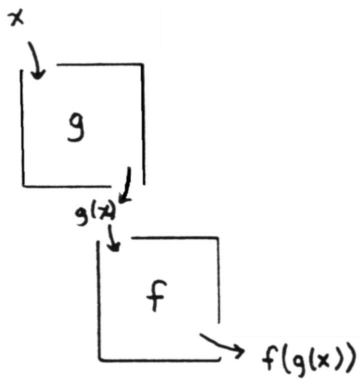
4.5 The Chain Rule
(Differentiating Composite Functions)

composite functions; review

Let f and g be functions of x. Recall that the composition $f \circ g$ is defined by

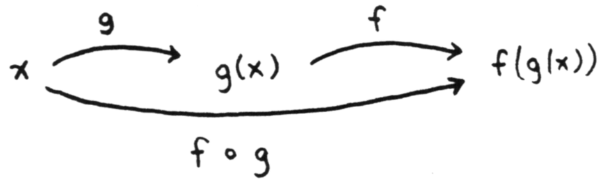
$$(f \circ g)(x) := f(g(x))$$



and has domain:

$$\mathcal{D}(f \circ g) = \{x \,|\, x \in \mathcal{D}(g) \text{ and } g(x) \in \mathcal{D}(f)\}$$

One reads $f \circ g$ as 'f *circle* g' or 'f *composed with* g'.

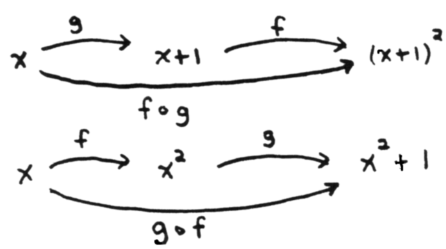


EXAMPLE

Let $f(x) = x^2$ and $g(x) = x + 1$. Find $f \circ g$ and $g \circ f$.

$$(f \circ g)(x) := f(g(x)) = f(x+1) = (x+1)^2 = x^2 + 2x + 1$$
$$(g \circ f)(x) := g(f(x)) = g(x^2) = x^2 + 1$$

Here are the corresponding mapping diagrams:



composition of functions is NOT, in general, commutative!

Observe that for these functions f and g, $f \circ g \neq g \circ f$. Indeed, the only value of x for which $f(g(x)) = g(f(x))$ is $x = 0$ (♣ check this).

Thus, composition of functions is *not*, in general, commutative!

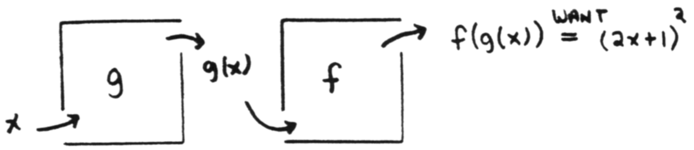
differentiating composite functions; introduction

If f and g are *differentiable*, then it would be reasonable to hope that $f \circ g$ is also differentiable. This is the idea investigated in this section.

EXERCISE 1

♣ 1. For $f(x) = x^2 - 1$ and $g(x) = -2x$, find both $f \circ g$ and $g \circ f$.

♣ 2. View the function $h(x) = (2x+1)^2$ as a composition of functions. That is, find functions f and g for which $h = f \circ g$. (There is not a unique answer.) To do this, appropriately 'name' the function boxes below.



a motivating example

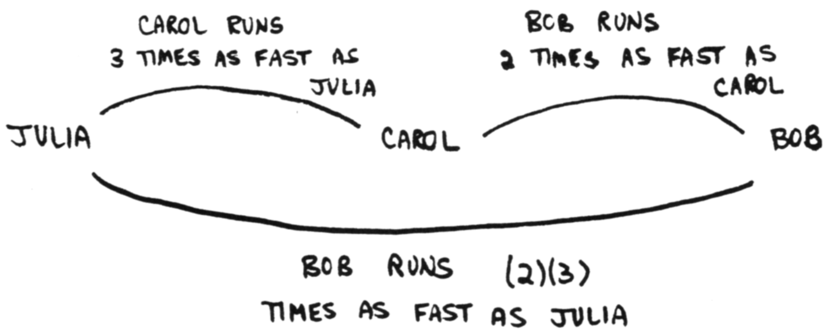
Consider the following scenario:

Suppose that Bob runs 2 times as fast as Carol, and Carol runs 3 times as fast as Julia.

If Julia runs 1 mile/hour, then Carol runs 3(1) miles/hour, and Bob runs 2(3) = 6 miles/hour. Thus, Bob runs 6 times as fast as Julia.

Observe that the rates *multiply.* This situation can be rephrased as follows (here, '*roc*' and '*wrt*' are abbreviations for '*rate of change*' and '*with respect to*', respectively):

$$\textit{roc of Bob wrt Julia} = (\textit{roc of Bob wrt Carol}) \cdot (\textit{roc of Carol wrt Julia})$$



rephrasing in terms of the derivative

Now reconsider the idea in the previous example, but this time in terms of functions and their derivatives. The critical idea is this:

If a function h is differentiable at c, then:

$$h(x) \text{ changes } h'(c) \text{ times as fast as } x \text{ at the point } (c, h(c)) \qquad (*)$$

Refer to the 'function box' sketch below as you read the following discussion.

We want to find $(f \circ g)'(c)$.

That is, we want to know how fast the function values $(f \circ g)(x)$ are changing with respect to x, at the point $\bigl(c, f(g(c))\bigr)$.

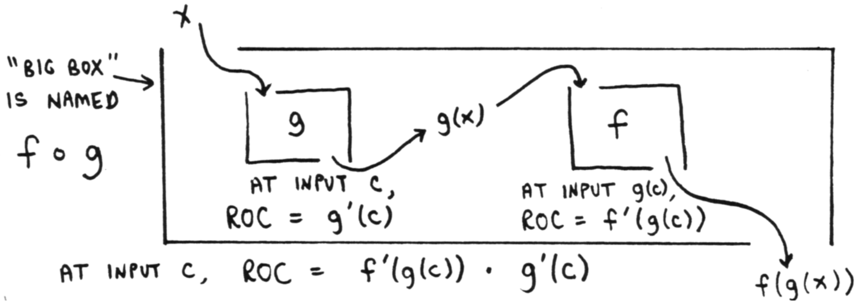
Well, $g(x)$ changes $g'(c)$ times as fast as x at $(c, g(c))$. (Rewrite (*), with 'h' replaced by 'g'.)

And, $f\bigl(g(x)\bigr)$ changes $f'(g(c))$ times as fast as $g(x)$ at $\bigl(g(c), f(g(c))\bigr)$. (Rewrite (*), with '$h$' replaced by '$f$', '$x$' replaced by '$g(x)$', and '$c$' replaced by '$g(c)$'.)

Thus, the rate of change of $(f \circ g)(x)$ wrt x at $(c, f(g(c)))$ should be:

$$f'(g(c)) \cdot g'(c)$$

And it is! The rule that tells us how to differentiate composite functions is called the *chain rule*; the name will be motivated shortly. A precise statement follows after a couple exercises.



EXERCISE 2

Consider the functions $f(x) = x^2$ and $g(x) = x + 1$ of an earlier example. It was found that $(f \circ g)(x) = x^2 + 2x + 1$. In this exercise, you will find the number $(f \circ g)'(2)$ in two different ways.

♣ 1. Differentiate $f \circ g$, and evaluate it at $x = 2$. What do you get?

Now, obtain the same result by doing the following:

♣ 2. Find out how fast the function values $g(x)$ are changing wrt x at $x = 2$. That is, find $g'(2)$.

♣ 3. Find out how fast the function values $f\bigl(g(x)\bigr)$ are changing wrt $g(x)$ at $g(2)$. That is, find $f'(g(2))$. (This is the function f', evaluated at the number $g(2)$.)

♣ 4. Multiply: $f'(g(2)) \cdot g'(2)$. Do your answers agree?

EXERCISE 3

♣ Repeat the previous exercise, except this time analyzing the function $g \circ f$ at $x = 2$. (You will need to make appropriate changes in the questions.)

THE CHAIN RULE Suppose that f and g are functions of x satisfying the following conditions:

- g is differentiable at x
- f is differentiable at $g(x)$

Then, the function $f \circ g$ is differentiable at x, and:

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

What is $f'(g(x))$? Be sure that you understand what $f'(g(x))$ represents: *it is the function f', evaluated at $g(x)$.*

Very roughly, in words, *to find out how fast $f \circ g$ changes with respect to x, we find how fast f changes wrt $g(x)$, and multiply by how fast g changes wrt x.*

Remember, *the chain rule tells you how to differentiate composite functions.*

EXAMPLE Problem: For the functions $f(x) = 3x^2 - 2x$ and $g(x) = x^3$, find $(f \circ g)'$ in two different ways.

Solution:

Method I: First find the function $f \circ g$:

$$(f \circ g)(x) = f(g(x)) = f(x^3) = 3(x^3)^2 - 2x^3 = 3x^6 - 2x^3$$

Then, differentiation yields:

$$(f \circ g)'(x) = 18x^5 - 6x^2$$

Method II: By the chain rule, $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$.

$$\begin{aligned} f'(x) &= 6x - 2 \\ f'(g(x)) &= f'(x^3) = 6x^3 - 2 \\ g'(x) &= 3x^2 \\ f'(g(x)) \cdot g'(x) &= (6x^3 - 2)(3x^2) = 18x^5 - 6x^2 \end{aligned}$$

EXERCISE 4 ♣ Let $f(x) = x^3$ and $g(x) = 3x^2 - 2x$. Find $(f \circ g)'$ in two different ways.

motivation for the name 'Chain Rule'

The chain rule can be extended to compositions of more than 2 functions, as follows:

$$\begin{aligned}(a\circ b\circ c)'(x) &= \big(a\circ(b\circ c)\big)'(x)\\ &= a'\big((b\circ c)(x)\big)\cdot(b\circ c)'(x)\\ &= a'(b(c(x)))\cdot b'(c(x))\cdot c'(x)\end{aligned}$$

(Make sure you understand *every step* here! The Chain Rule was applied twice; once in going from the first line to the second line; once in going from the second line to the third line.)

Similarly:

$$(a\circ b\circ c\circ d)'(x) = a'(b(c(d(x))))\cdot b'(c(d(x)))\cdot c'(d(x))\cdot d'(x)$$

Granted, the notation gets a bit unwieldy, but the important point is: see the *chains* that are forming? This is precisely the motivation for the name.

EXERCISE 5

♣ Write down the formula for the derivative of $a\circ b\circ c\circ d\circ e$ at x. Under what condition(s) do you think your formula holds?

★

function composition is associative

The previous argument used the fact that *composition of functions is associative.* This allows us to write things like $a\circ b\circ c$ without ambiguity. Indeed:

$$a\circ b\circ c = (a\circ b)\circ c = a\circ(b\circ c)$$

Leibniz notation for the chain rule

If y is a function of u, and u is a function of x, then the chain rule becomes, in Leibniz notation:

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

Roughly, to find out how fast y changes with respect to x, we find how fast y changes with respect to u, and multiply by how fast u changes with respect to x.

EXAMPLE

Problem: Let $y = u^2$ and $u = 3x^2 - x$. Find $\frac{dy}{dx}$ in two ways.

Method I: Write y as a function of x, and differentiate.

$$
\begin{aligned}
y &= u^2 = (3x^2 - x)^2 = 9x^4 - 6x^3 + x^2 \\
\frac{dy}{dx} &= 36x^3 - 18x^2 + 2x
\end{aligned}
$$

Method II: Use the chain rule.

$$
\begin{aligned}
\frac{dy}{du} &= 2u \\
\frac{du}{dx} &= 6x - 1 \\
\frac{dy}{dx} &= \frac{dy}{du} \cdot \frac{du}{dx} \\
&= (2u) \cdot (6x - 1) \\
&= 2(3x^2 - x)(6x - 1) \\
&= (6x^2 - 2x)(6x - 1) \\
&= 36x^3 - 18x^2 + 2x
\end{aligned}
$$

Note that since we want $\frac{dy}{dx}$ as a function of x, it was necessary to write $\frac{dy}{du}$ in terms of x.

EXERCISE 6

♣ 1. Let $y = 3u$ and $u = x^2 - 1$. Find $\frac{dy}{dx}$ in two ways.

♣ 2. Suppose y is a function of u, u is a function of v, and v is a function of x. Write down the formula for $\frac{dy}{dx}$, using Leibniz notation.

♣ 3. Let $y = u^2$, $u = 3v$ and $v = x^3$. Find $\frac{dy}{dx}$ in two ways.

some remarks on the proof of the chain rule

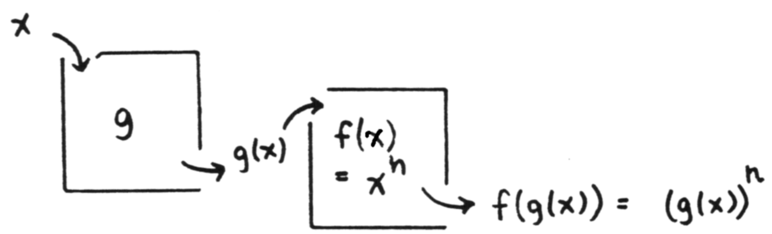
The proof of the chain rule is nontrivial; even in more advanced calculus books, it usually appears in an appendix, or as a supplement to the section on the chain rule.

appreciating the chain rule

It is often hard for the beginning calculus student to *appreciate the importance of the chain rule*. Perhaps this appreciation can begin by seeing all the new differentiation formulas that are an easy consequence of this rule…

how to differentiate $(g(x))^n$

Recall that the Simple Power Rule tells us how to differentiate x^n. However, we don't yet know a simple way to differentiate a *function* raised to a power, $(g(x))^n$. The chain rule will be used to tell us how to differentiate $(g(x))^n$! The trick comes in viewing $(g(x))^n$ as a *composition* of functions, as shown below:



Let $f(x) = x^n$, so that $(f \circ g)(x) = f(g(x)) = (g(x))^n$.

Thus, finding the derivative of $(g(x))^n$ reduces to finding the derivative of the composite function $f \circ g$.

This is easy, by the chain rule. First observe that $f'(x) = nx^{n-1}$ (by the Simple Power Rule), and then:

$$\begin{aligned}(f \circ g)'(x) &= f'(g(x)) \cdot g'(x) \\ &= n(g(x))^{n-1} \cdot g'(x)\end{aligned}$$

This result is summarized below.

GENERAL POWER RULE

differentiating $(g(x))^n$

The *general power rule* tells us how to differentiate $(g(x))^n$:

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

Observe that the General Power Rule looks a lot like the Simple Power Rule. The new part is that you must remember to *multiply by the derivative of the function that is being raised to the power*.

EXERCISE 7

♣ Think about what restrictions are necessary (say, on the exponent n and the function g) in order for the formula

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

to make sense.

EXAMPLE

using the General Power Rule

Problem: Differentiate $f(x) = (3x - 1)^7$.

Solution: Before the chain rule, we *could* differentiate this function f, but we would first need to multiply it out, and then differentiate term-by-term. The chain rule, however (under the guise of the General Power Rule) makes the problem easy:

$$f'(x) = 7(3x - 1)^6 \cdot (3) = 21(3x - 1)^6$$

The final form of the derivative obtained from using the chain rule is also *much more desirable* than the form obtained if we first multiplied f out, and then differentiated!

EXAMPLE

Problem: Differentiate $y = [x^2 - (x+1)^{-4}]^4$.

Solution: Be sure to write down complete mathematical sentences!

$$\begin{aligned}\frac{dy}{dx} &= 4[x^2 - (x+1)^{-4}]^3 \cdot \frac{d}{dx}[x^2 - (x+1)^{-4}] \\ &= 4[x^2 - (x+1)^{-4}]^3 \cdot [2x - (-4)(x+1)^{-5}(1)] \\ &= 4[x^2 - (x+1)^{-4}]^3 \cdot [2x + 4(x+1)^{-5}]\end{aligned}$$

Make sure you understand *every line* of this example. The General Power Rule was used twice—do you see where?

To find $\frac{dy}{dx}|_{x=0}$, just evaluate the formula at $x = 0$: $\frac{dy}{dx}|_{x=0} = 4[-1]^3 \cdot [4] = -16$.

EXAMPLE

Problem: Let f be a (differentiable) function of one variable.

Find $\frac{d}{dx}f(x^2+2x+1)$. (In other words, define h by $h(x) := f(x^2+2x+1)$, and find $\frac{d}{dx}h(x)$.)

Solution:

$$\frac{d}{dx}f(\overbrace{x^2+2x+1}^{g(x)}) = f'(\overbrace{x^2+2x+1}^{g(x)})\cdot(\overbrace{2x+2}^{g'(x)})$$

The Chain Rule was applied, taking $g(x) = x^2+2x+1$. The result is the function f', evaluated at x^2+2x+1, and then multiplied by $2x+2$.

Problem: Now, find $\frac{d}{dx}f(x^2+2x+1)|_{x=0}$.

Solution:

$$\begin{aligned}\frac{d}{dx}f(x^2+2x+1)|_{x=0} &= f'(x^2+2x+1)\cdot(2x+2)|_{x=0}\\ &= f'(0^2+2(0)+1)\cdot(2\cdot 0+2)\\ &= f'(1)\cdot 2\\ &= 2f'(1)\end{aligned}$$

This result cannot be simplified further, unless additional information is obtained about the function f.

EXERCISE 8

Differentiate the following functions. Use any appropriate method. It may be necessary to rewrite the functions before differentiating.

Then, find $f'(0)$ and $f'(1)$ (if they exist).

♣ 1. $f(x) = (2x+1)^7$

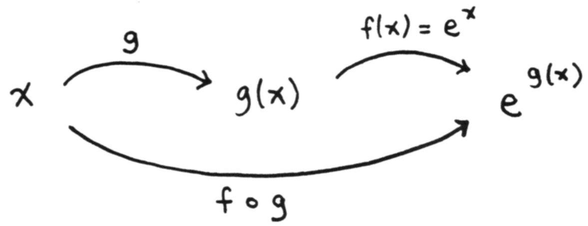
♣ 2. $f(x) = -\frac{1}{\sqrt{x^2+3}}$

♣ 3. $f(x) = (g(h(x)))^3$, where g and h are differentiable functions of one variable

♣ 4. $f(x) = [x+(x^2-1)^{-2}]^{-3}$

differentiating $e^{g(x)}$

To differentiate $e^{g(x)}$, the technique is again to view it as a composition:



First, define $f(x) = e^x$, so that $(f\circ g)(x) = f(g(x)) = e^{g(x)}$.

We seek $(f\circ g)'(x)$.

Recall that $f'(x) = e^x$. Then:

$$\begin{aligned}(f\circ g)'(x) &= f'(g(x))\cdot g'(x)\\ &= e^{g(x)}\cdot g'(x)\end{aligned}$$

This result is summarized next.

DIFFERENTIATION TOOL

differentiating $e^{g(x)}$

Suppose that g is differentiable at x. Then:

$$\frac{d}{dx}e^{g(x)} = e^{g(x)} \cdot g'(x)$$

EXERCISE 9

$\frac{d}{dx}\ln g(x) = \frac{1}{g(x)} \cdot g'(x)$

♣ Use the chain rule to show that:

$$\frac{d}{dx}\ln g(x) = \frac{1}{g(x)} \cdot g'(x)$$

What restrictions must be placed on g in order that this formula make sense?

EXAMPLE

Problem: Differentiate the following functions. Use any appropriate techniques. Be sure to write complete mathematical sentences.

a) $y = e^{x^2-1}$

b) $f(x) = e^{\sqrt{2x+1}}$

c) $g(t) = \ln\sqrt{t}$

d) $y = \frac{3}{\ln(2x-1)}$

Solutions:

a) $\frac{dy}{dx} = e^{x^2-1} \cdot (2x) = 2xe^{x^2-1}$

b)

$$\begin{aligned} f'(x) &= e^{\sqrt{2x+1}} \cdot \frac{d}{dx}(\sqrt{2x+1}) = e^{\sqrt{2x+1}} \cdot \frac{d}{dx}\big((2x+1)^{1/2}\big) \\ &= e^{\sqrt{2x+1}} \cdot (\frac{1}{2})(2x+1)^{\frac{1}{2}-1}(2) = e^{\sqrt{2x+1}} \cdot (2x+1)^{-\frac{1}{2}} \\ &= \frac{e^{\sqrt{2x+1}}}{\sqrt{2x+1}} \end{aligned}$$

c) Whenever possible, simplify the function by using properties of logarithms, before differentiating: $g(t) = \ln\sqrt{t} = \ln(t^{1/2}) = \frac{1}{2}\ln t$

Then, $g'(t) = \frac{1}{2} \cdot \frac{1}{t} = \frac{1}{2t}$.

d) First, rewrite y in a form that 'fits' the general power rule:

$$y = 3[\ln(2x-1)]^{-1}$$

Then:

$$\begin{aligned} \frac{dy}{dx} &= 3(-1)[\ln(2x-1)]^{-2}\frac{d}{dx}(\ln(2x-1)) \\ &= -3[\ln(2x-1)]^{-2} \cdot \frac{1}{2x-1} \cdot 2 \\ &= \frac{-6}{[\ln(2x-1)]^2(2x-1)} \end{aligned}$$

We have now added several important results to the list of Differentiation Tools:

DIFFERENTIATION TOOLS

prime notation	$\frac{d}{dx}$ operator
if $f(x) = k$, then $f'(x) = 0$	$\frac{d}{dx}(k) = 0$
$(kf)'(x) = k \cdot f'(x)$	$\frac{d}{dx}\big(kf(x)\big) = k \cdot f'(x)$
$(f+g)'(x) = f'(x) + g'(x)$	$\frac{d}{dx}\big(f(x)+g(x)\big) = f'(x) + g'(x)$
$(f-g)'(x) = f'(x) - g'(x)$	$\frac{d}{dx}\big(f(x)-g(x)\big) = f'(x) - g'(x)$
if $f(x) = x^n$, then $f'(x) = nx^{n-1}$	$\frac{d}{dx}(x^n) = nx^{n-1}$
if $f(x) = (g(x))^n$, then $f'(x) = n(g(x))^{n-1} \cdot g'(x)$	$\frac{d}{dx}(g(x))^n = n(g(x))^{n-1} \cdot g'(x)$
if $f(x) = e^x$, then $f'(x) = e^x$	$\frac{d}{dx}(e^x) = e^x$
if $f(x) = e^{g(x)}$, then $f'(x) = e^{g(x)} \cdot g'(x)$	$\frac{d}{dx}(e^{g(x)}) = e^{g(x)} \cdot g'(x)$
if $f(x) = \ln x$, then $f'(x) = \frac{1}{x}$	$\frac{d}{dx}(\ln x) = \frac{1}{x}$
if $f(x) = \ln(g(x))$, then $f'(x) = \frac{1}{g(x)} \cdot g'(x)$	$\frac{d}{dx}\big(\ln(g(x))\big) = \frac{1}{g(x)} \cdot g'(x)$

QUICK QUIZ
sample questions

1. Give a precise statement of the Chain Rule for differentiation. What type of function(s) does the Chain Rule tell you how to differentiate?
2. Let $f(x) = \sqrt{2}(1-x)^7$. Find $f'(x)$.
3. Suppose y is a function of w, w is a function of v, v is a function of u, and u is a function of t. Write a formula for $\frac{dy}{dt}$, using Leibniz notation.
4. Fill in the blanks: roughly, the formula

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

tells us that to find out how fast ____________ changes with respect to ____________, we find out how fast ____________ changes with respect to ____________, and multiply by how fast ____________ changes with respect to ____________.

5. Differentiate: $f(x) = \ln \sqrt[3]{2x+1}$

KEYWORDS
for this section

The chain rule (differentiating composite functions), motivation for the name 'chain rule', Leibniz notation for the chain rule, general power rule, differentiating $e^{g(x)}$ *and* $\ln(g(x))$.

END-OF-SECTION EXERCISES

The purpose of these exercises is to give you additional practice with all the differentiation formulas.

♣ Differentiate each of the following functions. Use any appropriate tools and notation. Be sure to write complete mathematical sentences. Write the derivative in a form that resembles, as closely as possible, the original function.

1. $f(x) = \frac{2}{\sqrt{e^x - 1}} + x$
2. $g(x) = \sqrt[3]{x^2 - 1}$
3. $y = (e^x)^3$
4. $y = e^{3x}$
5. $y = (3t - 4)^{11}$
6. $y = (2 - t)^8$
7. $g(t) = 3\sqrt[6]{t^2 + t + 1}$
8. $h(t) = -\sqrt[3]{\frac{1}{t^2 - 1}}$
9. $f(y) = 7e^{-y} + \ln(-y)$
10. $g(y) = \ln \sqrt[3]{-y}$
11. $y = (\ln x)^3$
12. $y = \ln(\sqrt{x}(x + 1))$
13. $y = \dfrac{-1}{t + \sqrt{t - 1}}$
14. $y = \dfrac{2}{(e^{3x} - 1)^4}$