The Chain Rule

The **Chain Rule** computes the derivative of the **composite** of two functions. The **composite** $(f\circ g)(x)$ is just "g inside f" — that is,

$$(f\circ g)(x) = f\left(g(x)\right).$$

(Note that this is *not* multiplication!)

Here are some examples:

$$(x^3+x^2-7x+1)^{99} \quad \text{is} \quad g(x)=x^3+x^2-7x+1 \quad \text{inside} \quad f(x)=x^{99}.$$

$$\begin{array}{c}(\qquad\qquad\qquad)^{99}\\ \uparrow \\ x^3+x^2-7x+1\end{array}$$

$$\frac{1}{x^2-x-1} \quad \text{is} \quad g(x)=x^2-x-1 \quad \text{inside} \quad f(x)=\frac{1}{x}.$$

$$\begin{array}{c}\dfrac{1}{(\qquad\qquad)}\\ \uparrow \\ x^2-x-1\end{array}$$

$$\sin(x^2) \quad \text{is} \quad g(x)=x^2 \quad \text{inside} \quad f(x)=\sin x.$$

$$\begin{array}{c}\sin\ (\quad)\\ \uparrow \\ x^2\end{array}$$

Here's a more complicated example:

$$\cos\frac{1}{x^2-2x+5} \quad \text{is} \quad h(x)=x^2-2x+5 \quad \text{inside} \quad g(x)=\frac{1}{x} \quad \text{inside} \quad f(x)=\cos x.$$

$$\begin{array}{c}\cos\,(\qquad\qquad\qquad)\\ \uparrow \\ \dfrac{1}{(\qquad\qquad)}\\ \uparrow \\ x^2-2x+5\end{array}$$

One way to tell which function is "inside" and which is "outside" is to think about how you would plug numbers in. For example, take $p(x)=\sin(x^2)$. What would you do to compute $p(1.7)$ on your calculator? First, you'd square 1.7 — $1.7^2 = 1.89$. Next, you'd take the sine of that — $\sin 1.89 \approx 0.94949$.

The function you did first — squaring — is the *inner* function. The function you did second — sine — is the *outer* function.

Example. Suppose

$$f(x)=\frac{1}{x} \quad \text{and} \quad g(x)=x^2+1.$$

Compute $(f \circ g)(x)$, $(g \circ f)(x)$, and $(f \circ f)(x)$.

$$(f \circ g)(x) = f(g(x)) = f\left(x^2+1\right) = \frac{1}{x^2+1}.$$

$$(g \circ f)(x) = g(f(x)) = g\left(\frac{1}{x}\right) = \frac{1}{\left(\frac{1}{x}\right)^2+1} = \frac{x^2}{1+x^2}.$$

$$(f \circ f)(x) = f(f(x)) = f\left(\frac{1}{x}\right) = \frac{1}{\frac{1}{x}} = x. \quad \square$$

Theorem. (**Chain Rule**) If f is differentiable at a and g is differentiable at $f(a)$, then the composite function $(g \circ f)(x) = g(f(x))$ is differentiable at a, and its derivative is

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

In functional form, this is

$$\frac{d}{dx}(g \circ f)(x) = \frac{d}{dx} g(f(x)) = g'(f(x))\, f'(x).$$

In words, you differentiate the outer function while holding the inner function fixed, then you differentiate the inner function.

The proof is pretty technical, and you can omit it if you're taking a typical first-term calculus course. It is given at the end. In the examples, I'll focus on how you use the Chain Rule to compute derivatives.

Example. Compute $\dfrac{d}{dx}(x^3+x^2-7x+1)^{99}$.

$(x^3+x^2-7x+1)^{99}$ looks like this:

$$\begin{array}{c}(\qquad\qquad\qquad\qquad)^{99}\\ \uparrow \\ x^3+x^2-7x+1\end{array}$$

Differentiate the outer function $(\text{junk})^{99}$, obtaining $99(\text{junk})^{98}$. What is "junk"? It's x^3+x^2-7x+1. The first term in the Chain Rule is $99(x^3+x^2-7x+1)^{98}$. (Notice that I differentiated the outer function, temporarily leaving the inner one untouched.)

Next, differentiate the inner function. The derivative of x^3+x^2-7x+1 is $3x^2+2x-7$.

Therefore,

$$\frac{d}{dx}(x^3+x^2-7x+1)^{99} = 99(x^3+x^2-7x+1)^{98} \cdot (3x^2+2x-7). \quad \square$$

Example. Compute $\dfrac{d}{dx}\left(\dfrac{1}{x^2-x-1}\right)$.

While it would be correct to use the Quotient Rule, it's unnecessary.

$$\frac{d}{dx}\left(\frac{1}{x^2-x-1}\right) = \underset{\substack{\uparrow \\ \text{the derivative of} \\ \frac{1}{\text{junk}}}}{-\frac{1}{(x^2-x-1)^2}} \cdot \underset{\substack{\uparrow \\ \text{the derivative of} \\ x^2-x-1}}{(2x-1)}$$

That is,

$$\frac{d}{dx}\left(\frac{1}{x^2-x-1}\right) = \left(-\frac{1}{(x^2-x-1)^2}\right)(2x-1).$$

In general, you do not need to use the Quotient Rule to differentiate things of the form

$$\frac{\text{number}}{\text{junk}} \quad \text{or} \quad \frac{\text{junk}}{\text{number}}.$$

In the first case, use the Chain Rule as above. In the second case, divide the top by the number on the bottom. □

Example. Compute $\dfrac{d}{dx}\dfrac{1}{x+5x^3}$.

$$\frac{d}{dx}\frac{1}{x+5x^3} = \frac{d}{dx}(x+5x^3)^{-1} = -(x+5x^3)^{-2}\cdot(1+15x^2). \quad \square$$

In some of the examples which follow, I'll use the derivative formulas for $\sin x$ and $\cos x$. They are:

$$\frac{d}{dx}\sin x = \cos x \quad \text{and} \quad \frac{d}{dx}\cos x = -\sin x.$$

Example. Compute $\dfrac{d}{dx}\sin(x^2)$.

$$\frac{d}{dx}\sin(x^2) = \underset{\substack{\uparrow \\ \text{the derivative of} \\ \sin(\text{junk})}}{\left(\cos(x^2)\right)} \cdot \underset{\substack{\uparrow \\ \text{the derivative of} \\ x^2}}{2x} \quad \square$$

Example. Compute $\dfrac{d}{dx}\cos\dfrac{1}{x^2-2x+5}$.

Differentiating	$\cos(\text{junk})$	gives	$-\sin\dfrac{1}{x^2-2x+5}$
Differentiating	$\dfrac{1}{\text{junk}}$	gives	$-\dfrac{1}{(x^2-2x+5)^2}$
Differentiating	x^2-2x+5	gives	$2x-2$

Therefore,

$$\frac{d}{dx}\cos\frac{1}{x^2-2x+5} = \left(-\sin\frac{1}{x^2-2x+5}\right)\left(-\frac{1}{(x^2-2x+5)^2}\right)(2x-2). \quad \square$$

Example. f and g are differentiable functions. A table of some values for these functions is shown below.

	$x = 3$	$x = 7$
$f(x)$	7	14
$g(x)$	-5	0
$f'(x)$	6	2
$g'(x)$	10	11

Find $(g \circ f)'(3)$.

By the Chain Rule,

$$(g \circ f)'(3) = g'\left(f(3)\right) \cdot f'(3) = g'(7) \cdot f'(3) = 11 \cdot 6 = 66. \quad \square$$

Example. Compute $\dfrac{d}{dx}\sin(\sin x)$.

$$\frac{d}{dx}\sin(\sin x) = [\cos(\sin x)] \cdot \cos x. \quad \square$$

Example. (a) Compute $\dfrac{d}{dx}\left[(\sin x)^2 + \sin(x^2)\right]$.

(b) Draw a picture to show the difference between the functions $(\sin x)^2$ and $\sin(x^2)$, considered as composites of $f(x) = \sin x$ and $g(x) = x^2$.

(a)

$$\frac{d}{dx}\left[(\sin x)^2 + \sin(x^2)\right] = 2(\sin x)(\cos x) + 2x \cdot \cos(x^2). \quad \square$$

(b) Here's a picture showing the difference between $(\sin x)^2$ and $\sin(x^2)$:

$$\begin{array}{ccc} (\quad\quad)^2 & \qquad\qquad & \sin(\quad) \\ \uparrow & & \uparrow \\ \sin x & & x^2 \end{array}$$

In the first case, the outer function is the squaring function; in the second case, the outer function is the sine function. $\square$

Example. The derivative formulas for $\tan x$ and $\cot x$ are

$$\frac{d}{dx}\tan x = (\sec x)^2 \quad \text{and} \quad \frac{d}{dx}\cot x = -(\csc x)^2.$$

Taking these for granted, find:

(a) $\dfrac{d}{dx}\tan\dfrac{1}{x}$.

(b) $\frac{d}{dx}\sqrt{\cot(3x+1)}$.

(a)

$$\frac{d}{dx}\tan\frac{1}{x} = \left(\sec\frac{1}{x}\right)^2 \cdot \left(-\frac{1}{x^2}\right). \quad \square$$

(b)

$$\frac{d}{dx}\sqrt{\cot(3x+1)} = \frac{1}{2}\left(\cot(3x+1)\right)^{-1/2} \cdot \left[-\cot(3x+1)\csc(3x+1)\right] \cdot (3). \quad \square$$

Example. Compute $\frac{d}{dx}\left(1+\left(1+x^2\right)^2\right)^2$.

Differentiate from the outside in:

$$\frac{d}{dx}\left(1+\left(1+x^2\right)^2\right)^2 = 2\left(1+\left(1+x^2\right)^2\right)\cdot 2\left(1+x^2\right)\cdot(2x). \quad \square$$

Example. Where does the graph of $f(x) = (x^2-2x+7)^{-50}$ have a horizontal tangent?

$$f'(x) = (-50)(x^2-2x+7)^{-51}\cdot(2x-2) = \frac{(-50)(2x-2)}{(x^2-2x+7)^{51}}.$$

Set $f'(x) = 0$ and solve for x:

$$\begin{aligned}\frac{(-50)(2x-2)}{(x^2-2x+7)^{51}} &= 0\\ -50(2x-2) &= 0 \quad \square\\ 2x &= 2\\ x &= 1\end{aligned}$$

The proof of the Chain Rule.

This section is fairly technical, so you can probably skip it if you're reading this for first-term calculus.

Lemma. If f is differentiable at a, there is a continuous function $p(h)$ which satisfies:

(a) $\lim_{h\to 0} p(h) = 0$.

(b)

$$f(a+h) - f(a) = (f'(a) + p(h))\cdot h.$$

Proof. Define

$$p(h) = \begin{cases} \dfrac{f(a+h)-f(a)}{h} - f'(a) & \text{if } h \ne 0\\ 0 & \text{if } h = 0\end{cases}$$

Then

$$\begin{aligned}\lim_{h\to 0} p(h) &= \lim_{h\to 0}\left(\frac{f(a+h)-f(a)}{h} - f'(a)\right)\\ &= \lim_{h\to 0}\frac{f(a+h)-f(a)}{h} - f'(a)\\ &= f'(a) - f'(a)\\ &= 0\end{aligned}$$

Thus,

$$\lim_{h \to 0} p(h) = 0 = p(0).$$

Hence, p is a continuous function. This proves (a).

Note that for $h \neq 0$,

$$\begin{aligned}
\frac{f(a+h) - f(a)}{h} - f'(a) &= p(h) \\
\frac{f(a+h) - f(a)}{h} &= f'(a) + p(h) \\
f(a+h) - f(a) &= (f'(a) + p(h)) \cdot h
\end{aligned}$$

For $h = 0$, this equation is true, since both sides are 0. This proves (b). □

Theorem. (**Chain Rule**) Suppose that $f(a)$. Assume f is differentiable at a and g is differentiable at $f(a)$. Then the composite $(g \circ f)(x) = g(f(x))$ is differentiable at a, and

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

Proof. By the lemma, there are functions p and q such that

$$\lim_{h \to 0} p(h) = 0 \quad \text{and} \quad \lim_{k \to 0} q(k) = 0,$$

$$f(a+h) - f(a) = (f'(a) + p(h)) \cdot h,$$

$$g(f(a) + k) - g(f(a)) = (g'(f(a)) + q(k)) \cdot k.$$

Here $k = f(a+h) - f(a)$. Thus, as $h \to 0$, I have $k \to 0$. By the rule for the limit of a composite, this means that as $h \to 0$, I have $q(k) \to 0$.

The next few steps may be a little hard to follow, so I'll give some detail before I do the computation. I will take the equation $g(f(a) + k) - g(f(a)) = (g'(f(a)) + q(k)) \cdot k$ and substitute as follows:

1. On the left side, I'll plug in $k = f(a+h) - f(a)$.
2. On the right side I'll plug in $k = f(a+h) - f(a) = (f'(a) + p(h)) \cdot h$ in for k.

Now here's the computation:

$$\begin{aligned}
g(f(a) + k) - g(f(a)) &= (g'(f(a)) + q(k)) \cdot k \\
g\left(f(a) + f(a+h) - f(a)\right) - g(f(a)) &= (g'(f(a)) + q(k)) \cdot ((f'(a) + p(h)) \cdot h) \\
g(f(a+h)) - g(f(a)) &= (g'(f(a)) + q(k)) \cdot (f'(a) + p(h)) \cdot h \\
\frac{g(f(a+h)) - g(f(a))}{h} &= (g'(f(a)) + q(k)) \cdot (f'(a) + p(h))
\end{aligned}$$

Now take the limit as $h \to 0$ on both sides. Remember that as $h \to 0$, I have both $p(h) \to 0$ and $q(k) \to 0$.

$$(g \circ f)(x)'(a) = \lim_{h \to 0} \frac{g(f(a+h)) - g(f(a))}{h} = \lim_{h \to 0} (g'(f(a)) + q(k)) \cdot (f'(a) + p(h)) = g'(f(a)) \cdot f'(a). \quad \square$$