2.4 The chain rule

Leibniz notation for derivatives

In this section, we wish to understand how to differentiate chains of functions, like $\sin(x^2)$ (which equals $f(g(x))$, with $f(x) = \sin(x)$ and $g(x) = x^2$), $\sqrt{3 + \tan(t)}$ (which is the chain of $f(t) = \sqrt{t}$ with $g(t) = 3 + \tan(t)$), and so on. To do so, it will be useful to first introduce yet another way of writing derivatives.

Definition 2.4.1. Leibniz notation for the derivative. If $y = f(x)$ is a differentiable function, then we write

$$\frac{dy}{dx}$$

for the derivative $f'(x)$.

For example:

$$\text{If} \quad y = x^9 + x^{3/8} - \frac{\tan(x)}{5}, \quad \text{then} \quad \frac{dy}{dx} = 9x^8 + \frac{3}{8}x^{-5/8} - \frac{\sec^2(x)}{5};$$

$$\text{If} \quad r = 5^q + q^5, \quad \text{then} \quad \frac{dr}{dq} = \ln(5)5^q + 5q^4,$$

etc.

So if $y = f(x)$ is differentiable, then $f'(x)$, $\frac{d}{dx}[f(x)]$, and $\frac{dy}{dx}$ all mean the same thing.

A couple of observations are worth making:

1. $\frac{dy}{dx}$ **is not** a fraction; it's a derivative. BUT:

2. $\frac{dy}{dx}$ **is** *a limit of fractions*, since, by definition of the derivative,

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}. \tag{2.4.1}$$

(Here as usual, if $y = f(x)$, then Δy denotes $f(x + \Delta x) - f(x)$.)

One might think of formula (2.4.1), informally, as saying "As $\Delta x \to 0$, the Δ's become d's." This is informal because it's only true as far as the above *notation* goes. It's not true about the actual quantities Δx, Δy, dx, and dy. The latter two aren't even quantities *per se*; they're just pieces of the symbol $\frac{dy}{dx}$. Still, this way of thinking is suggestive. In particular, as we will see in this section, it's quite useful in understanding derivatives of chains.

The chain rule, first version

In Section 2.3, we saw that the derivative of a sum is the sum of the derivatives (the sum rule), and that the derivative of a constant multiple is the corresponding constant multiple of the derivative (the constant multiple rule). Now what about chains, or compositions? Is the derivative of a chain equal to the chain of the derivatives? As the following example shows, the answer is *no*, though there *is* a nice "algebra" to differentiation of chains.

Example 2.4.1. At a particular point in time on the elliptic trainer, the monitor tells you that you are burning 0.31 calories (cal) per step, and are climbing 40 steps per minute (min). What is your time rate of energy expenditure at this point in time, in calories per minute?

Solution. We can answer by "following the units." That is: calories per step times steps per minute equals calories per minute. So the rate of energy expenditure, in calories per minute, equals

$$0.31\,\frac{\text{cal}}{\text{step}} \times 40\,\frac{\text{step}}{\text{min}} = 0.31 \times 40\,\frac{\text{cal}}{\text{min}} = 12.4\,\frac{\text{cal}}{\text{min}}.$$

Note that, in the above example, we *are* considering a chain of functions: the number y of calories burned is a function of the number u of steps taken, and u is a function of time x, so y itself is, ultimately, also a function of x.

Moreover, the above example illustrates the fact that *the rate of change of y with respect to x equals the rate of change of y with respect to u* ***times*** *the rate of change of u with respect to x.* This can be summarized quite nicely in Leibniz notation:

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

The chain rule, first version

Very roughly speaking, the chain rule says that "the derivative of the chain equals the product of the derivatives." However, one needs to be a bit careful about such an interpretation. We'll discuss this further in the next subsection, after introduction of our second version of the chain rule.

Note that we haven't actually proved the chain rule. We hope that the above example makes it believable, but those who still need convincing might consider the following. First of all, we have a simple formula for *average* rates of change:

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u}\frac{\Delta u}{\Delta x}. \tag{2.4.2}$$

(This is true just by algebra: cancelling a Δu top and bottom on the right-hand side gives the left-hand side.) But if we apply, to equation (2.4.2), the philosophy that "Δ's become d's as $\Delta x \to 0$," then we get the above chain rule exactly. (This argument is not completely rigorous. But it could be made so with just a few additional details.)

Here are some further examples.

Example 2.4.2. (a) Find:

(i) $\dfrac{dy}{dx}$ if $y = \cos(u)$ and $u = \sin(x)$

(ii) $\dfrac{dz}{dw}$ if $z = 3^v$ and $v = w^2 - 7w$

(iii) $\dfrac{dy}{dx}$ if $y = \sin(x^2)$

(iv) $\dfrac{d}{dt}\left[\sqrt{3 + \tan(t)}\right]$

(b) A spherical snowball is melting, in such a way that its radius is decreasing at 0.75 centimeters (cm) per minute (min). How fast is the surface area of the snowball melting when the radius is 10 cm?

Solution. (a)(i) We have

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{d}{du}[\cos(u)] \cdot \frac{d}{dx}[\sin(x)] = -\sin(u) \cdot \cos(x).$$

Now we are asking for a derivative with respect to x, so our answer should be in terms of x. Since $u = \sin(x)$, the above result gives us

$$\frac{dy}{dx} = -\sin(\sin(x))\cos(x).$$

(ii)
$$\frac{dz}{dw} = \frac{dz}{dv}\frac{dv}{dw} = \frac{d}{dv}[3^v] \cdot \frac{d}{dw}[w^2 - 7w] = \ln(3)3^v \cdot (2w - 7) = \ln(3)(2w - 7)3^{w^2 - 7w}.$$

(iii) We are not explicitly given a "function u in the middle," so we need to introduce one. That is, we write $y = \sin(u)$ where $u = x^2$. Then

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{d}{du}[\sin(u)] \cdot \frac{d}{dx}[x^2] = \cos(u) \cdot 2x = 2x\cos(x^2).$$

(iv) Let $y = \sqrt{3 + \tan(t)}$: we write $y = \sqrt{u}$ where $u = 3 + \tan(t)$. Then

$$\frac{d}{dt}\left[\sqrt{3 + \tan(t)}\right] = \frac{dy}{dt} = \frac{dy}{du}\frac{du}{dt} = \frac{d}{du}\left[\sqrt{u}\right] \cdot \frac{d}{dt}[3 + \tan(t)] = \frac{1}{2\sqrt{u}} \cdot (0 + \sec^2(t)) = \frac{\sec^2(t)}{2\sqrt{3 + \tan(t)}}.$$

(b) The surface area S of a sphere is given in terms of its radius r by the formula $S = 4\pi r^2$.

If we were interested in the rate of change of S with respect to r, this would be easy: the power formula and the constant multiple rule would give us $dS/dr = d[4\pi r^2]/dr = 4\pi \cdot 2r = 8\pi r$. But because we want the rate of change of S *with respect to time* t (as is indicated by the phrase "how fast"), there is an extra step.

Specifically, the chain rule gives

$$\frac{dS}{dt} = \frac{dS}{dr}\frac{dr}{dt} = 8\pi r\frac{dr}{dt}. \tag{2.4.3}$$

We are given that $dr/dt = 0.75$ cm/min, so when $r = 10$ cm we have, by equation (2.4.3),

$$\frac{dS}{dt} = 8\pi \times 10 \text{ cm} \times 0.75 \ \frac{\text{cm}}{\text{min}} = 188.496 \ \frac{\text{cm}^2}{\text{min}}.$$

Note that the units work out: square centimeters per minute are the correct units for a rate of change of area (in centimeters) with respect to time (in minutes).

The chain rule, second version

The situation embodied by parts (a)(iii) and (a)(iv) of Example 2.4.2 is typical, in that the chains there were given to us in "final" form. That is, the two simpler functions being "chained together" to get the more complex one, whose derivative we wanted, were not specified explicitly. Our next version of the chain rule allows us to differentiate chains like this, without having to *explicitly write down* an auxiliary variable like u.

To see how this works, let $y = f(g(x))$. We can write $y = f(u)$ where $u = g(x)$. Then by the above chain rule,

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = f'(u)g'(x). \tag{2.4.4}$$

But again, $y = f(g(x))$ and $u = g(x)$; substituting these facts into equation (2.4.4) gives

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

The chain rule, second version

Note that, although we needed a variable u to *arrive at* this version of the chain rule, no such variable appears explicitly in the final result.

This second version of the chain rule says, roughly: to differentiate a chain, follow these two steps.

Step 1. Differentiate the outer function, and into this derivative, substitute the inner function.

Step 2. Multiply your result by the derivative of the inner function.

Steps for using the chain rule, second version

Example 2.4.3. Find:

(a) $h'(x)$ if $h(x) = (x^2 + 45x)^{37}$.

(b) $\dfrac{d}{dz}[\cos(3z)]$.

Solution. (a) [Implicitly, we are thinking that our "outer" function is the 37th power function (that is, our outer function is $f(u) = u^{37}$), and our "inner" function is $g(x) = x^2 + 45x$. The derivative of the outer function is 37 times the 36th power function (that is, $f'(u) = 37u^{36}$), while the derivative of the inner function is given by $g'(x) = 2x + 45$. So by the above chain rule – or, equivalently, by the two-step process described below it – we have the following.]

$$\begin{aligned} h'(x) &= [\text{ the derivative of the other function, evaluated at the inner function}] \\ &\quad \times [\text{the derivative of the inner function}] \\ &= 37(x^2 + 45)^{36}(2x + 45). \end{aligned}$$

(b) [Our outer function is the cosine function, whose derivative is minus the sine function. Our inner function is $3z$. So the chain rule gives us the following.]

$$\frac{d}{dz}[\cos(3z)] = -\sin(3z)\frac{d}{dz}[3z] = -3\sin(3z).$$

In the above example, everything in square-brackets (both in the preambles to the computations and in the computations themselves) is meant implicitly; it indicates the thought processes involved, but would generally not need to be written out *per se.* This is the point of our second version of the chain rule: it allows us to do a certain amount of "bookkeeping" – identifying inner and outer functions – in our head, rather than on paper.

The following examples further illustrate the idea. (You should do the mental bookkeeping of identifying outer and inner functions.)

Example 2.4.4. Find:

(a) $\dfrac{d}{dz}[\sin(\sin(z))]$.

(b) $H'(1)$ if $H(t) = Q(2t^2 + 3t)$ and $Q'(5) = 4$.

(c) $\dfrac{d}{dz}[\sin(\sin(\sin(z))]$.

Solution. (a) $\dfrac{d}{dz}[\sin(\sin(z))] = \cos(\sin(z))\dfrac{d}{dz}[\sin(z)] = \cos(\sin(z))\cos(z)$.

(b) We differentiate first, and *then* evaluate at the particular point $t = 1$. By the chain rule, we have

$$H'(t) = Q'(2t^2 + 3t)\frac{d}{dt}[2t^2 + 3t] = Q'(2t^2 + 3t) \cdot (4t + 3).$$

Substituting $t = 1$ then gives

$$H'(1) = Q'(2 \times 1^2 + 3 \times 1) \cdot (4 \times 1 + 3) = Q'(5) \cdot 7 = 4 \times 7 = 28.$$

(c) Applying the chain rule twice in succession gives

$$\begin{aligned}\frac{d}{dz}[\sin(\sin(\sin(z))] &= \cos(\sin(\sin(z))\frac{d}{dz}[\sin(\sin(z)]] = \cos(\sin(\sin(z))\cos(\sin(z))\frac{d}{dz}[\sin(z)] \\ &= \cos(\sin(\sin(z))\cos(\sin(z))\sin(z).\end{aligned}$$

(In part (c), we first thought of our outer function as the sine function, and our inner function as $\sin(\sin(z))$. Of course, Step 2 of the chain rule tells us that we need to differentiate this inner function; to do that, we applied the chain rule *again*, since the inner function is itself a chain.)

As part (c) of the above example illustrates, our second version of the chain rule is particularly convenient when we have chains of chains (or chains of chains of chains, and so on). Had we attempted this same exercise using our first version of the chain rule, we would have needed to introduce two auxiliary variables: that is, we would have had to write, say, $y = \sin(u)$ where $u = \sin(v)$ and $v = \sin(x)$. The second version of the chain rule allows us to do this kind of "unchaining" behind the scenes.

Warning. We noted in the previous subsection that, in a sense, the chain rule expresses the derivative of a chain as a product of derivatives. This is true, but only if properly interpreted. That is: the chain rule does *not* say that the derivative of $f(g(x))$ is $f'(x)$ times $g'(x)$. It *does* say that the derivative of $f(g(x))$ is $f'(g(x))$ times $g'(x)$. So: "the derivative of a chain equals the product of the derivatives, as long as the latter two are evaluated at the appropriate, distinct places."

Exercises

Part 1: Chain rule, first version

For these exercises, you should refer back to the subsection "The chain rule, first version" above, and especially Example 2.4.2.

1. Use the chain rule to find dy/dx, when y is given as a function of x in the following way.

(a) $y = 5u - 3$, where $u = 4 - 7x$.

(b) $y = \sin u$, where $u = 4 - 7x$.

(c) $y = \tan u$, where $u = x^3$.

(d) $y = 10^u$, where $u = x^2$.

(e) $y = u^4$, where $u = x^3 + 5$.

2. A cube of sidelength ℓ has surface area $S = 6\ell^2$.

When the sides of a cube are 5 inches, its surface area is changing at the rate of 60 square inches per inch increase in the side. (That is: when $\ell = 5$, $dS/d\ell = 60$ square inches per inch.) If, at that moment, the sides are increasing at a rate of 3 inches per hour, how fast is the surface area increasing: is it 60, 3, 63, 20, 180, 5, or 15 square inches per hour?

3. An explorer is marooned on an iceberg. The top of the iceberg is shaped like a square with sides of length 100 feet. The length of the sides is shrinking at the rate of two feet per day. How fast is the area of the top of the iceberg shrinking? Assuming the sides continue to shrink at the rate of two feet per day, what will be the dimensions of the top of the iceberg in five days? How fast will the area of the top of the iceberg be shrinking then?

4. Suppose the iceberg of problem 3 is shaped like a cube. How fast is the volume of the cube shrinking when the sides have length 100 feet? How fast after five days? (A cube of sidelength ℓ has volume $V = \ell^3$.)

5. (Note: only parts (c) and (d) of this exercise involve the chain rule.) If the radius of a spherical balloon is r inches, its volume V is $\frac{4}{3}\pi r^3$ cubic inches.

(a) Find the rate of change dV/dr of V with respect to r.

(b) At what rate does the volume increase, in cubic inches per inch, when the radius is 4 inches?

(c) Use the chain rule to express the rate of change of volume V, *with respect to time* t, in terms of the radius r, and the rate of change dr/dt of the radius with respect to time.

(d) Suppose someone is inflating the balloon at the rate of 10 cubic inches of air per second. If the radius is 4 inches, at what rate is it increasing, in inches per second?

Part 2: Chain rule, second version

For these exercises, you should refer back to the subsection "The chain rule, second version" above, and especially Examples 2.4.3 and 2.4.4.

6. Find the derivatives of the following functions.

(a) $F(x) = (9x + 6x^3)^5$.

(b) $G(w) = \sqrt{4w^2 + 1}$.

(c) $R(x) = \dfrac{1}{1-x}$. Hint: $\dfrac{1}{1-x} = (1-x)^{-1}$.

(d) $D(z) = 3\tan\left(\dfrac{1}{z}\right)$.

(e) $\text{pig}(t) = \cos(2^t)$.

(f) $\text{wombat}(x) = 5^{1/x}$.

7. Apply the chain rule more than once to find each of the following derivatives. (See Example 2.4.4(c).)

(a) $\text{dog}(w) = \sin^2(w^3 + 1)$. Hint: first write $\sin^2(w^3 + 1) = (\sin(w^3 + 1))^2$.

(b) $q(y) = \tan(\sin(\cos(y)))$.

(c) $R(x) = 3x + (x^2 + (7x^3 + 5)^2)^3$.

8. Let $S(w) = \sqrt{(4w^2 + 1)^3}$. Find $S'(w)$ in two ways:

(i) Write $\sqrt{(4w^2 + 1)^3} = (4w^2 + 1)^{3/2}$ and use the chain rule once.

(ii) Write $\sqrt{(4w^2 + 1)^3} = ((4w^2 + 1)^3)^{1/2}$ and use the chain rule twice.

Make sure to show that your two answers are the same.

9. Let $f(t) = t^2 + 2t$ and $g(t) = 5t^3 - 3$. Determine all of the following: $f'(t)$, $g'(t)$, $g(f(t))$, $f(g(t))$, $g'(f(t))$, $f'(g(t))$, $\frac{d}{dt}[f(g(t))]$, $\frac{d}{dt}[g(f(t))]$.

Part 3: Particular values

For the following exercises remember that, to evaluate a derivative at a point, you differentiate first, and *then* plug in the point in question. See Example 2.4.4(b) above.

10. If $h(x) = (f(x))^6$ where f is some function satisfying $f(93) = 2$ and $f'(93) = -4$, what is $h'(93)$?

11. If $H(x) = F(x^2 - 4x + 2)$ where F is some function satisfying $F'(2) = 3$, what is $H'(4)$?

12. If $f(x) = (1 + x^2)^5$, what are the numerical values of $f'(0)$ and $f'(1)$?

13. If $h(t) = \cos(\sin t)$, what are the numerical values of $h'(0)$ and $h'(\pi)$?

Part 3: Miscellaneous

14. (a) What is the derivative of $f(x) = 2^{-x^2}$?

(b) Sketch the graphs of f and its derivative on the interval $-2 \leq x \leq 2$.

(c) For what values(s) of x is $f'(x) = 0$? What is true about the graph of f at the corresponding points?

(d) Where does the graph of f have positive slope, and where does it have negative slope?

15. (a) With a graphing utility, find the point x where the function $y = 1/(3x^2 - 5x + 7)$ takes its maximum value. Obtain the numerical value of x accurately to two decimal places.

(b) Find the derivative of $y = 1/(3x^2 - 5x + 7)$, and determine where it takes the value 0.

[Answer: $y' = -(6x - 5)(3x^2 - 5x + 7)^{-2}$, and $y' = 0$ when $x = 5/6$.]

(c) Using part (b), find the *exact* value of x where $y = 1/(3x^2 - 5x + 7)$ takes its maximum value.

(d) At what point is the graph of $y = 1/(3x^2 - 5x + 7)$ rising most steeply? Describe how you determined the location of this point.

16. (a) Find a function $f(x)$ for which $f'(x) = 3x^2(5 + x^3)^{10}$. A useful way to proceed is to guess. For instance, you might guess $f(x) = (5 + x^3)^{11}$. Differentiate $f(x)$ and see what you get; see if you can use this information to modify $f(x)$, to get the correct answer.

(b) Find a function $p(x)$ for which $p'(x) = x^2(5 + x^3)^{10}$.

17. Find a function $g(t)$ for which $g'(t) = t/\sqrt{1 + t^2}$. Hint: guess $g(t) = \sqrt{1 + t^2}$. As in Exercise 16 above, differentiate $g(t)$, and then modify your guess to correct it.

18. Find a function $h(x)$ for which $h'(x) = x7^{x^2}$. Hint: guess $h(x) = 7^{x^2}$, then proceed as in the previous two exercises.