x$ (by a trigonometric identity known as the double-angle formula).

COMBINING THE CHAIN RULE

In Example 2 a, we combined the Chain Rule with the rule for differentiating the sine function.

COMBINING THE CHAIN RULE

In general, if $y = \sin u$, where u is a differentiable function of x, then, by the Chain Rule,

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \cos u \frac{du}{dx}$$

Thus,

$$\frac{d}{dx}(\sin u) = \cos u \frac{du}{dx}$$

COMBINING THE CHAIN RULE

In a similar fashion, all the formulas for differentiating trigonometric functions can be combined with the Chain Rule.

COMBINING CHAIN RULE WITH POWER RULE

Let’s make explicit the special case of the Chain Rule where the outer function is a power function.

- If $y = [g(x)]^n$, then we can write $y = f(u) = u^n$ where $u = g(x)$.
- By using the Chain Rule and then the Power Rule, we get:

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = nu^{n-1}\frac{dy}{dx} = n[g(x)]^{n-1}g'(x)$$

POWER RULE WITH CHAIN RULE **Rule 4**

If n is any real number and $u = g(x)$ is differentiable, then $$\frac{d}{dx}(u^n) = nu^{n-1}\frac{du}{dx}$$

Alternatively, $$\frac{d}{dx}[g(x)]^n = n[g(x)]^{n-1} \cdot g'(x)$$

POWER RULE WITH CHAIN RULE

Notice that the derivative in Example 1 could be calculated by taking $n = ½$ in Rule 4.

Differentiate $y = (x^3 - 1)^{100}$

- Taking $u = g(x) = x^3 - 1$ and $n = 100$ in the rule, we have:

$$\frac{dy}{dx} = \frac{d}{dx}(x^3 - 1)^{100}$$

$$= 100(x^3 - 1)^{99} \frac{d}{dx}(x^3 - 1)$$

$$= 100(x^3 - 1)^{99} \cdot 3x^2$$

$$= 300x^2(x^3 - 1)^{99}$$

POWER RULE WITH CHAIN RULE

Example 4

Find *f'* (*x*) if $f(x) = \dfrac{1}{\sqrt[3]{x^2 + x + 1}}$.

- First, rewrite *f* as $f(x) = (x^2 + x + 1)^{-1/3}$
- Thus, $f'(x) = -\frac{1}{3}(x^2 + x + 1)^{-4/3}\frac{d}{dx}(x^2 + x + 1)$

$$= -\frac{1}{3}(x^2 + x + 1)^{-4/3}(2x + 1)$$

POWER RULE WITH CHAIN RULE

Example 5

Find the derivative of $g(t)=\left(\frac{t-2}{2t+1}\right)^9$

- Combining the Power Rule, Chain Rule, and Quotient Rule, we get:

$$g'(t)=9\left(\frac{t-2}{2t+1}\right)^8\frac{d}{dt}\left(\frac{t-2}{2t+1}\right)$$

$$=9\left(\frac{t-2}{2t+1}\right)^8\frac{(2t+1)\cdot 1-2(t-2)}{(2t+1)^2}=\frac{45(t-2)^8}{(2t+1)^{10}}$$

Example 6

Differentiate:

$$y = (2x + 1)^5 (x^3 - x + 1)^4$$

- In this example, we must use the Product Rule before using the Chain Rule.

Thus,

$$\frac{dy}{dx} = (2x+1)^5 \frac{d}{dx}(x^3 - x + 1)^4 + (x^3 - x + 1)^4 \frac{d}{dx}(2x+1)^5$$

$$= (2x+1)^5 \cdot 4(x^3 - x + 1)^3 \frac{d}{dx}(x^3 - x + 1)$$

$$+(x^3 - x + 1)^4 \cdot 5(2x+1)^4 \frac{d}{dx}(2x+1)$$

$$= 4(2x+1)^5 (x^3 - x + 1)^3 (3x^2 - 1)$$

$$+5(x^3 - x + 1)^4 (2x+1)^4 \cdot 2$$

Example 6

Noticing that each term has the common factor $2(2x + 1)^4(x^3 – x + 1)^3$, we could factor it out and write the answer as:

$$\frac{dy}{dx} = 2(2x+1)^4(x^3 - x + 1)^3(17x^3 + 6x^2 - 9x + 3)$$

Example 7

Differentiate $y = e^{\sin x}$

- Here, the inner function is $g(x) = \sin x$ and the outer function is the exponential function $f(x) = e^x$.

- So, by the Chain Rule:

$$\frac{dy}{dx} = \frac{d}{dx}(e^{\sin x})$$

$$= e^{\sin x} \frac{d}{dx}(\sin x) = e^{\sin x} \cos x$$

CHAIN RULE

We can use the Chain Rule to differentiate an exponential function with any base $a > 0$.

- Recall from Section 1.6 that $a = e^{\ln a}$.
- So, $a^x = (e^{\ln a})^x = e^{(\ln a)x}$.

Thus, the Chain Rule gives

$$\frac{d}{dx}(a^x) = \frac{d}{dx}(e^{(\ln a)x}) = e^{(\ln a)x}\frac{d}{dx}(\ln a)x$$

$$= e^{(\ln a)x} \cdot \ln a = a^x \ln a$$

because $\ln a$ is a constant.

Therefore, we have the formula:

$$\frac{d}{dx}(a^x) = a^x \ln a$$

In particular, if *a* = 2, we get:

$$\frac{d}{dx}(2^x) = 2^x \ln 2$$

In Section 3.1, we gave the estimate

$$\frac{d}{dx}(2^x) \approx (0.69)2^x$$

- This is consistent with the exact Formula 6 because $\ln 2 \approx 0.693147$

CHAIN RULE

The reason for the name ‘Chain Rule’ becomes clear when we make a longer chain by adding another link.

Suppose that $y = f(u)$, $u = g(x)$, and $x = h(t)$, where *f, g*, and *h* are differentiable functions, then, to compute the derivative of *y* with respect to *t*, we use the Chain Rule twice:

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

If $f(x) = \sin(\cos(\tan x))$, then

$$f'(x) = \cos(\cos(\tan x))\frac{d}{dx}\cos(\tan x)$$

$$= \cos(\cos(\tan x))[-\sin(\tan x)]\frac{d}{dx}(\tan x)$$

$$= -\cos(\cos(\tan x))\sin(\tan x)\sec^2 x$$

- Notice that we used the Chain Rule twice.

Example 9

Differentiate $y = e^{\sec 3\theta}$

- The outer function is the exponential function, the middle function is the secant function and the inner function is the tripling function.
- Thus, we have: $\frac{dy}{d\theta} = e^{\sec 3\theta} \frac{d}{d\theta}(\sec 3\theta)$

$$= e^{\sec 3\theta} \sec 3\theta \tan 3\theta \frac{d}{d\theta}(3\theta)$$

$$= 3e^{\sec 3\theta} \sec 3\theta \tan 3\theta$$

HOW TO PROVE THE CHAIN RULE

Recall that if $y = f(x)$ and x changes from a to $a + \Delta x$, we defined the increment of y as:

$$\Delta y = f(a + \Delta x) - f(a)$$

HOW TO PROVE THE CHAIN RULE

According to the definition of a derivative, we have:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f\,'(a)$$

HOW TO PROVE THE CHAIN RULE

So, if we denote by ε the difference between the difference quotient and the derivative, we obtain:

$$\lim_{\Delta x \to 0} \varepsilon = \lim_{\Delta x \to 0} \left(\frac{\Delta y}{\Delta x} - f'(a) \right)$$

$$= f'(a) - f'(a) = 0$$

HOW TO PROVE THE CHAIN RULE

However,

$$\varepsilon = \frac{\Delta y}{\Delta x} - f'(a) \quad \Rightarrow \quad \Delta y = f'(a)\Delta x + \varepsilon\Delta x$$

- If we define ε to be 0 when $\Delta x = 0$, then ε becomes a continuous function of Δx .

HOW TO PROVE THE CHAIN RULE **Equation 7**

Thus, for a differentiable function *f*, we can write:

$$\Delta y = f'(a)\Delta x + \varepsilon\ \Delta x \quad \text{where } \varepsilon \to 0 \text{ as } \Delta x \to 0$$

- ε is a continuous function of Δx.
- This property of differentiable functions is what enables us to prove the Chain Rule.

PROOF OF THE CHAIN RULE

Equation 8

Suppose $u = g(x)$ is differentiable at a and $y = f(u)$ at $b = g(a)$.

If Δx is an increment in x and Δu and Δy are the corresponding increments in u and y, then we can use Equation 7 to write

$$\Delta u = g'(a)\, \Delta x + \varepsilon_1\, \Delta x = [g'(a) + \varepsilon_1]\, \Delta x$$

where $\varepsilon_1 \to 0$ as $\Delta x \to 0$

Equation 9

Similarly,

$$\Delta y = f'(b)\,\Delta u + \varepsilon_2\,\Delta u = [f'(b) + \varepsilon_2]\,\Delta u$$

where $\varepsilon_2 \to 0$ as $\Delta u \to 0$.

PROOF OF THE CHAIN RULE

If we now substitute the expression for Δu from Equation 8 into Equation 9, we get:

$$\Delta y = [f'(b) + \varepsilon_2][g'(a) + \varepsilon_1]\Delta x$$

So,

$$\frac{\Delta y}{\Delta x} = [f'(b) + \varepsilon_2][g'(a) + \varepsilon_1]$$

PROOF OF THE CHAIN RULE

As $\Delta x \rightarrow 0$, Equation 8 shows that

$\Delta u \rightarrow 0$.

So, both $\varepsilon_1 \rightarrow 0$ and $\varepsilon_2 \rightarrow 0$ as $\Delta x \rightarrow 0$.

PROOF OF THE CHAIN RULE

Therefore, $$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

$$= \lim_{\Delta x \to 0} [f'(b) + \varepsilon_2][g'(a) + \varepsilon_1]$$

$$= f'(b)g'(a)$$

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

This proves the Chain Rule.