Introducing the chain rule

Consider having a gas in a cylinder with an adjustable piston so that we can control or measure quantities such as volume V, pressure p, temperature T, and the number of gas atoms/molecules n. Each of these quantities has units as given in the table below.

Quantity	Units
volume V	liter (L)
pressure p	atmosphere (atm)
temperature T	Kelvin (K)
number n	mol

We'll assume that these quantities are related by the *ideal gas law*

$$pV = nRT$$

where R is a constant with value $R = 0.08206 \, \frac{\text{L}\cdot\text{atm}}{\text{mol}\cdot\text{K}}$.

Let's consider a specific situation in which we keep the number of gas particles constant at $n = 1$ mol and the temperature constant at $T = 300$ K. We can then think of volume as a function of pressure. We can solve to get

$$V = nRT \cdot \frac{1}{p}$$

where nRT is a constant in the context we've set up.

To get the rate of change in volume with respect to pressure, we compute the derivative dV/dp and get

$$\frac{dV}{dp} = -nRT \cdot \frac{1}{p^2}.$$

We can input a specific value of pressure to get a specific rate of change. For example, at pressure $p = 1.09$ atm, we get

$$\left.\frac{dV}{dp}\right|_{p=1.09\text{ atm}} = -(1\text{ mol})(0.08206\,\frac{\text{L}\cdot\text{atm}}{\text{mol}\cdot\text{K}})(300\,\text{K})\frac{1}{(1.09\,\text{atm})^2} = -20.72\frac{\text{L}}{\text{atm}}$$

Now, let's introduce time t as a new variable with t measured in minutes (min). Imagine that we control pressure over time. To be specific, suppose we control pressure so that

$$p = p_0 + at^2$$

with constants $p_0 = 1$ atm and $a = 0.01\,\frac{\text{atm}}{\text{min}^2}$. To get the rate of change in pressure with respect to time, we compute

$$\frac{dp}{dt} = 0 + a(2t) = 2at.$$

Now focus on a specific time, say $t = 3$ min. For $t = 3$ min, the pressure has the value

$$p = 1\,\text{atm} + (0.01\frac{\text{atm}}{\text{min}^2})(3\,\text{min})^2 = 1.09\,\text{atm}.$$

(Note that this is the value of pressure we used above. This is not a coincidence.) At this specific time, the pressure is changing with respect to time at the rate

$$\left.\frac{dp}{dt}\right|_{t=3\text{ min}} = 2(0.01\frac{\text{atm}}{\text{min}^2})(3\,\text{min}) = 0.06\frac{\text{atm}}{\text{min}}.$$

Here's the main question: How do we combine

$$\left.\frac{dV}{dp}\right|_{p=1.09\text{ atm}} = -20.72\frac{\text{L}}{\text{atm}} \qquad \text{and} \qquad \left.\frac{dp}{dt}\right|_{t=3\text{ min}} = 0.06\frac{\text{atm}}{\text{min}}$$

to get the rate of change in volume *with respect to time* at $t = 3$ min? In other words, how do we get

$$\left.\frac{dV}{dt}\right|_{t=3\text{ min}}$$

from

$$\left.\frac{dV}{dp}\right|_{p=1.09\text{ atm}} = -20.72\frac{\text{L}}{\text{atm}} \qquad \text{and} \qquad \left.\frac{dp}{dt}\right|_{t=3\text{ min}} = 0.06\frac{\text{atm}}{\text{min}}$$

Looking at units makes it reasonable to conjecture that we should *multiply* to get

$$\left.\frac{dV}{dt}\right|_{t=3\text{ min}} = \left.\frac{dV}{dp}\right|_{p=1.09\text{ atm}} \cdot \left.\frac{dp}{dt}\right|_{t=3\text{ min}} = -20.72\frac{\text{L}}{\text{atm}} \cdot 0.06\frac{\text{atm}}{\text{min}} = -1.2432\frac{\text{L}}{\text{min}}.$$

This is, in fact, correct. We'll justify this more completely later.

The rule that says

$$\frac{dV}{dt} = \frac{dV}{dp}\frac{dp}{dt}$$

is called the *chain rule*. For our example, we have

$$\frac{dV}{dt} = \frac{dV}{dp}\frac{dp}{dt} = -nRT\frac{1}{p^2} \cdot 2at.$$

To get the specific value for $t = 3$ min, we need to substitute in $t = 3$ min *and* we need to substitute in $p = 1.09$ atm since that is the value of pressure for $t = 3$ min. Putting in these values (and the values of all our constants), we have

$$\left.\frac{dV}{dt}\right|_{t=3\text{ min}} = -(1\text{ mol})(0.08206\,\frac{\text{L}\cdot\text{atm}}{\text{mol}\cdot\text{K}})(300\,\text{K})\frac{1}{(1.09\,\text{atm})^2}(2)(0.01\frac{\text{atm}}{\text{min}^2})(3\,\text{min}) = -1.2432\frac{\text{L}}{\text{min}}$$

In mathematical terms, the chain rule is used for *differentiating a composition*. Let's express the chain rule in the "prime" notation for dervivatives. First, use function notation to write

$$V(p) = nRT \cdot \frac{1}{p} \qquad \text{and} \qquad p(t) = p_0 + at^2$$

To write volume V as a function of time t, we compose these to get

$$(V \circ p)(t) = V(p(t)) = nRT \cdot \frac{1}{p(t)} = nRT \cdot \frac{1}{p_0 + at^2}.$$

The chain rule tells us that to differentiate the composition $V \circ p$, we *multiply* the derivatives of V and p. Specifically,

$$(V \circ p)'(t) = V'(p(t)) \cdot p'(t).$$

If you are not comfortable with the notation $V \circ p$, you could express the same rule as

$$\frac{d}{dt}\big[V(p(t))\big] = V'(p(t)) \cdot p'(t).$$

Our first example is complicated by the fact that we had to keep track of units in order to let units guide us to a conjecture. Let's forget units and look at an example without context. Suppose we want to differentiate $f(x) = \sin(x^2)$. We can think of this as a composition of the functions $u = \sin(v)$ and $v = x^2$. So $f(x) = u(v(x))$ and the derivative of f is

$$\frac{df}{dx} = \frac{du}{dv}\frac{dv}{dx} = \cos(v) \cdot 2x = 2x\cos(v).$$

As a final step, we can substitute $v = x^2$ to get

$$\frac{df}{dx} = 2x\cos(x^2).$$

Here's the same calculation using the "prime" notation:

$$f'(x) = u'(v(x)) \cdot v'(x) = \cos(x^2) \cdot 2x = 2x\cos(x^2).$$

Here are two key things to note about the chain rule

- We multiply the derivatives of u and v.
- The derivative of u is evaluated at $v(x)$.