A FEW NOTES ON NOTATION

1. **Variables, and why we often use letters x, y and f.** There are two types variables: *independent* and *dependent* variables:
    - The *independent* variable is the variable whose value can be anything; that is, its value does not depend on the value of some other variable. It also known as the *input* variable, as its value is input into a function, which then outputs some other value which depends on what value was input. The independent variable is often denoted as x (see a brief explanation below about why).
    - The *dependent* variable is the variable whose value depends on the independent variable. The dependent variable is often denoted y, simply because it is the next letter after x.

    Historically, x basically comes from the word that means "thing" in Arabic. See a slightly more detailed explanation at

    http://www.newton.dep.anl.gov/askasci/math99/math99228.htm.

    Functions are often given the letter y or f, which simply stands for "function." We write $y = f(x)$ to say "y is a function of x" i.e. the value of y depends on the value of x." For example, we might write $y = 5x^2 + 1$, or $y = f(x)$ where $f(x) = 5x^2 + 1$. So in this case, x is the *independent* variable, and y is the *dependent* variable, i.e. the variable whose value is a function of (i.e. depends on) the independent variable x.

2. **Why Δ?** Δ is "delta," the Greek letter for "d" which stands for "difference"—we tend to use the word "change" rather than "difference." We are often interested in how much a function f changes relative to how much x changes; that is, we are often interested in $\frac{\Delta f}{\Delta x} = \frac{\text{change in } f}{\text{change in } x}$. How much the function f changes will of course depend on how much x changes. Given a particular value of x, if the value x changes by Δx to $x + \Delta x$, then the value of f changes from $f(x)$ to $f(x + \Delta x)$. The change in x is simply the difference between its new value and its old value: $\Delta x = (x + \Delta x) - x = \Delta x$ (which is redundant, of course). Similarly, the change in function f's value is $\Delta f = f(x + \Delta x) - f(x)$, so that $\frac{\Delta f}{\Delta x} = \frac{f(x+\Delta x) - f(x)}{\Delta x}$. Sometimes instead of Δx we use a letter, such as h. So, for example, as x increases by h from x to $x + h$ (so Δx, the change in x, is simply h) we can write $\frac{\Delta f}{\Delta x} = \frac{f(x+h) - f(x)}{h}$.

3. **Infinitely small change**. The letter "d" (where again "d" is "difference") is used when the change or difference mentioned in Note 2 is infinitely small. "Infinitely small" is another way of saying "is getting smaller and smaller" or "getting closer and closer to 0." So $\frac{df}{dx} = \frac{\Delta f}{\Delta x}$ where Δx and Δf are infinitely small. We can write this as $\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$. It is automatically true that $\Delta f \to 0$ if $\Delta x \to 0$. In other words, $\Delta f = f(x+\Delta x) - f(x)$ automatically gets closer and closer to 0 (since $f(x+\Delta x)$ gets closer and closer to $f(x)$), as Δx gets closer and closer to 0, which is why we do not need to write something like $\frac{df}{dx} = \lim_{\substack{\Delta x \to 0 \\ \Delta f \to 0}} \frac{\Delta f}{\Delta x}$.

4. **Units, interpretation, and other variable letters**. Since $f'(x) = \frac{df}{dx} = \frac{\text{change in } f}{\text{change in } x}$, where these changes Δf and Δx are infinitely small, then the units of $\frac{df}{dx}$ are simply $\frac{\text{units of } f}{\text{units of } x}$. For example, $f'(2) = 5$ means that at the current value $x = 2$, f would increase by approximately 5 units for each increase in x of 1 unit (so if x were to increase by 2 units, f would increase by approximately 10 units, if x were to increase by .5 units, f would increase by 2.5 units, if x were to *decrease* by 1 unit, f would *decrease* by 5 units, and so on). Consider a more realistic example: suppose demand for coffee D, measured in tons per day, in a certain city is a function of its price p, measured in dollars per pound. Then $D'(p) = \frac{dD}{dp} = \frac{\text{change in } D}{\text{change in } p}$ is how demand would change if p were to increase by 1 unit, that is, by \$1 per pound. Suppose that $D'(3) = -1.3$. Adding units makes this a bit easier to interpret: $D'(\$3/\text{pound}) = -1.3 \frac{\text{tons/day}}{\$/\text{pound}}$, which means that if the current price is \$3/pound, a \$1/pound increase in price would result in a drop (a *drop*, since -1.3 is a *negative* number) in demand of approximately 1.3 tons/day. On the other hand, $D'(30) = -0.05$ means that if the current price is \$30/pound, a \$1/pound increase in price would result in a drop in demand of approximately 0.05 tons/day. This very small drop in demand is probably due to the fact that the demand would be so low to begin with if coffee were currently \$30/pound, so that there would not be much room for demand to drop even further if the price were to rise by \$1 per pound.

SLOPE OF A FUNCTION

We next briefly discuss slope and what it means. For now we'll just consider straight lines. Consider the line $y = 5x + 7$ or equivalently $f(x) = 5x + 7$. We list several values of x and y in the table below.

x	y
0	7
1	12
2	17
$\vdots$	$\vdots$
k	$5k + 7$
$k+1$	$5(k+1) + 7 = 5k + 12$

The slope of a line tells us how much y changes as x increases by 1. In the above example, the slope is 5, which means that y changes (in this case, it *increases*, since the slope 5 is *positive*) by 5 if x *increases* by 1. Notice this in the table. For example, when x changes from 1 to 2, y changes from 12 to 17, and when x changes from k to $k+1$, y changes from $5k+7$ to $5k+12$, and of course $5k+12$ is 5 greater than $5k+7$. Similarly, if we use the notation f instead of y, as we often do, so that we have $f(x) = 5x + 7$, then f *increases* by 5 if x *increases* by 1.

As another example, now with a negative slope, if $f(x) = -3.7x + 11$, then f *decreases* by 3.7 if x *increases* by 1. In general, the slope of any function (not just a straight line) at any particular point is the amount by which the function increases as its input variable x increases by 1 unit.

CHANGE IN FUNCTION RELATIVE TO CHANGE IN ITS INPUT VARIABLE

Given a function $f(x)$, quite often we are interested in how much f changes relative to how much x changes. That is, we are interested in $\frac{\Delta f}{\Delta x}$. Be sure you read Note 2 at the beginning of this Review.

For example, suppose that f is the number of toasters a factory produces and x is the number of employees the factory has employed. The number of toasters produced each day will depend on the number of employees working. That is, f is a function of, x: presumably the larger x is, the larger f will be. Suppose that $f(5)=100$ and $f(8)=190$, that is, this factory produces 100 toasters with 5 employees and 190 toasters with 8 employees. If the factory currently has 5 employees, then adding 3 more employees will increase the number of toasters produced from 100 to 190. It is handy to write the change in f relative to the change in x as

$$\frac{\Delta f}{\Delta x}=\frac{f(8)-f(5)}{8-5}=\frac{190-100}{3}=\frac{90\text{ toasters}}{3\text{ employees}}=\frac{30\text{ toasters}}{1\text{ employee}}=30\text{ toasters per employee}$$

That is, if there are currently 5 employees, then on average for each additional employee, we will be able to produce 30 additional toasters.

A SIMPLE REAL-LIFE EXAMPLE TO MOTIVATE DERIVATIVES

Sometimes we are interested in $\frac{\Delta f}{\Delta x}$ where Δx is extremely small. Be sure you read Notes 2 and 3 at the beginning of this Review. Consider the following example.

Suppose you are driving from Malibu to San Diego, and suppose that San Diego is exactly 100 miles away. Let position $p = p(t)$. That is, position p is a function of time t (remember we don't always have to use the letters x and f), where t is the number of hours after you leave Malibu on your trip to San Diego. Then $p(0) = 0$ means that at time $t = 0$ your position is $p = 0$. Similarly, $p(2) = 100$ means that at time $t = 2$ your position is $p = 100$, that is, after 2 hours you are 100 miles from where you started (hopefully you are now in San Diego, if you were driving in the right direction). Then of course your average speed for the entire trip would simply be the total distance traveled divided by the total time you were traveling, which for this example is

$$\frac{100 \text{ miles}}{2 \text{ hours}} = \frac{50 \text{ miles}}{1 \text{ hour}} = 50 \text{ miles/hour.}$$

Suppose you would like to know your speed at a specific time in your trip (this is often referred to as *instantaneous* speed), but that your car does not have a speedometer. Suppose you are interested in your instantaneous speed at $t = 1$, exactly one hour into your trip. Without a speedometer, you could simply use your average speed of 50 mph for the entire trip as an estimate for your instantaneous speed at $t = 1$. Unfortunately, this average speed may or (more likely) may not be a very good estimate for your instantaneous speed, unless you happen to drive at exactly the same speed the entire trip. A better estimate for your instantaneous speed at $t = 1$ would be to find your average speed between $t = 1$ and some other time just a bit later than $t = 1$, say one minute ($\frac{1}{60}$ of an hour) later. Suppose that $p(1) = 53.4$ and $p(1\frac{1}{60}) = 54.3$, that is, after one hour you have traveled 53.4 miles and after one hour and one minute you have traveled 54.3 miles. Then your average speed during that one minute of your drive is

$$\frac{p(1\frac{1}{60}) - p(1)}{1\frac{1}{60} - 1} = \frac{54.3 - 53.4}{\frac{1}{60}} = \frac{0.9 \text{ mile}}{\frac{1}{60} \text{ hour}} = 54 \text{ miles per hour}.$$

This average speed of 54 mph is probably a better estimate for your instantaneous speed at $t = 1$ than the average speed of 50 mph for the entire trip. If your average speed over this single minute is a better estimate of your instantaneous speed at $t = 1$ than the average speed over the entire two hours, then your average speed over an even smaller interval of time is probably an even better estimate. Let's assume still that $p(1) = 53.4$. Let h be a certain amount of time after $t = 1$, so for example $h = \frac{1}{60}$ means we are using our position at time $t = 1 + \frac{1}{60} = 1\frac{1}{60}$, as we just described. For different values of h, ranging from larger to smaller, we can find the average speed for each interval of time between $t = 1$ and $t = 1 + h$.

h	$t = 1 + h$	$p(t)$	Average speed between $t = 1$ and $t = 1 + h$
1	2	100.00	$\dfrac{p(2) - p(1)}{2 - 1} = \dfrac{100.00 - 53.40 \text{ miles}}{1 \text{ hour}} = 47.6 \text{ miles per hour}$
0.1	1.1	58.60	$\dfrac{p(1.1) - p(1)}{1.1 - 1} = \dfrac{58.60 - 53.40 \text{ miles}}{0.1 \text{ hour}} = 52 \text{ miles per hour}$
0.01	1.01	53.95	$\dfrac{p(1.01) - p(1)}{1.01 - 1} = \dfrac{53.95 - 53.40 \text{ miles}}{0.01 \text{ hour}} = 55 \text{ miles per hour}$
0.001	1.001	53.46	$\dfrac{p(1.001) - p(1)}{1.001 - 1} = \dfrac{53.46 - 53.40 \text{ miles}}{0.001 \text{ hour}} = 60 \text{ miles per hour}$

As we said above, the smaller h, the better the estimate is for the instantaneous speed at $t = 1$. The best estimate is probably the estimate for which we used the smallest interval of time, the estimate of 60 mph measured over 0.001 hour (3.6 seconds). That is, your average speed of 60 mph from $t = 1$ hour to $t = 1$ hour and 3.6 seconds is probably a very good estimate of your instantaneous speed right at $t = 1$.

In each estimate above, the average speed between time $t=1$ and $t=1+h$ is the change in position divided by the change in time, that is,

$$\text{average speed} = \frac{\Delta p}{\Delta t} = \frac{p(1+h)-p(1)}{(1+h)-1} = \frac{p(1+h)-p(1)}{h}.$$

Also remember that we know that the smaller h is, the better $\frac{p(1+h)-p(1)}{h}$ is as an estimate for the instantaneous speed at $t=1$. Theoretically, the instantaneous speed at $t=1$ is simply $\frac{p(1+h)-p(1)}{h}$ where we let h get smaller and smaller, that is, $\lim\limits_{h\to 0}\frac{p(1+h)-p(1)}{h}$. In this case, at $t=1$, we write $\frac{dp}{dt} = \lim\limits_{h\to 0}\frac{p(1+h)-p(1)}{h}$. (Again, be sure you understand Notes 2 and 3 at the beginning of this Review.) To summarize notation, where $t=1$, where h is the change in time,

$$\frac{\Delta p}{\Delta t} = \frac{p(1+h)-p(1)}{h} \quad \text{and} \quad \frac{dp}{dt} = \lim_{h\to 0}\frac{p(1+h)-p(1)}{h}.$$

In general, for any function $p = p(t)$ (not necessarily *position* as a function of *time*), $\frac{dp}{dt}$ is the rate of change of p with respect to t, that is, how much p changes relative to t, as discussed in Note 4 at the beginning of this Review. For the above example of driving from Malibu to San Diego, $\frac{dp}{dt}$ is the rate of change of position with respect to time (how much did our position change relative to how much time went by)—of course we usually simply call this *speed* or *velocity*.

ANOTHER EXAMPLE TO MOTIVATE DERIVATIVES

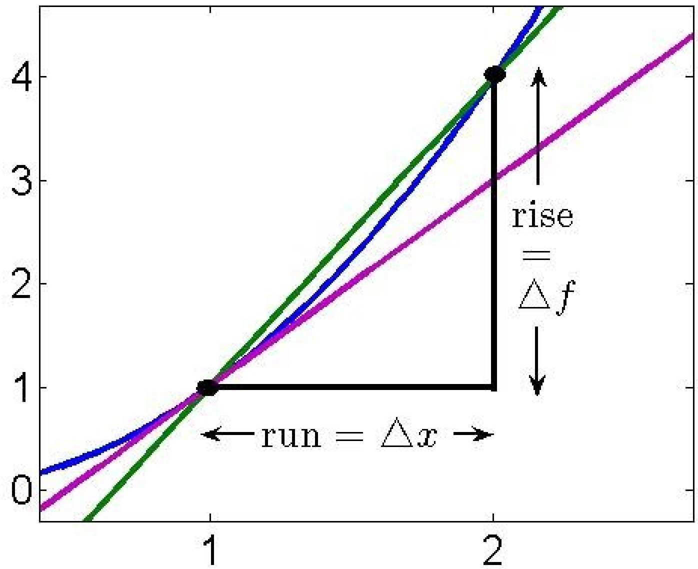
Sometimes we are interested in the slope of a function at a particular point. Suppose that we are interested in the slope of the function $f(x) = x^2$ at $x = 1$, as seen in the figure below. The slope of the purple line is the slope of $f(x) = x^2$ at $x = 1$.

Remember that slope is rise over run, the change in the function f divided by the change in the variable x that caused the change in the function. For example, given two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$, the slope of the line between these two points is

$$slope = \frac{rise}{run} = \frac{\Delta f}{\Delta x} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

To find the slope of the function $f(x) = x^2$ at $x = 1$, we take a second point near $x = 1$, say $x = 1 + h$ where h is small, and then find the slope of the line between the two points $(1,1)$ and $(1+h, (1+h)^2)$. The slope of the line that passes through these two points is an estimate for the slope of the function right at $x = 1$. The smaller h is (that is, the closer this second point is to the point we are interested in), the better the estimate will be. In the following figure we illustrate this where $h = 1$, so that the second point used is at $x = 1 + h = 1 + 1 = 2$. In the figure:

- The blue curve is the function $f(x) = x^2$.
- The purple line is the line passing through and tangent to the function at the point $(1,1)$ and whose slope of 2 is the true slope of the function at $x = 1$.
- The green line is the line passing through the function at the two points $(1,1)$ and $(2,4)$. The slope of this green line is $\frac{p(2) - p(1)}{2-1} = \frac{2^2 - 1^2}{1} = 3$. We use this slope as an estimate for the slope of the function at $x = 1$. 3 is not a very good estimate for the actual slope of 2, since the point $x = 2$ is not really that close to $x = 1$, and we generally only get good estimates if the second point is relatively close to the first point.



We list (but we don't graph) some other estimates for the slope of the purple line by taking points $x = 1 + h$ closer (i.e. smaller values of h) to the point of interest $x = 1$

h	$1+h$	Slope of line through $(1, f(1))$ and $(1+h, f(1+h))$
1	2	$\dfrac{p(2)-p(1)}{2-1} \quad = \dfrac{2^2-1^2}{1} \quad = 3.0000$
0.1	1.1	$\dfrac{p(1.1)-p(1)}{1.1-1} \quad = \dfrac{1.1^2-1^2}{0.1} \quad = 2.1000$
0.01	1.01	$\dfrac{p(1.01)-p(1)}{1.01-1} \quad = \dfrac{1.01^2-1^2}{0.01} \quad = 2.0100$
0.001	1.001	$\dfrac{f(1.001)-f(1)}{1.001-1} \quad = \dfrac{1.001^2-1^2}{0.001} \quad = 2.0010$

Again, the smaller h is, the closer the second point is to the first point, and the better the estimate is for the slope of the green line. It would appear that the true slope (rate of change) of the function at $x = 1$ is somewhere around 2.

HOW DO WE FIND $f'(x) = \dfrac{df}{dx}$ EXACTLY RATHER THAN JUST ESTIMATING IT?

Remember that $f'(x) = \dfrac{df}{dx} = \lim\limits_{h\to 0} \dfrac{f(x+h)-f(x)}{h}$. See beginning Notes 2 and 3.

For example, if $f(x) = x^2$, then at a particular value of x,

$f'(x) = \lim\limits_{h\to 0} \dfrac{f(x+h)-f(x)}{h}$	Can we plug $h=0$ in yet? No—we would divide by 0.
$= \lim\limits_{h\to 0} \dfrac{(x+h)^2 - x^2}{h}$	Can we plug $h=0$ in yet? No—we would divide by 0.
$= \lim\limits_{h\to 0} \dfrac{x^2+2xh+h^2-x^2}{h}$	Can we plug $h=0$ in yet? No—we would divide by 0.
$= \lim\limits_{h\to 0} \dfrac{2xh+h^2}{h}$	Can we plug $h=0$ in yet? No—we would divide by 0.
$= \lim\limits_{h\to 0} 2x+h$	Can we plug $h=0$ in yet? Yes.
$= 2x+0$	
$= 2x$	

So, for example, just above we found the slope of $f(x) = x^2$ at $x=1$ is something close to 2. Now we see that $f'(1) = 2(1) = 2$ exactly.

Note that in finding the above limit, we cannot plug the value of $h=0$ into the fraction until the second to last step, as otherwise we would be dividing by 0 in the previous steps. We can compute virtually all of the limits given in the table on the next page using the definition of a derivative $f'(x) = \dfrac{df}{dx} = \lim\limits_{h\to 0} \dfrac{f(x+h)-f(x)}{h}$.

TABLE OF RULES FOR FINDING DERIVATIVES

Rule		Example	
Function	Derivative	Function	Derivative
C	0	5	0
Cx	C	$5x$	5
x^n	nx^{n-1}	x^7	$7x^6$
e^x	e^x	$5e^x$	$5e^x$
$\ln x$	$\frac{1}{x}$	$5\ln x$	$\frac{5}{x}$
$Cf(x)$	$Cf'(x)$	$5x^3$	$5\cdot 3x^2$
$f(x)\pm g(x)$	$f'(x)\pm g'(x)$	$x^3+x^{-\frac{1}{3}}-\frac{1}{7}x-7$	$3x^2-\frac{1}{3}x^{-\frac{4}{3}}-\frac{1}{7}$
$f(x)g(x)$	$f'(x)g(x)+f(x)g'(x)$	$x^3\cdot\ln x$	$3x^2\cdot\ln x+x^3\cdot\frac{1}{x}$
$f(g(x))$	$f'(g(x))g'(x)$	$(x^2+5x+3)^7$	$7(x^2+5x+3)^6(2x+5)$
$e^{f(x)}$	$e^{f(x)}f'(x)$	e^{x^2+5x+3}	$e^{x^2+5x+3}(2x+5)$
$\ln f(x)$	$\frac{f'(x)}{f(x)}$	$\ln(x^2+5x+3)$	$\frac{2x+5}{x^2+5x+3}$

REVIEW OF THE PRODUCT RULE

We will briefly review the product rule, mostly by looking at an example which helps explain it, which states that

$$\tfrac{d}{dx}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x) = \tfrac{df}{dx}g(x) + f(x)\tfrac{dg}{dx}.$$

Recall that $\frac{df}{dx} = \frac{\text{change in } f}{\text{change in } x}$. See Notes 2 through 4.

Consider the following example. At Disneyland (as well as most any other business that sells tickets), the revenue due to ticket sales is simply the product of the number of tickets sold and the price per ticket. For simplicity, we will assume there is one single ticket price. So we have $r(t) = n(t)c(t)$ where

- $r(t)$ is the revenue (in dollars per year)
- $n(t)$ is the number of tickets sold (in tickets per year)
- $c(t)$ is the cost per ticket (in dollars per ticket)

Suppose that in 2006, a total of 1,570,000 tickets were sold at a cost of \$45 per ticket. That is,

$$n(2006) = \frac{1{,}570{,}000 \text{ tickets}}{\text{year}} \quad \text{and} \quad c(2006) = \frac{\$45}{\text{ticket}}$$

so that

$$r(2006) = n(2006)c(2006) = \frac{1{,}570{,}000 \text{ tickets}}{\text{year}} \cdot \frac{\$45}{\text{ticket}} = \$70{,}650{,}000 \text{ per year}$$

Notice that the “tickets” units cancel out to leave us with units of “\$ / year.”

There are two ways for revenue to increase: raise ticket prices and/or sell more tickets. Suppose that from 2006 to 2007 Disneyland increased the number of tickets it sold by 50,000. Then the increase in revenue from 2006 to 2007 *due to an increase in the number of tickets sold* would be

$$50{,}000 \text{ tickets} \cdot \frac{\$45}{\text{ticket}} = \$2{,}250{,}000.$$

If this increase in tickets were a yearly event, so the increase in number of tickets sold were 50,000 tickets *per year*, then the increase in revenue *per year* would be

$$\frac{50{,}000 \text{ tickets}}{\text{year}} \cdot \frac{\$45}{\text{ticket}} = \$2{,}250{,}000 \textit{ per year}$$

Similarly, suppose that at the current 1,570,000 of tickets sold per year, the ticket price is increasing at \$1 per ticket per year. Then the revenue increase per year *due to an increase in ticket price* is

$$1{,}570{,}000 \text{ tickets} \cdot \frac{\$1/\text{ticket}}{\text{year}} = \$1{,}570{,}000 \text{ per year}.$$

So the total change in revenue (due to selling 50,000 more tickets plus raising the ticket price by \$1) per year is

$$\frac{50{,}000 \text{ tickets}}{\text{year}} \cdot \frac{\$45}{\text{ticket}} \quad + \quad 1{,}570{,}000 \text{ tickets} \cdot \frac{\$1/\text{ticket}}{\text{year}}$$

$$= \quad \$2{,}250{,}000 \text{ per year} \quad + \quad \$1{,}570{,}000 \text{ per year}.$$

So where

$$\frac{dn(2006)}{dt} = \frac{50{,}000 \text{ tickets}}{\text{year}} \quad \text{and} \quad \frac{dc(2006)}{dt} = \frac{\$1/\text{ ticket}}{\text{year}},$$

the *rate* of increase in annual revenue $\frac{dr}{dt}$ from 2006 to 2007 is

$$\frac{dr(2006)}{dt} = \frac{dn(2006)}{dt} c(2006) + n(2006) \frac{dc(2006)}{dt}$$

In general, the product rule says that

$$\frac{dr(t)}{dt} = \frac{dn(t)}{dt} c(t) + n(t) \frac{dc(t)}{dt},$$

which can also be written,

$$r'(t) = n'(t)c(t) + n(t)c'(t)$$

that is,

$$[n(t)c(t)]' = n'(t)c(t) + n(t)c'(t).$$

Using the more familiar notation of $f(x)$ and $g(x)$ we have

$$\boxed{[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x)}$$

or

$$\boxed{\frac{d}{dx}[f(x)g(x)] = \frac{df(x)}{dx} g(x) + f(x) \frac{dg(x)}{dx}}.$$

REVIEW OF THE CHAIN RULE

We will briefly review the quotient rule, mostly by looking at an example which helps explain it.

First, suppose you own a factory that produces toasters. The more employees you hire, the more toasters you produce (of course your employees do all the work—you just sit in your office counting the money as it pours in), and the more toasters you produce, the most money you make. So the amount of money you make is a function of the number of toasters you produce, which is a function of how many employees you hire, then the amount of money you make is a function of how many employees you hire.

If you make \$5 for each additional toaster you produce, and you get 3 additional toasters for each additional employee, then you make \$15 for each additional employee. Let's write this out in mathematical notation. Let

M = the amount of money made

T = the number of toasters produced

E = the number of employees working for you

Then "each you make \$5 for each additional toaster you produce" can be written

$$\frac{\Delta M}{\Delta T} = \frac{\text{change in amount of money}}{\text{change in number of toasters}} = \frac{\$5}{1\,\text{toaster}}$$

and "you get 3 additional toasters for each additional employee" can be written

$$\frac{\Delta T}{\Delta E} = \frac{\text{change in number of toasters}}{\text{change in number of employees}} = \frac{3\,\text{toasters}}{1\,\text{employee}}.$$

which means that

$$\frac{\text{change in amount of money}}{\text{change in number of employees}} = \frac{\text{change in amount of money}}{\text{change in number of toasters}} \cdot \frac{\text{change in number of toasters}}{\text{change in number of employees}}$$

can be written

$$\frac{\Delta M}{\Delta E} = \frac{\Delta M}{\Delta T} \cdot \frac{\Delta T}{\Delta E} = \frac{\$5}{1\,\text{toaster}} \cdot \frac{3\,\text{toasters}}{1\,\text{employee}} = \frac{\$15}{1\,\text{employee}},$$

that is, "you make \$15 for each additional employee."

In general, if we have some quantity f that is a function of some other quantity g, that is, $f = f(g)$, and if g is a function of some quantity x, that is, $g = g(x)$, then f is really a function of x, since $f = f(g) = f(g(x))$. In other words, the value of x will determine (the value of g which will determine) the value of f. So if you change the value of x, you will change the value of f, so f is a function of x. (Sorry this is a bit redundant—I want to make sure this is clear.)

Since f is a function of x, we can find $\frac{\Delta f}{\Delta x}$ and it is simply $\frac{\Delta f}{\Delta x} = \frac{\Delta f}{\Delta g} \cdot \frac{\Delta g}{\Delta x}$. In words, $\frac{\text{change in } f}{\text{change in } x} = \frac{\text{change in } f}{\text{change in } g} \cdot \frac{\text{change in } g}{\text{change in } x}$. This leads to

$$\boxed{\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}}$$

which can also be written

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

since $g = g(x)$, that is, g is a function of x.

For example, if $f(g) = g^3$ and $g(x) = x^2 + 5x + 1$, then $\frac{df}{dg} = 3g^2$ and $\frac{dg}{dx} = 2x + 5$, so that

$$\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx} = 3g^2 \cdot (2x+5) = 3(x^2 + 5x + 1)^2 (2x+5)$$

We can also write this as $f(g(x)) = f(x^2 + 5x + 1) = (x^2 + 5x + 1)^3$, and

$$\frac{d}{dx} f(g(x)) = f'(g(x))g'(x) = 3(g(x))^2 g'(x) = 3(x^2 + 5x + 1)^2 (2x+5).$$

One more example. If $f(g) = e^g$ and $g(x) = x^4 + x$, then $\frac{df}{dg} = e^g$ and $\frac{dg}{dx} = 4x^3 + 1$, so that

$$\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx} = e^g \cdot (4x^3 + 1) = e^{x^4 + x} \cdot (4x^3 + 1).$$

We can also write this as $f(g(x)) = f(x^4 + x) = e^{x^4 + x}$, and

$$\frac{d}{dx} f(g(x)) = f'(g(x))g'(x) = e^g g'(x) = e^{x^4 + x} \cdot (4x^3 + 1).$$

In general, $\frac{d}{dx} e^{g(x)} = e^{g(x)} g'(x)$.

FINDING THE DERIVATIVE "WITH RESPECT TO…"

$\frac{d}{dx}$ means "find the derivative with respect to x of …" For example, $\frac{d}{dx}(x^2+5x+3)=2x+5$. In general, $\frac{d}{dx}f=\frac{df}{dx}$ and $\frac{d}{dx}f(x)=\frac{df(x)}{dx}$.

The "with respect to x" part is important. It means think of x as the variable and anything else—in particular, other letters—as constants. For example, just as

$$\frac{d}{dx}5x^2=5\cdot 2x=10x,$$

then where a is a constant we have

$$\frac{d}{dx}ax^2=a\cdot 2x=2ax,$$

We can also use letters other than x as the variable. (Of course we can—otherwise it would be blatant case of variable bias, which in our modern society is quite illegal, not to mention just plain wrong!) Suppose b were instead the variable. Then we would have

$$\frac{d}{db}ab^2=a\cdot 2b=2ab.$$

In this example, $\frac{d}{db}$ means "take the derivative with respect to b of … ," in which case we think of b as the variable and think of everything else as a constant. Here are some other examples.

In finding	Think of ____ as the variable	Think of ____ as constants
$\frac{d}{da}ab^2c^3=\frac{d}{da}b^2c^3a=b^2c^3$	a	b , c
$\frac{d}{db}ab^2c^3=\frac{d}{db}ac^3b^2=ac^3\cdot 2b=2abc^3$	b	a , c
$\frac{d}{dc}ab^2c^3=ab^2\cdot 3c^2=3ab^2c^2$	c	a , b

A FEW MORE EXAMPLES

Let a, b and c are constants.

$$\frac{d}{dx}ae^{\frac{x}{b}} = ae^{\frac{x}{b}} \cdot \frac{d}{dx}\frac{x}{b} = ae^{\frac{x}{b}} \cdot \frac{1}{b} = \frac{a}{b}e^{\frac{x}{b}}$$

$$\frac{d}{dx}ae^{\frac{b}{x}} = ae^{\frac{b}{x}} \cdot \frac{d}{dx}\frac{b}{x} = ae^{\frac{b}{x}} \cdot \frac{d}{dx}bx^{-1} = ae^{\frac{x}{b}} \cdot b(-1x^{-2}) = -\frac{abe^{\frac{x}{b}}}{x^2}$$

$$\frac{d}{dx}xe^{\frac{a}{b}} = \frac{d}{dx}e^{\frac{a}{b}}x = e^{\frac{a}{b}}$$

$$\frac{d}{dx}\sqrt{x^2+a^2} = \frac{d}{dx}(x^2+a^2)^{\frac{1}{2}} = \tfrac{1}{2}(x^2+a^2)^{-\frac{1}{2}}\frac{d}{dx}(x^2+a^2) = \tfrac{1}{2}(x^2+a^2)^{-\frac{1}{2}} \cdot 2x = \frac{x}{\sqrt{x^2+a^2}}$$

$$\frac{d}{dx}(e^{x^3}) = e^{x^3} \cdot \frac{d}{dx}(x^3) = e^{x^3} \cdot 3x^2 = 3x^2e^{x^3}$$

$$\frac{d}{dx}(x^3e^{x^3}) = \frac{d}{dx}(x^3) \cdot e^{x^3} + x^3 \cdot \frac{d}{dx}(e^{x^3}) = 3x^2 \cdot e^{x^3} + x^3 \cdot 3x^2e^{x^3} = 3x^2e^{x^3} + 3x^5e^{x^3}$$

$$\frac{d}{dx}3\sqrt{ax} = 3\sqrt{a} \cdot \frac{d}{dx}\sqrt{x} = 3\sqrt{a} \cdot \frac{d}{dx}x^{\frac{1}{2}} = 3\sqrt{a} \cdot \tfrac{1}{2}x^{-\frac{1}{2}} = \tfrac{3}{2}\sqrt{\frac{a}{x}}$$

$$\frac{d}{dx}(\frac{x^2}{a^2+b^2}) = \frac{1}{a^2+b^2} \cdot \frac{d}{dx}x^2 = \frac{1}{a^2+b^2} \cdot 2x = \frac{2x}{a^2+b^2}$$

$$\frac{d}{dx}(\frac{a^2}{x^2+b^2}) = a^2 \cdot \frac{d}{dx}(x^2+b^2)^{-1} = a^2 \cdot (-1)(x^2+b^2)^{-2}\frac{d}{dx}(x^2+b^2)$$

$$= a^2 \cdot (-1)(x^2+b^2)^{-2}(2x) = \frac{-2a^2x}{(x^2+b^2)^2}$$

$$\frac{d}{dx}(\frac{a^2}{x^2+b^2})^3 = 3(\frac{a^2}{x^2+b^2})^2 \cdot \frac{d}{dx}(\frac{a^2}{x^2+b^2}) = 3(\frac{a^2}{x^2+b^2})^2\frac{-2a^2x}{(x^2+b^2)^2} = \frac{-6a^2x^5}{(x^2+b^2)^3}$$

$$\frac{d}{dx}(\frac{x^2}{a^2+b^2})^3 = 3(\frac{x^2}{a^2+b^2})^2 \cdot \frac{d}{dx}(\frac{x^2}{a^2+b^2}) = 3(\frac{x^2}{a^2+b^2})^2\frac{2x}{a^2+b^2} = \frac{6x^5}{(a^2+b^2)^3}$$

or $\frac{d}{dx}(\frac{x^2}{a^2+b^2})^3 = (\frac{1}{a^2+b^2})^3\frac{d}{dx}x^6 = \frac{6x^5}{(a^2+b^2)^3}$, the same as found on the line before.

$$\frac{d}{dx}\ln(x^2+a^2) = \frac{1}{x^2+a^2} \cdot \frac{d}{dx}(x^2+a^2) = \frac{1}{x^2+a^2} \cdot 2x = \frac{2x}{x^2+a^2}$$

BRIEF REVIEW OF HOW TO FIND A MINIMUM OR MAXIMUM OF A FUNCTIOn.

While we will not give a complete discussion here of how to use first and second derivatives to find the minima and maxima (minimums and maximums, if you prefer) of a function, we will give a brief summary. If a function has a relative minimum or maximum at a certain value $x = a$, then is must be that $f'(a) = 0$. However, $f'(a) = 0$ does not guarantee that $f(x)$ has a minimum or maximum at $x = a$. If $f'(a) = 0$, then the value of the second derivative at $f''(a)$ tells us whether $f(x)$ has a minimum or maximum at $x = a$. We give a bit more detail.

In general, the derivative of $f(x)$ tells us whether $f(x)$ is decreasing, increasing or neither:

If $f'(x)$	then $f(x)$ is
< 0	decreasing
$= 0$	not changing
> 0	increasing

Similarly, the derivative of $f'(x)$ (in other words, $f''(x)$) tells us whether $f'(x)$ is decreasing, increasing or neither:

If $f''(x)$	then $f'(x)$ is
< 0	decreasing
$= 0$	not changing
> 0	increasing

Suppose we know a function has a minimum or maximum at a certain value $x = a$. The function will have a minimum at $x = a$ if the slope of the function (not the function itself) is increasing at $x = a$. Similarly, the function will have a maximum at $x = a$ if the slope of the function (not the function itself) is decreasing at $x = a$. To summarize, if $f'(a) = 0$, then

if $f''(a)$	then $f'(a)$ is	and at $x = a$, $f(x)$ has a
< 0	decreasing	maximum
$= 0$	not changing	(don't know)
> 0	increasing	minimum

Consider the function $f(x) = 3x^5 - 5x^3$, which has four critical points, that is, points at which $f'(x) = 0$:

$$f'(x) = 15x^4 - 15x^2 = 15x^2(x^2 - 1) = 15x^2(x-1)(x+1),$$

so

$$f'(x) = 0 \text{ at } x = 0,\ x = 1 \text{ and } x = -1.$$

(The critical point $x = 0$ counts as two of the four critical points.) Since $f''(x) = 60x^3 - 30x$, then $f''(-1) = -30 < 0$, which means f has a relative maximum at $x = -1$. Similarly, $f''(1) = 60 > 0$, which means f has a relative minimum at $x = 1$. Since $f''(0) = 0$, we can't tell just from this information whether at $x = 0$ the function has a minimum, maximum or neither (it turns out it has neither, as you can see in the figure). As you see in the graph, the values of the function at these three points are $f(-1) = 2$, $f(0) = 0$ and $f(1) = -2$.