Chain of functions. On a Ferris wheel, your height H (in feet) depends on the angle θ of the wheel (in radians): $H = 100 + 100\sin(\theta)$. The wheel is turning at one revolution per minute, meaning the angle at t minutes is $\theta = 2\pi t$ radians. At $t = \frac{1}{12}$, we have $\theta = \frac{\pi}{6}$ and:

$$H \;=\; 100 + 100\sin(2\pi t) \;=\; 100 + 100\sin(\tfrac{\pi}{6}) = 150 \text{ ft.}$$

At this moment, how fast are you rising (in ft/min)?

The answer is given by the *Chain Rule*, which computes the derivative for a chain of functional dependencies: one variable H depends on a second variable θ, which depends on a third variable t. The Rule states:

$$\frac{dH}{dt} \;=\; \frac{dH}{d\theta} \cdot \frac{d\theta}{dt}$$

$$\frac{\text{ft}}{\text{min}} \;=\; \frac{\text{ft}}{\text{rad}} \cdot \frac{\text{rad}}{\text{min}}$$

The rate of change of height with respect to angle is:

$$\frac{dH}{d\theta} \;=\; \frac{d}{d\theta}(100 + 100\sin(\theta)) \;=\; 0 + 100\sin'(\theta)$$

$$= \; 100\cos(\theta) \;=\; 100\cos(\tfrac{\pi}{6}) \;\cong\; 86.6\,\frac{\text{ft}}{\text{rad}}.$$

The rate of change of angle with respect to time is:

$$\frac{d\theta}{dt} \;=\; \frac{d}{dt}(2\pi t) \;=\; 2\pi \;\cong\; 6.28\,\frac{\text{rad}}{\text{min}}.$$

Thus, the Chain Rule says the rate of change of height with respect to time is the product:

$$\frac{dH}{dt} \;\cong\; 86.6\frac{\text{ft}}{\text{rad}} \times 6.28\frac{\text{rad}}{\text{min}} \;\cong\; 544\frac{\text{ft}}{\text{min}}.$$

Your rate of rise is about 544 feet per minute, at time $t = \frac{1}{12}$.

Chain Rule: Let y, u, x be variables related by $y = f(u)$ and $u = g(x)$, so that $y = f(g(x))$. Then, in Leibnitz notation:

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

or in Newton notation:

$$f(g(x))' \;=\; f'(g(x)) \cdot g'(x).$$

This holds at any value of x where $g'(x)$ and $f'(g(x))$ are both defined.

Notes by Peter Magyar `magyar@math.msu.edu`

The function $f(g(x))$ is called the *composition* of f following g, sometimes denoted $f \circ g$, so that we may write $f(g(x))'$ as $(f \circ g)'(x)$.

Proof.[*] First we assume that the value $g(a)$ is different from all other nearby output values $g(x)$: that is, for x close enough (but unequal) to a, we have $g(x) \neq g(a)$. Then we compute, using the alternative definition of derivative:

$$\begin{aligned}
(f \circ g)'(a) \;&=\; \tfrac{d}{dx} f(g(x))\Big|_{x=a} \;=\; \lim_{x \to a} \frac{f(g(x)) - f(g(a))}{x - a} \\
&= \lim_{x \to a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \cdot \lim_{u \to a} \frac{g(x) - g(a)}{x - a} \\
&= \lim_{u \to g(a)} \frac{f(u) - f(g(a))}{u - g(a)} \cdot \lim_{u \to a} \frac{g(x) - g(a)}{x - a} \\
&= f'(g(a)) \cdot g'(a)\,.
\end{aligned}$$

Here we used the Limit Substitution Theorem from Notes §1.7, substituting u for $g(x)$ so that $x \to a$ forces $u \to g(a)$. (Since $g(x)$ is differentiable at $x = a$, it is also continuous.)

Finally, if there is a sequence of inputs $x_1, x_2, \cdots \to a$ with $g(x_i) = g(a)$, then we clearly have $g'(a) = 0$, and the right side of our formula becomes $f'(g(a)) \cdot g'(a) = 0$. On the left side, we have values $(f(g(x_i)) - f(g(a)))/(x_i - a) = 0$, which is consistent with the desired limit $(f \circ g)'(a) = 0$, and the previous argument is still valid when restricted to the set of x with $g(x) \neq g(a)$.

Differentiation Rules. Along with our previous Derivative Rules from Notes §2.3, and the Basic Derivatives from Notes §2.3 and §2.4, the Chain Rule is the last fact needed to compute the derivative of any function defined by a formula.

EXAMPLE: Find the derivative of $(x+\frac{1}{x})^{10}$. First, we use Leibnitz notation: let $y = u^{10}$ and $u = x + \frac{1}{x}$, so that $y = (x+\frac{1}{x})^{10}$. Then:

$$\begin{aligned}
\frac{dy}{dx} \;&=\; \frac{dy}{du} \cdot \frac{du}{dx} \;=\; \frac{d}{du}(u^{10}) \cdot \frac{d}{dx}\left(x+\tfrac{1}{x}\right) \;=\; 10u^9 \cdot \frac{d}{dx}\left(x+x^{-1}\right) \\
&= 10\big(x+\tfrac{1}{x}\big)^9 \cdot \big(1+(-1x^{-2})\big) \;=\; 10\big(x+\tfrac{1}{x}\big)^9\big(1-\tfrac{1}{x^2}\big)\,.
\end{aligned}$$

Next, we redo this in Newton notation, without introducing new letters y, u. Let $f(x) = x^{10}$ with $f'(x) = 10x^9$, and $g(x) = x + \frac{1}{x} = x + x^{-1}$ with $g'(x) = 1 - x^{-2} = 1 - \frac{1}{x^2}$, so that:

$$f(g(x))' \;=\; f'(g(x)) \cdot g'(x) \;=\; 10\big(x+\tfrac{1}{x}\big)^9\big(1-\tfrac{1}{x^2}\big)\,.$$

A third way (the quickest in practice) is to think of the composite function as an outside function $out = (\;\;)^{10}$ wrapped around an inside function $in = x + \frac{1}{x}$, so the Chain Rule becomes:

$$out(in)' \;=\; out'(in) \cdot in'$$

[*] For another proof based on linear approximations, see the Stewart text §2.5, p. 153.

Here $out' = 10(\;\;)^9$, so:

$$out(in)' \;=\; 10\big(x+\tfrac{1}{x}\big)^9 \cdot \big(x+\tfrac{1}{x}\big)' \;=\; 10\big(x+\tfrac{1}{x}\big)^9 \cdot \big(1-\tfrac{1}{x^2}\big)$$

EXAMPLE: For any function $u = g(x)$, and any number n, we have:

$$\tfrac{d}{dx}(u^n) = nu^{n-1}\,\tfrac{du}{dx} \qquad \text{and} \qquad (g(x)^n)' = n\,g(x)^{n-1}g'(x).$$

EXAMPLE: Find the derivative of $\frac{1}{\sqrt{x\cos(x)}}$. Here the outer function is $out = \frac{1}{\sqrt{\;\;}} = (\;\;)^{-1/2}$ with $out' = -\frac{1}{2}(\;\;)^{-3/2}$. Thus:

$$\left(\frac{1}{\sqrt{x\cos(x)}}\right)' \;=\; -\tfrac{1}{2}(x\cos(x))^{-3/2} \cdot (x\cos(x))'$$

$$=\; -\tfrac{1}{2}(x\cos(x))^{-3/2} \cdot ((x)'\cos(x) + x\cos'(x)) \;=\; -\tfrac{1}{2}(x\cos(x))^{-3/2}(\cos(x) - x\sin(x))$$

Here we used the Chain Rule, then the Product Rule.

EXAMPLE: Compare the derivatives of $\sin(x^2)$ and $\sin^2(x)$. Note that if $f(x) = \sin(x)$ and $g(x) = x^2$, we have $\sin(x^2) = f(g(x))$, but $\sin^2(x) = g(f(x))$. Thus:

$$\big(\sin(x^2)\big)' \;=\; \sin'(x^2)\cdot(x^2)' \;=\; \cos(x^2)\cdot 2x \;=\; 2x\cos(x^2)$$

$$\big(\sin^2(x)\big)' \;=\; \big((\sin(x))^2\big)' \;=\; 2(\sin(x))\cdot\sin'(x) \;=\; 2\sin(x)\cos(x)$$

EXAMPLE: Find the derivative of $\sin\Big(\tan\Big(\frac{x}{x+1}\Big)\Big)$, a composition of three functions. We start by applying the Chain Rule to the outermost function $\sin(\;\;)$, with inner function $\tan\Big(\frac{x}{x+1}\Big)$; then we use the Chain Rule again on this.

$$\Big(\sin\Big(\tan\Big(\tfrac{x}{x+1}\Big)\Big)\Big)' \;=\; \sin'\Big(\tan\Big(\tfrac{x}{x+1}\Big)\Big)\cdot\Big(\tan\Big(\tfrac{x}{x+1}\Big)\Big)'$$

$$=\; \sin'\Big(\tan\Big(\tfrac{x}{x+1}\Big)\Big)\cdot\tan'\Big(\tfrac{x}{x+1}\Big)\cdot\Big(\tfrac{x}{x+1}\Big)'$$

$$=\; \sin'\Big(\tan\Big(\tfrac{x}{x+1}\Big)\Big)\cdot\tan'\Big(\tfrac{x}{x+1}\Big)\cdot\tfrac{(x)'(x+1)-x(x+1)'}{(x+1)^2}$$

$$=\; \cos\Big(\tan\Big(\tfrac{x}{x+1}\Big)\Big)\cdot\sec^2\Big(\tfrac{x}{x+1}\Big)\cdot\tfrac{(x+1)-x}{(x+1)^2}$$

The last factor uses the Quotient Rule.

EXAMPLE: What if we apply the Chain Rule to a complicated constant like π^3, where we consider x^3 as the outside function and the constant function $p(x) = \pi$ as the inside? Then:

$$\big(\pi^3\big)' \;=\; 3\pi^2\cdot(\pi)' \;=\; 3\pi^2\cdot 0 \;=\; 0,$$

since $(\pi)' = c' = 0$. *Any* expression with no variable in it is constant, with derivative *zero.*

Degrees versus radians. In higher mathematics, we always use radian measure (full circle $= 2\pi$ radians)†, so that $\sin(x)$ always means sine of x radians. This is essential to get the formula $\sin'(x) = \cos(x)$.

The sine with input x in degrees (full circle $= 360$ deg) is actually a different function, which we can denote as $\sin_{deg}(x)$. Remember that a function is a rule which converts input numbers to output numbers: it does not know that we interpret some numbers as angles, or what their units should be. Since $\sin(x)$ and $\sin_{deg}(x)$ produce different outputs from a given number x, they are different functions. In fact, we have:

$$\sin_{deg}(x) = \sin\big(\tfrac{2\pi}{360}x\big)\,.$$

The inside operation converts x from degrees to radians, then feeds this into the ordinary (radian) sine function.

This makes a crucial difference in the derivative:

$$\sin'_{deg}(x) \;=\; \big(\sin\big(\tfrac{2\pi}{360}x\big)\big)' \;=\; \cos\big(\tfrac{2\pi}{360}x\big)\cdot\big(\tfrac{2\pi}{360}x\big)' \;=\; \cos_{deg}(x)\cdot\big(\tfrac{2\pi}{360}\big)\,.$$

This is why we stay away from degree measure in calculus!

†The geometric definition is that an angle of x radians spans an arc of length x radius-lengths, or xr. Thus, 2π radians spans an arc of length $2\pi r$, meaning the circumference of the full circle.