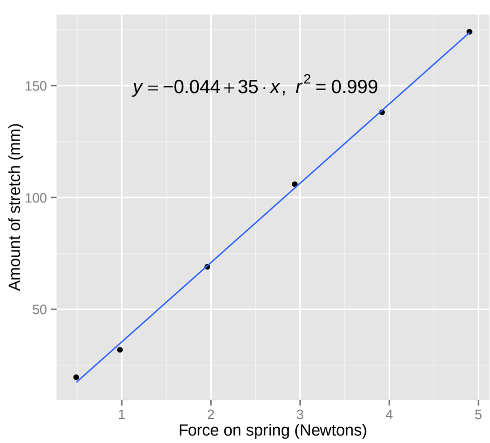
Chapter 3

Linear Regression

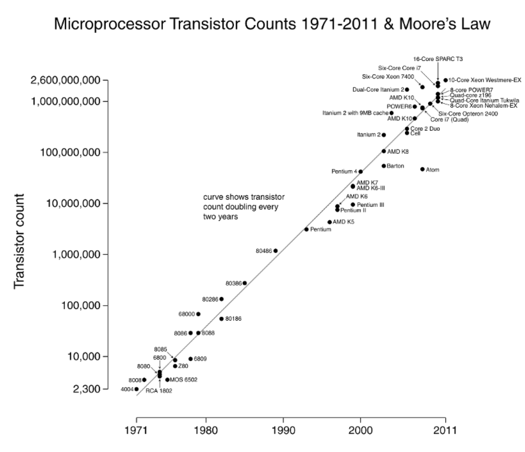
Once we've acquired data with multiple variables, one very important question is how the variables are related. For example, we could ask for the relationship between people's weights and heights, or study time and test scores, or two animal populations. **Regression** is a set of techniques for estimating relationships, and we'll focus on them for the next two chapters.

In this chapter, we'll focus on finding one of the simplest type of relationship: linear. This process is unsurprisingly called **linear regression**, and it has many applications. For example, we can relate the force for stretching a spring and the distance that the spring stretches (Hooke's law, shown in Figure 3.1a), or explain how many transistors the semiconductor industry can pack into a circuit over time (Moore's law, shown in Figure 3.1b).

Despite its simplicity, linear regression is an incredibly powerful tool for analyzing data. While we'll focus on the basics in this chapter, the next chapter will show how just a few small tweaks and extensions can enable more complex analyses.



(a) In classical mechanics, one could empirically verify Hooke's law by dangling a mass with a spring and seeing how much the spring is stretched.



(b) In the semiconductor industry, Moore's law is an observation that the number of transistors on an integrated circuit doubles roughly every two years.

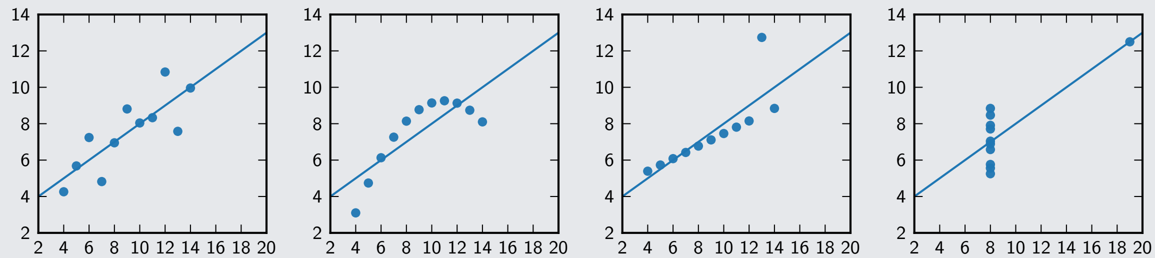
Figure 3.1: Examples of where a line fit explains physical phenomena and engineering feats.[1]

[1]The Moore's law image is by Wgsimon (own work) [CC-BY-SA-3.0 or GFDL], via Wikimedia Commons.

But just because fitting a line is easy doesn't mean that it always makes sense. Let's take another look at Anscombe's quartet to underscore this point.

Example: Anscombe's Quartet Revisited

Recall Anscombe's Quartet: 4 datasets with very similar statistical properties under a simple quantitative analysis, but that look very different. Here they are again, but this time with linear regression lines fitted to each one:



For all 4 of them, the slope of the regression line is 0.500 (to three decimal places) and the intercept is 3.00 (to two decimal places). This just goes to show: visualizing data can often reveal patterns that are hidden by pure numeric analysis!

We begin with **simple linear regression** in which there are only two variables of interest (e.g., weight and height, or force used and distance stretched). After developing intuition for this setting, we'll then turn our attention to **multiple linear regression**, where there are more variables.

Disclaimer: While some of the equations in this chapter might be a little intimidating, it's important to keep in mind that as a user of statistics, the most important thing is to understand their uses and limitations. Toward this end, make sure not to get bogged down in the details of the equations, but instead focus on understanding how they fit in to the big picture.

3.1 Simple linear regression

We're going to fit a line $y = \beta_0 + \beta_1 x$ to our data. Here, x is called the **independent variable** or **predictor variable**, and y is called the **dependent variable** or **response variable**.

Before we talk about how to do the fit, let's take a closer look at the important quantities from the fit:

- β_1 is the slope of the line: this is one of the most important quantities in any linear regression analysis. A value very close to 0 indicates little to no relationship; large positive or negative values indicate large positive or negative relationships, respectively. For our Hooke's law example earlier, the slope is the spring constant[2].

[2]Since the spring constant k is defined as $F = -kx$ (where F is the force and x is the stretch), the slope in Figure 3.1a is actually the inverse of the spring constant.

- β_0 is the intercept of the line.

In order to actually fit a line, we'll start with a way to quantify how good a line is. We'll then use this to fit the "best" line we can.

One way to quantify a line's "goodness" is to propose a probabilistic model that generates data from lines. Then the "best" line is the one for which data generated from the line is "most likely". This is a commonly used technique in statistics: proposing a probabilistic model and using the probability of data to evaluate how good a particular model is. Let's make this more concrete.

A probabilistic model for linearly related data

We observe paired data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, where we assume that as a function of x_i, each y_i is generated by using some true underlying line $y = \beta_0 + \beta_1 x$ that we evaluate at x_i, and then adding some Gaussian noise. Formally,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i. \tag{3.1}$$

Here, the noise ε_i represents the fact that our data won't fit the model perfectly. We'll model ε_i as being Gaussian: $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. Note that the intercept β_0, the slope β_1, and the noise variance σ^2 are all treated as fixed (i.e., deterministic) but unknown quantities.

Solving for the fit: least-squares regression

Assuming that this is actually how the data $(x_1, y_1), \ldots, (x_n, y_n)$ we observe are generated, then it turns out that we can find the line for which the probability of the data is highest by solving the following optimization problem[3]:

$$\min_{\beta_0,\beta_1} : \sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_i)\right]^2, \tag{3.2}$$

where $\min_{\beta_0,\beta_1}$ means "minimize over β_0 and β_1". This is known as the **least-squares linear regression problem**. Given a set of points, the solution is:

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n}\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - \frac{1}{n}(\sum_{i=1}^n x_i)^2} && (3.3)\\
&= r\frac{s_y}{s_x}, && (3.4)\\
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x}, && (3.5)
\end{aligned}$$

[3]This is an important point: the assumption of Gaussian noise leads to squared error as our minimization criterion. We'll see more regression techniques later that use different distributions and therefore different cost functions.

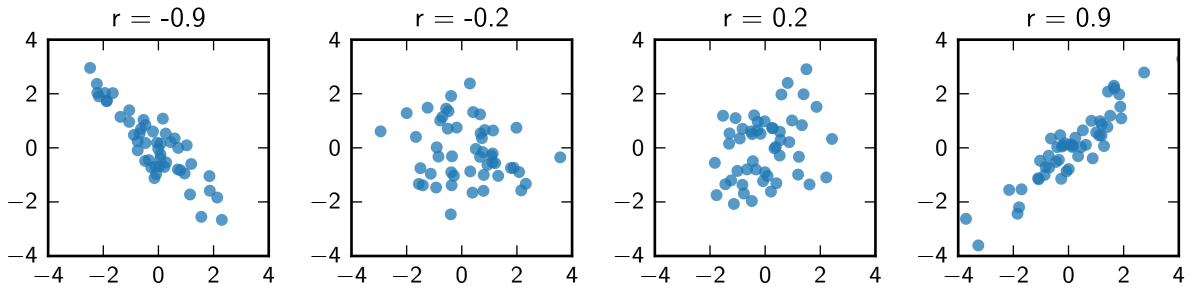


Figure 3.2: An illustration of correlation strength. Each plot shows data with a particular correlation coefficient r. Values farther than 0 (outside) indicate a stronger relationship than values closer to 0 (inside). Negative values (left) indicate an inverse relationship, while positive values (right) indicate a direct relationship.

where $\bar{x}$, $\bar{y}$, s_x and s_y are the sample means and standard deviations for x values and y values, respectively, and r is the **correlation coefficient**, defined as

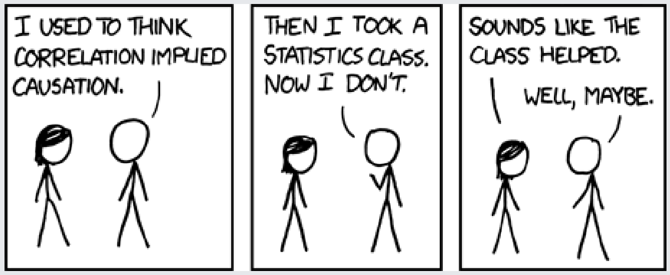
$$r = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right). \tag{3.6}$$

By examining the second equation for the estimated slope $\hat{\beta}_1$, we see that since sample standard deviations s_x and s_y are positive quantities, the correlation coefficient r, which is always between -1 and 1, measures how much x is related to y and whether the trend is positive or negative. Figure 3.2 illustrates different correlation strengths.

The square of the correlation coefficient r^2 will always be positive and is called the **coefficient of determination**. As we'll see later, this also is equal to the proportion of the total variability that's explained by a linear model.

As an extremely crucial remark, correlation does not imply causation! We devote the entire next page to this point, which is one of the most common sources of error in interpreting statistics.

Example: Correlation and Causation



Just because there's a strong correlation between two variables, there isn't necessarily a causal relationship between them. For example, drowning deaths and ice-cream sales are strongly correlated, but that's because both are affected by the season (summer vs. winter). In general, there are several possible cases, as illustrated below:

$$x \longrightarrow y \qquad x \longleftarrow y$$

(a) **Causal link**: Even if there is a causal link between x and y, correlation alone cannot tell us whether y causes x or x causes y.

$$\begin{tikzcd} & z \arrow[dl] \arrow[dr] & \\ x & & y \end{tikzcd}$$

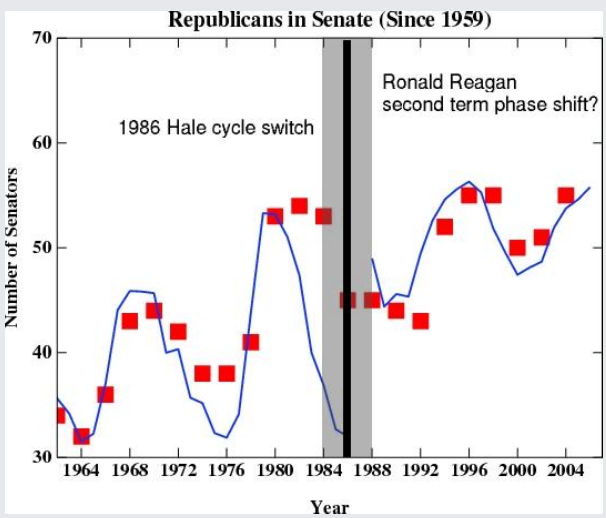
(b) **Hidden Cause**: A hidden variable z causes both x and y, creating the correlation.

$$\begin{tikzcd} & z \arrow[d] \\ x \arrow[r] & y \end{tikzcd}$$

(c) **Confounding Factor**: A hidden variable z and x both affect y, so the results also depend on the value of z.

$$x \qquad y$$

(d) **Coincidence**: The correlation just happened by chance (e.g. the strong correlation between sun cycles and number of Republicans in Congress, as shown below).



(e) The number of Republican senators in congress (red) and the sunspot number (blue, before 1986)/inverted sunspot number (blue, after 1986). This figure comes from http://www.realclimate.org/index.php/archives/2007/05/fun-with-correlations/.

Figure 3.3: Different explanations for correlation between two variables. In this diagram, arrows represent causation.

3.2 Tests and Intervals

Recall from last time that in order to do hypothesis tests and compute confidence intervals, we need to know our test statistic, its standard error, and its distribution. We'll look at the standard errors for the most important quantities and their interpretation. Any statistical analysis software can compute these quantities automatically, so we'll focus on interpreting and understanding what comes out.

Warning: All the statistical tests here crucially depend on the assumption that the observed data actually comes from the probabilistic model defined in Equation (3.1)!

3.2.1 Slope

For the slope β_1, our test statistic is

$$t_{\beta_1} = \frac{\hat{\beta}_1 - \beta_1}{s_{\beta_1}}, \tag{3.7}$$

which has a Student's t distribution with $n-2$ degrees of freedom. The standard error of the slope s_{β_1} is

$$s_{\beta_1} = \frac{\hat{\sigma}}{\sqrt{\underbrace{\sum_{i=1}^{n}(x_i - \bar{x})^2}_{\text{how close together } x \text{ values are}}}} \tag{3.8}$$

and the mean squared error $\hat{\sigma}^2$ is

$$\hat{\sigma}^2 = \frac{\overbrace{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}^{\text{how large the errors are}}}{n-2} \tag{3.9}$$

These terms make intuitive sense: if the x-values are all really close together, it's harder to fit a line. This will also make our standard error s_{β_1} larger, so we'll be less confident about our slope. The standard error also gets larger as the errors grow, as we should expect it to: larger errors should indicate a worse fit.

3.2.2 Intercept

For the intercept β_0, our test statistic is

$$t_{\beta_0} = \frac{\hat{\beta}_0 - \beta_0}{s_{\beta_0}}, \tag{3.10}$$

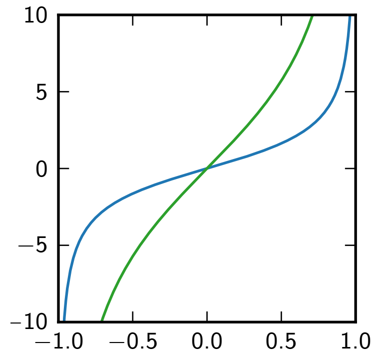


Figure 3.4: The test statistic for the correlation coefficient r for $n = 10$ (blue) and $n = 100$ (green).

which is also t-distributed with $n-2$ degrees of freedom. The standard error is

$$s_{\beta_0} = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_i (x_i - \bar{x})^2}}, \tag{3.11}$$

and $\hat{\sigma}$ is given by Equation (3.9).

■ 3.2.3 Correlation

For the correlation coefficient r, our test statistic is the standardized correlation

$$t_r = r\sqrt{\frac{n-2}{1-r^2}}, \tag{3.12}$$

which is t-distributed with $n-2$ degrees of freedom. Figure 3.4 plots t_r against r.

■ 3.2.4 Prediction

Let's look at the prediction at a particular value x^*, which we'll call $\hat{y}(x^*)$. In particular:

$$\hat{y}(x^*) = \hat{\beta}_0 + \hat{\beta}_1 x^*.$$

We can do this even if x^* wasn't in our original dataset.

Let's introduce some notation that will help us distinguish between predicting the line versus predicting a particular point generated from the model. From the probabilistic model given by Equation (3.1), we can similarly write how y is generated for the new point x^*:

$$y(x^*) = \underbrace{\beta_0 + \beta_1 x^*}_{\text{defined as } \mu(x^*)} + \varepsilon, \tag{3.13}$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$.

Then it turns out that the standard error $s_{\hat{\mu}}$ for estimating $\mu(x^*)$ (i.e., the mean of the line at point x^*) using $\hat{y}(x^*)$ is:

$$s_{\hat{\mu}} = \hat{\sigma}\sqrt{\frac{1}{n} + \underbrace{\frac{(x^* - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}_{\text{distance from "comfortable prediction region"}}}.$$

This makes sense because if we're trying to predict for a point that's far from the mean, then we should be less sure, and our prediction should have more variance. To compute the standard error for estimating a particular point $y(x^*)$ and not just its mean $\mu(x^*)$, we'd also need to factor in the extra noise term ε in Equation (3.13):

$$s_{\hat{y}} = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})}{\sum_i(x_i - \bar{x})^2} \underbrace{+1}_{\text{added}}}.$$

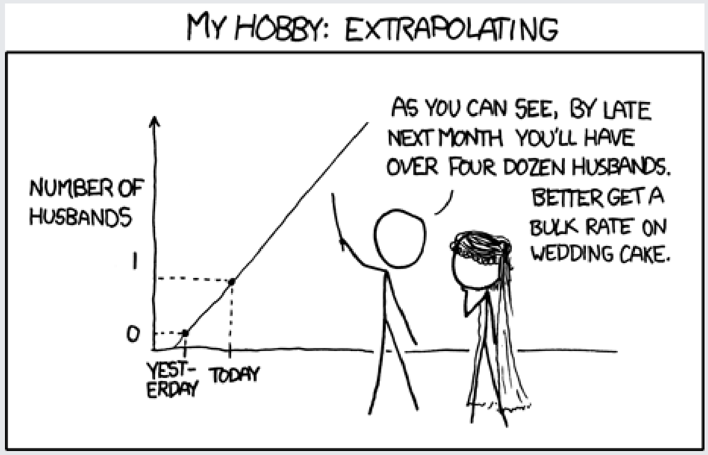
While both of these quantities have the same value when computed from the data, when analyzing them, we have to remember that they're different random variables: $\hat{y}$ has more variation because of the extra ε.

Interpolation vs. extrapolation

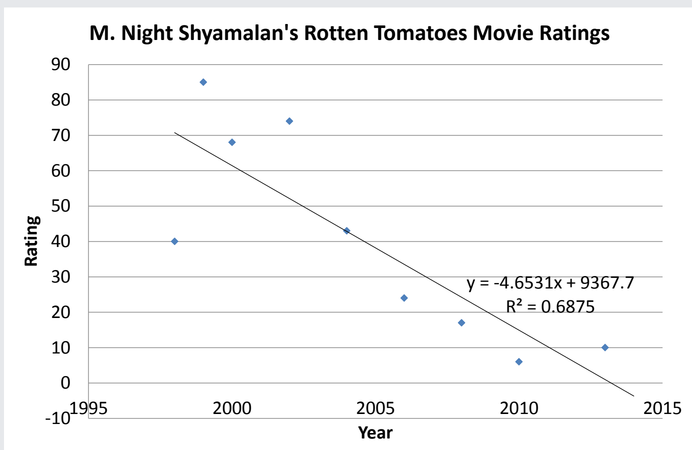
As a reminder, everything here crucially depends on the probabilistic model given by Equation (3.1) being true. In practice, when we do prediction for some value of x we haven't seen before, we need to be very careful. Predicting y for a value of x that is within the interval of points that we saw in the original data (the data that we fit our model with) is called **interpolation**. Predicting y for a value of x that's outside the range of values we actually saw for x in the original data is called **extrapolation**.

For real datasets, even if a linear fit seems appropriate, we need to be extremely careful about extrapolation, which can often lead to false predictions!

Example: The perils of extrapolation



By fitting a line to the Rotten Tomatoes ratings for movies that M. Night Shyamalan directed over time, one may erroneously be led to believe that in 2014 and onward, Shyamalan's movies will have negative ratings, which isn't even possible!



■ 3.3 Multiple Linear Regression

Now, let's talk about the case when instead of just a single scalar value x, we have a vector $(x_1, \ldots, x_p)$ for every data point i. So, we have n data points (just like before), each with p different predictor variables or **features**. We'll then try to predict y for each data point as a linear function of the different x variables:

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p. \tag{3.14}$$

Even though it's still linear, this representation is very versatile; here are just a few of the things we can represent with it:

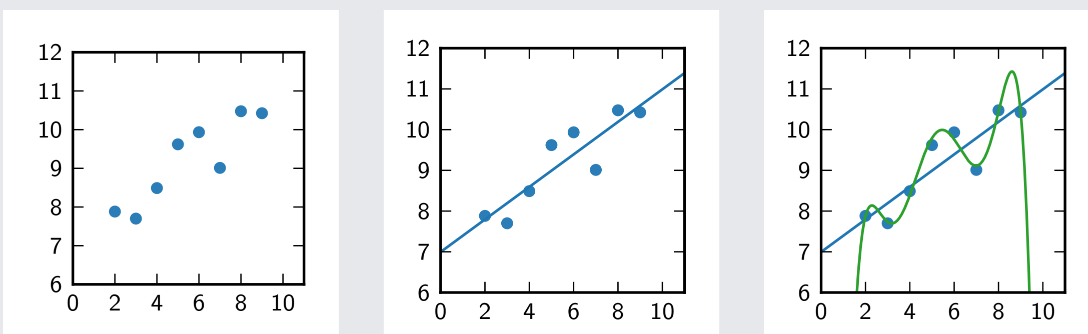
- Multiple dependent variables: for example, suppose we're trying to predict medical outcome as a function of several variables such as age, genetic susceptibility, and clinical diagnosis. Then we might say that for each patient, $x_1 = \text{age}, x_2 = \text{genetics}, x_3 = \text{diagnosis}$, and $y = \text{outcome}$.

- Nonlinearities: Suppose we want to predict a quadratic function $y = ax^2 + bx + c$: then for each data point we might say $x_1 = 1$, $x_2 = x$, and $x_3 = x^2$. This can easily be extended to any nonlinear function we want.

One may ask: why not just use multiple linear regression and fit an extremely high-degree polynomial to our data? While the model then would be much richer, one runs the risk of **overfitting**, where the model is so rich that it ends up fitting to the noise! We illustrate this with an example; it's also illustrated by a song[4].

Example: Overfitting

Using too many features or too complex of a model can often lead to overfitting. Suppose we want to fit a model to the points in Figure 3.3(a). If we fit a linear model, it might look like Figure 3.3(b). But, the fit isn't perfect. What if we use our newly acquired multiple regression powers to fit a 6th order polynomial to these points? The result is shown in Figure 3.3(c). While our errors are definitely smaller than they were with the linear model, the new model is far too complex, and will likely go wrong for values too far outside the range.



(a) A set of points with a simple linear relationship.

(b) The same set of points with a linear fit (blue).

(c) The same points with a 6th-order polynomial fit (green). As before, the linear fit is shown in blue.

We'll talk a little more about this in Chapters 4 and 5.

We'll represent our input data in matrix form as X, an $x \times p$ matrix where each row corresponds to a data point and each column corresponds to a feature. Since each output y_i is just a single number, we'll represent the collection as an n-element column vector y. Then our linear model can be expressed as

$$y = X\beta + \varepsilon \tag{3.15}$$

[4] Machine Learning A Cappella, Udacity. https://www.youtube.com/watch?v=DQWI1kvmwRg

where β is a p-element vector of coefficients, and ε is an n-element matrix where each element, like ε_i earlier, is normal with mean 0 and variance σ^2. Notice that in this version, we haven't explicitly written out a constant term like β_0 from before. We'll often add a column of 1s to the matrix X to accomplish this (try multiplying things out and making sure you understand why this solves the problem). The software you use might do this automatically, so it's something worth checking in the documentation.

This leads to the following optimization problem:

$$\min_{\beta} \sum_{i=1}^{n} (y_i - X_i\beta)^2, \tag{3.16}$$

where $\min_\beta$. just means "find values of β that minimize the following", and X_i refers to row i of the matrix X.

We can use some basic linear algebra to solve this problem and find the optimal estimates:

$$\hat{\beta} = (X^T X)^{-1} X^T y, \tag{3.17}$$

which most computer programs will do for you. Once we have this, what conclusions can we make with the help of statistics? We can obtain confidence intervals and/or hypothesis tests for each coefficient, which most statistical software will do for you. The test statistics are very similar to their counterparts for simple linear regression.

It's important not to blindly test whether all the coefficients are greater than zero: since this involves doing multiple comparisons, we'd need to correct appropriately using Bonferroni correction or FDR correction as described in the last chapter. But before even doing that, it's often smarter to measure whether the model even explains a significant amount of the variability in the data: if it doesn't, then it isn't even worth testing any of the coefficients individually. Typically, we'll use an **analysis of variance (ANOVA)** test to measure this. If the ANOVA test determines that the model explains a significant portion of the variability in the data, then we can consider testing each of the hypotheses and correcting for multiple comparisons.

We can also ask about which features have the most effect: if a feature's coefficient is 0 or close to 0, then that feature has little to no impact on the final result. We need to avoid the effect of scale: for example, if one feature is measured in feet and another in inches, even if they're the same, the coefficient for the feet feature will be twelve times larger. In order to avoid this problem, we'll usually look at the standardized coefficients $\frac{\hat{\beta}_k}{s_{\hat{\beta}_k}}$.

■ 3.4 Model Evaluation

How can we measure the performance of our model? Suppose for a moment that every point y_i was very close to the mean $\bar{y}$: this would mean that each y_i wouldn't depend on x_i, and that there wasn't much random error in the value either. Since we expect that this shouldn't be the case, we can try to understand how much the prediction from x_i and random error

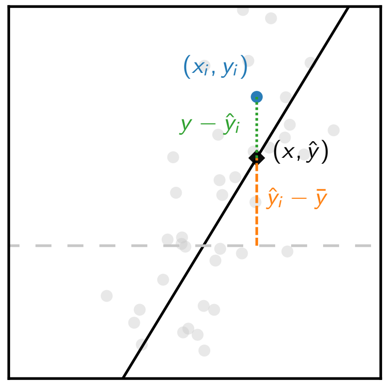


Figure 3.5: An illustration of the components contributing to the difference between the average y-value $\bar{y}$ and a particular point (x_i, y_i) (blue). Some of the difference, $\hat{y}_i - \bar{y}$, can be explained by the model (orange), and the remainder, $y_i - \hat{y}_i$, is known as the residual (green).

contribute to y_i. In particular, let's look at how far y_i is from the mean $\bar{y}$. We'll write this difference as:

$$y_i - \bar{y} = \underbrace{(\hat{y}_i - \bar{y})}_{\text{difference explained by model}} + \underbrace{(y_i - \hat{y}_i)}_{\text{difference not explained by model}} \tag{3.18}$$

In particular, the **residual** is defined to be $y_i - \hat{y}_i$: the distance from the original data point to the predicted value on the line. You can think of it as the error left over after the model has done its work. This difference is shown graphically in Figure 3.5. Note that the residual $y_i - \hat{y}$ isn't quite the same as the **noise** ε! We'll talk a little more about analyzing residuals (and why this distinction matters) in the next chapter.

If our model is doing a good job, then it should explain most of the difference from $\bar{y}$, and the first term should be bigger than the second term. If the second term is much bigger, then the model is probably not as useful.

If we square the quantity on the left, work through some algebra, and use some facts about linear regression, we'll find that

$$\underbrace{\sum_i (y_i - \bar{y})^2}_{\text{SS}_{\text{total}}} = \underbrace{\sum_i (\hat{y}_i - \bar{y})^2}_{\text{SS}_{\text{model}}} + \underbrace{\sum_i (y_i - \hat{y}_i)^2}_{\text{SS}_{\text{error}}}, \tag{3.19}$$

where "SS" stands for "sum of squares". These terms are often abbreviated as SST, SSM, and SSE respectively.

If we divide through by SST, we obtain

$$1 = \underbrace{\frac{\text{SSM}}{\text{SST}}}_{r^2} + \underbrace{\frac{\text{SSE}}{\text{SST}}}_{1-r^2},$$

where we note that r^2 is precisely the coefficient of determination mentioned earlier. Here, we see why r^2 can be interpreted as the fraction of variability in the data that is explained by the model.

One way we might evaluate a model's performance is to compare the ratio SSM/SSE. We'll do this with a slight tweak: we'll instead consider the mean values, $\mathrm{MSM} = \mathrm{SSM}/(p-1)$ and $\mathrm{MSE} = \mathrm{SSE}/(n-p)$, where the denominators correspond to the degrees of freedom. These new variables MSM and MSE have χ^2 distributions, and their ratio

$$f = \frac{\mathrm{MSM}}{\mathrm{MSE}} \tag{3.20}$$

has what's known as an F **distribution** with parameters $p-1$ and $n-p$. The widely used ANOVA test for categorical data, which we'll see in Chapter 6, is based on this F statistic: it's a way of measuring how much of the variability in the data is from the model and how much is from random error, and comparing the two.

Derivatives

Definition and Notation

If $y = f(x)$ then the derivative is defined to be $f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$.

If $y = f(x)$ then all of the following are equivalent notations for the derivative.

$$f'(x) = y' = \frac{df}{dx} = \frac{dy}{dx} = \frac{d}{dx}\left(f(x)\right) = Df(x)$$

If $y = f(x)$ all of the following are equivalent notations for derivative evaluated at $x = a$.

$$f'(a) = y'\big|_{x=a} = \frac{df}{dx}\bigg|_{x=a} = \frac{dy}{dx}\bigg|_{x=a} = Df(a)$$

Interpretation of the Derivative

If $y = f(x)$ then,

1. $m = f'(a)$ is the slope of the tangent line to $y = f(x)$ at $x = a$ and the equation of the tangent line at $x = a$ is given by $y = f(a) + f'(a)(x-a)$.
2. $f'(a)$ is the instantaneous rate of change of $f(x)$ at $x = a$.
3. If $f(x)$ is the position of an object at time x then $f'(a)$ is the velocity of the object at $x = a$.

Basic Properties and Formulas

If $f(x)$ and $g(x)$ are differentiable functions (the derivative exists), c and n are any real numbers,

1. $(cf)' = c f'(x)$
2. $(f \pm g)' = f'(x) \pm g'(x)$
3. $(fg)' = f'g + fg'$ – **Product Rule**
4. $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$ – **Quotient Rule**
5. $\frac{d}{dx}(c) = 0$
6. $\frac{d}{dx}(x^n) = n x^{n-1}$ – **Power Rule**
7. $\frac{d}{dx}\left(f(g(x))\right) = f'(g(x))\,g'(x)$
   This is the **Chain Rule**

Common Derivatives

$\frac{d}{dx}(x) = 1$	$\frac{d}{dx}(\csc x) = -\csc x \cot x$	$\frac{d}{dx}(a^x) = a^x \ln(a)$
$\frac{d}{dx}(\sin x) = \cos x$	$\frac{d}{dx}(\cot x) = -\csc^2 x$	$\frac{d}{dx}(\mathbf{e}^x) = \mathbf{e}^x$
$\frac{d}{dx}(\cos x) = -\sin x$	$\frac{d}{dx}(\sin^{-1} x) = \frac{1}{\sqrt{1-x^2}}$	$\frac{d}{dx}(\ln(x)) = \frac{1}{x},\ x > 0$
$\frac{d}{dx}(\tan x) = \sec^2 x$	$\frac{d}{dx}(\cos^{-1} x) = -\frac{1}{\sqrt{1-x^2}}$	$\frac{d}{dx}(\ln\lvert x\rvert) = \frac{1}{x},\ x \neq 0$
$\frac{d}{dx}(\sec x) = \sec x \tan x$	$\frac{d}{dx}(\tan^{-1} x) = \frac{1}{1+x^2}$	$\frac{d}{dx}(\log_a(x)) = \frac{1}{x \ln a},\ x > 0$

Chain Rule Variants

The chain rule applied to some specific functions.

1. $\frac{d}{dx}\left(\left[f(x)\right]^n\right) = n\left[f(x)\right]^{n-1} f'(x)$
2. $\frac{d}{dx}\left(\mathbf{e}^{f(x)}\right) = f'(x)\,\mathbf{e}^{f(x)}$
3. $\frac{d}{dx}\left(\ln\left[f(x)\right]\right) = \frac{f'(x)}{f(x)}$
4. $\frac{d}{dx}\left(\sin\left[f(x)\right]\right) = f'(x)\cos\left[f(x)\right]$
5. $\frac{d}{dx}\left(\cos\left[f(x)\right]\right) = -f'(x)\sin\left[f(x)\right]$
6. $\frac{d}{dx}\left(\tan\left[f(x)\right]\right) = f'(x)\sec^2\left[f(x)\right]$
7. $\frac{d}{dx}\left(\sec\left[f(x)\right]\right) = f'(x)\sec\left[f(x)\right]\tan\left[f(x)\right]$
8. $\frac{d}{dx}\left(\tan^{-1}\left[f(x)\right]\right) = \frac{f'(x)}{1+\left[f(x)\right]^2}$

Higher Order Derivatives

The Second Derivative is denoted as

$f''(x) = f^{(2)}(x) = \frac{d^2 f}{dx^2}$ and is defined as

$f''(x) = \left(f'(x)\right)'$, *i.e.* the derivative of the first derivative, $f'(x)$.

The nth Derivative is denoted as

$f^{(n)}(x) = \frac{d^n f}{dx^n}$ and is defined as

$f^{(n)}(x) = \left(f^{(n-1)}(x)\right)'$, *i.e.* the derivative of the $(n\text{-}1)^{\text{st}}$ derivative, $f^{(n-1)}(x)$.

Implicit Differentiation

Find y' if $\mathbf{e}^{2x-9y} + x^3 y^2 = \sin(y) + 11x$. Remember $y = y(x)$ here, so products/quotients of x and y will use the product/quotient rule and derivatives of y will use the chain rule. The "trick" is to differentiate as normal and every time you differentiate a y you tack on a y' (from the chain rule). After differentiating solve for y'.

$$\mathbf{e}^{2x-9y}(2 - 9y') + 3x^2y^2 + 2x^3 y\, y' = \cos(y)\, y' + 11$$
$$2\mathbf{e}^{2x-9y} - 9y'\mathbf{e}^{2x-9y} + 3x^2y^2 + 2x^3 y\, y' = \cos(y)\, y' + 11 \quad \Rightarrow \quad y' = \frac{11 - 2\mathbf{e}^{2x-9y} - 3x^2y^2}{2x^3 y - 9\mathbf{e}^{2x-9y} - \cos(y)}$$
$$\left(2x^3 y - 9\mathbf{e}^{2x-9y} - \cos(y)\right) y' = 11 - 2\mathbf{e}^{2x-9y} - 3x^2y^2$$

Increasing/Decreasing – Concave Up/Concave Down

Critical Points

$x = c$ is a critical point of $f(x)$ provided either **1.** $f'(c) = 0$ or **2.** $f'(c)$ doesn't exist.

Increasing/Decreasing

1. If $f'(x) > 0$ for all x in an interval I then $f(x)$ is increasing on the interval I.
2. If $f'(x) < 0$ for all x in an interval I then $f(x)$ is decreasing on the interval I.
3. If $f'(x) = 0$ for all x in an interval I then $f(x)$ is constant on the interval I.

Concave Up/Concave Down

1. If $f''(x) > 0$ for all x in an interval I then $f(x)$ is concave up on the interval I.
2. If $f''(x) < 0$ for all x in an interval I then $f(x)$ is concave down on the interval I.

Inflection Points

$x = c$ is a inflection point of $f(x)$ if the concavity changes at $x = c$.

Extrema

Absolute Extrema

1. $x = c$ is an absolute maximum of $f(x)$ if $f(c) \geq f(x)$ for all x in the domain.
2. $x = c$ is an absolute minimum of $f(x)$ if $f(c) \leq f(x)$ for all x in the domain.

Fermat's Theorem

If $f(x)$ has a relative (or local) extrema at $x = c$, then $x = c$ is a critical point of $f(x)$.

Extreme Value Theorem

If $f(x)$ is continuous on the closed interval $[a,b]$ then there exist numbers c and d so that, **1.** $a \leq c, d \leq b$, **2.** $f(c)$ is the abs. max. in $[a,b]$, **3.** $f(d)$ is the abs. min. in $[a,b]$.

Finding Absolute Extrema

To find the absolute extrema of the continuous function $f(x)$ on the interval $[a,b]$ use the following process.

1. Find all critical points of $f(x)$ in $[a,b]$.
2. Evaluate $f(x)$ at all points found in Step 1.
3. Evaluate $f(a)$ and $f(b)$.
4. Identify the abs. max. (largest function value) and the abs. min.(smallest function value) from the evaluations in Steps 2 & 3.

Relative (local) Extrema

1. $x = c$ is a relative (or local) maximum of $f(x)$ if $f(c) \geq f(x)$ for all x near c.
2. $x = c$ is a relative (or local) minimum of $f(x)$ if $f(c) \leq f(x)$ for all x near c.

1st Derivative Test

If $x = c$ is a critical point of $f(x)$ then $x = c$ is

1. a rel. max. of $f(x)$ if $f'(x) > 0$ to the left of $x = c$ and $f'(x) < 0$ to the right of $x = c$.
2. a rel. min. of $f(x)$ if $f'(x) < 0$ to the left of $x = c$ and $f'(x) > 0$ to the right of $x = c$.
3. not a relative extrema of $f(x)$ if $f'(x)$ is the same sign on both sides of $x = c$.

2nd Derivative Test

If $x = c$ is a critical point of $f(x)$ such that $f'(c) = 0$ then $x = c$

1. is a relative maximum of $f(x)$ if $f''(c) < 0$.
2. is a relative minimum of $f(x)$ if $f''(c) > 0$.
3. may be a relative maximum, relative minimum, or neither if $f''(c) = 0$.

Finding Relative Extrema and/or Classify Critical Points

1. Find all critical points of $f(x)$.
2. Use the 1st derivative test or the 2nd derivative test on each critical point.

Mean Value Theorem

If $f(x)$ is continuous on the closed interval $[a,b]$ and differentiable on the open interval (a,b) then there is a number $a < c < b$ such that $f'(c) = \dfrac{f(b) - f(a)}{b - a}$.

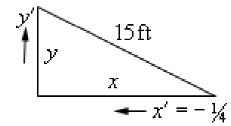
Newton's Method

If x_n is the n^{th} guess for the root/solution of $f(x) = 0$ then $(n+1)^{\text{st}}$ guess is $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$ provided $f'(x_n)$ exists.

Related Rates

Sketch picture and identify known/unknown quantities. Write down equation relating quantities and differentiate with respect to t using implicit differentiation (*i.e.* add on a derivative every time you differentiate a function of t). Plug in known quantities and solve for the unknown quantity.

Ex. A 15 foot ladder is resting against a wall. The bottom is initially 10 ft away and is being pushed towards the wall at $\frac{1}{4}$ ft/sec. How fast is the top moving after 12 sec?



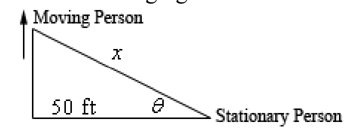
x' is negative because x is decreasing. Using Pythagorean Theorem and differentiating,

$$x^2 + y^2 = 15^2 \quad \Rightarrow \quad 2x\,x' + 2y\,y' = 0$$

After 12 sec we have $x = 10 - 12\left(\frac{1}{4}\right) = 7$ and so $y = \sqrt{15^2 - 7^2} = \sqrt{176}$. Plug in and solve for y'.

$$7\left(-\tfrac{1}{4}\right) + \sqrt{176}\;y' = 0 \;\Rightarrow\; y' = \frac{7}{4\sqrt{176}} \text{ ft/sec}$$

Ex. Two people are 50 ft apart when one starts walking north. The angle θ changes at 0.01 rad/min. At what rate is the distance between them changing when $\theta = 0.5$ rad?



We have $\theta' = 0.01$ rad/min. and want to find x'. We can use various trig fcns but easiest is,

$$\sec\theta = \frac{x}{50} \quad \Rightarrow \quad \sec\theta\tan\theta\;\theta' = \frac{x'}{50}$$

We know $\theta = 0.5$ so plug in θ' and solve.

$$\sec(0.5)\tan(0.5)(0.01) = \frac{x'}{50}$$

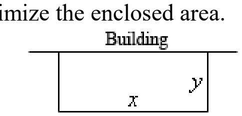
$$x' = 0.3112 \text{ ft/min}$$

Remember to have calculator in radians!

Optimization

Sketch picture if needed, write down equation to be optimized and constraint. Solve constraint for one of the two variables and plug into first equation. Find critical points of equation in range of variables and verify that they are min/max as needed.

Ex. We're enclosing a rectangular field with 500 ft of fence material and one side of the field is a building. Determine dimensions that will maximize the enclosed area.



Maximize $A = xy$ subject to constraint of $x + 2y = 500$. Solve constraint for x and plug into area.

$$x = 500 - 2y \quad \Rightarrow \quad \begin{aligned} A &= y(500 - 2y) \\ &= 500y - 2y^2 \end{aligned}$$

Differentiate and find critical point(s).

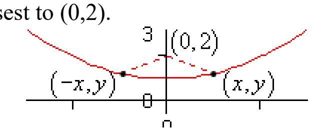
$$A' = 500 - 4y \quad \Rightarrow \quad y = 125$$

By 2nd deriv. test this is a rel. max. and so is the answer we're after. Finally, find x.

$$x = 500 - 2(125) = 250$$

The dimensions are then 250 x 125.

Ex. Determine point(s) on $y = x^2 + 1$ that are closest to (0,2).



Minimize $f = d^2 = (x-0)^2 + (y-2)^2$ and the constraint is $y = x^2 + 1$. Solve constraint for x^2 and plug into the function.

$$\begin{aligned} x^2 = y - 1 \;\Rightarrow\; f &= x^2 + (y-2)^2 \\ &= y - 1 + (y-2)^2 = y^2 - 3y + 3 \end{aligned}$$

Differentiate and find critical point(s).

$$f' = 2y - 3 \quad \Rightarrow \quad y = \tfrac{3}{2}$$

By the 2nd derivative test this is a rel. min. and so all we need to do is find x value(s).

$$x^2 = \tfrac{3}{2} - 1 = \tfrac{1}{2} \quad \Rightarrow \quad x = \pm\tfrac{1}{\sqrt{2}}$$

The 2 points are then $\left(\frac{1}{\sqrt{2}}, \frac{3}{2}\right)$ and $\left(-\frac{1}{\sqrt{2}}, \frac{3}{2}\right)$.

CHAPTER 4

The Derivative of a Function

OUTLINE

Pooled Samples

If you had to give a blood test to 3000 people to test for the presence of Hepatitis C, how would you do it? Would you test each person individually, which would require 3000 tests? That could be expensive, especially if each test were to cost $100. How can you use less tests? One way is to pool the samples. What sample size would you use? If you pooled the blood of 100 people and the test came out negative, then one test was used instead of 100. But if the test came out positive, then you would need to test each one, which now requires $1 + 100 = 101$ tests. Maybe the sample size to use is 200. Could it be 300? And the likelihood of a positive test must play a role as well. Sounds like a very hard problem. But with the discussion in this chapter and the Chapter Project to guide you, you can find the best sample size to use so cost is least.

A LOOK BACK, A LOOK FORWARD

In Chapter 1, we discussed various properties that functions have, such as intercepts, even/odd, increasing/decreasing, local maxima and minima, and average rate of change. In Chapter 2 we discussed classes of functions and listed some properties that these classes possess. In Chapter 3 we began our study of the calculus by discussing limits of functions and continuity of functions. Now we are ready to define another property of functions: the *derivative of a function.*

The cofounders of calculus are generally recognized to be Gottfried Wilhelm von Leibniz (1646–1716) and Sir Isaac Newton (1642–1727). Newton approached calculus by solving a physics problem involving falling objects, while Leibniz approached calculus by solving a geometry problem. Surprisingly, the solution of these two problems led to the same mathematical concept: the derivative. We shall discuss the physics problem later in this chapter. We shall address the geometry problem, referred to as *The Tangent Problem,* now.

4.1 The Definition of a Derivative

PREPARING FOR THIS SECTION ***Before getting started, review the following:***

- Average Rate of Change (Section 1.3, pp. xx–xx)
- Secant Line (Section 1.3, pp. xx–xx)
- Factoring (Section 0.3, pp. xx–xx)
- Point–Slope Form of a Line (Section 0.8, pp. xx–xx)
- Difference Quotient (Section 1.2, pp. xx–xx)

OBJECTIVES

1. Find an equation of the tangent line to the graph of a function
2. Find the derivative of a function at a number c
3. Find the derivative of a function using the difference quotient
4. Find the instantaneous rate of change of a function
5. Find marginal cost and marginal revenue

The Tangent Problem

The geometry question that motivated the development of calculus was "What is the slope of the tangent line to the graph of a function $y = f(x)$ at a point P on its graph?" See Figure 1.

FIGURE 1

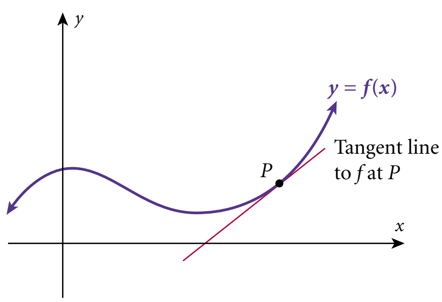
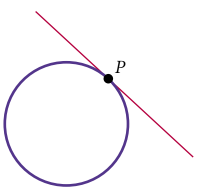


FIGURE 2



We first need to define what we mean by a *tangent* line. In high school geometry, the tangent line to a circle is defined as the line that intersects the graph in exactly one point. Look at Figure 2. Notice that the tangent line just touches the graph of the circle.

This definition, however, does not work in general. Look at Figure 3. The lines L_1 and L_2 only intersect the graph in one point P, but neither touches the graph at P. Additionally, the tangent line L_T shown in Figure 4 touches the graph of f at P, but also intersects the graph elsewhere. So how should we define the tangent line to the graph of f at a point P?

FIGURE 3

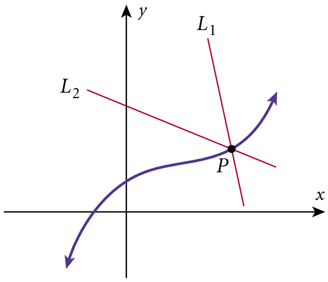
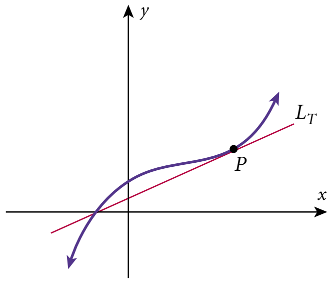


FIGURE 4

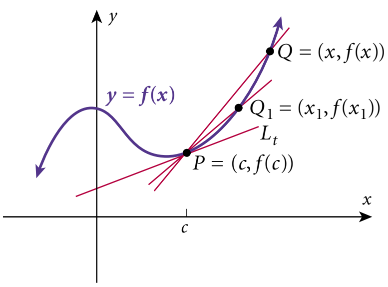


The tangent line L_T to the graph of a function $y = f(x)$ at a point P necessarily contains the point P. To find an equation for L_T using the point–slope form of the equation of a line, it remains to find the slope $m_{\tan}$ of the tangent line.

Suppose that the coordinates of the point P are $(c, f(c))$. Locate another point $Q = (x, f(x))$ on the graph of f. The line containing P and Q is a secant line. (Refer to Section 1.3.) The slope $m_{\sec}$ of the secant line is

$$m_{\sec} = \frac{f(x) - f(c)}{x - c}$$

FIGURE 5



Now look at Figure 5.

As we move along the graph of f from Q toward P, we obtain a succession of secant lines. The closer we get to P, the closer the secant line is to the tangent line. The limiting position of these secant lines is the tangent line. Therefore, the limiting value of the slopes of these secant lines equals the slope of the tangent line. But, as we move from Q toward P, the values of x get closer to c. Therefore,

$$m_{\tan} = \lim_{x \to c} m_{\sec} = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

The **tangent line** to the graph of a function $y = f(x)$ at a point $P = (c, f(c))$ on its graph is defined as the line containing the point P whose slope is

$$\boxed{m_{\tan} = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}} \qquad \textbf{(1)}$$

provided that this limit exists.

If $m_{\tan}$ exists, an equation of the tangent line is

$$\boxed{y - f(c) = m_{\tan}(x - c)} \qquad \textbf{(2)}$$

1 EXAMPLE 1 Finding an Equation of the Tangent Line

Find an equation of the tangent line to the graph of $f(x) = \dfrac{x^2}{4}$ at the point $\left(1, \dfrac{1}{4}\right)$. Graph the function and the tangent line.

SOLUTION The tangent line contains the point $\left(1, \dfrac{1}{4}\right)$. The slope of the tangent line to the graph of $f(x) = \dfrac{x^2}{4}$ at $\left(1, \dfrac{1}{4}\right)$ is

$$m_{\tan} = \lim_{x\to 1} \frac{f(x) - f(1)}{x - 1} = \lim_{x\to 1} \frac{\dfrac{x^2}{4} - \dfrac{1}{4}}{x - 1} = \lim_{x\to 1} \frac{\cancel{(x-1)}(x+1)}{4\cancel{(x-1)}}$$

$$= \lim_{x\to 1} \frac{x+1}{4} = \frac{1}{2}$$

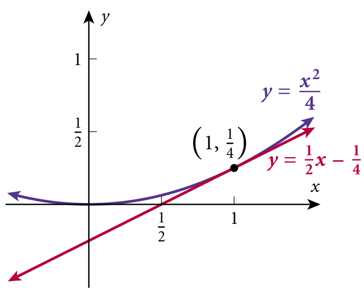
An equation of the tangent line is

$$y - \frac{1}{4} = \frac{1}{2}(x - 1) \qquad y - f(c) = m_{\tan}(x - c)$$

$$y = \frac{1}{2}x - \frac{1}{4}$$

Figure 6 shows the graph of $y = \dfrac{x^2}{4}$ and the tangent line at $\left(1, \dfrac{1}{4}\right)$.

FIGURE 6



NOW WORK PROBLEM 3.

The limit in formula (1) has an important generalization: it is called the *derivative of f at c*.

The Derivative of a Function at a Number c.

Let $y = f(x)$ denote a function f. If c is a number in the domain of f, the **derivative of f at c,** denoted by $f'(c)$, read "f prime of c," is defined as

$$\boxed{f'(c) = \lim_{x\to c} \frac{f(x) - f(c)}{x - c}} \qquad (3)$$

provided that this limit exists.

The steps for finding the derivative of a function are listed below:

Steps For Finding the Derivative of a Function at c

STEP 1 Find $f(c)$.

STEP 2 Subtract $f(c)$ from $f(x)$ to get $f(x) - f(c)$ and form the quotient

$$\boxed{\frac{f(x) - f(c)}{x - c}}$$

STEP 3 Find the limit (if it exists) of the quotient found in Step 2 as $x \to c$:

$$\boxed{f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}}$$

2

EXAMPLE 2 Finding the Derivative of a Function at a Number

Find the derivative of $f(x) = 2x^2 - 5x$ at 2. That is, find $f'(2)$.

SOLUTION **Step 1:** $f(2) = 2(4) - 5(2) = -2$

Step 2: $\dfrac{f(x) - f(2)}{x - 2} = \dfrac{(2x^2 - 5x) - (-2)}{x - 2} = \dfrac{2x^2 - 5x + 2}{x - 2} = \dfrac{(2x - 1)(x - 2)}{x - 2}$

Step 3: The derivative of f at 2 is

$$f'(2) = \lim_{x \to 2} \frac{f(x) - f(2)}{x - 2} = \lim_{x \to 2} \frac{(2x - 1)\cancel{(x - 2)}}{\cancel{x - 2}} = 3$$

NOW WORK PROBLEM 13.

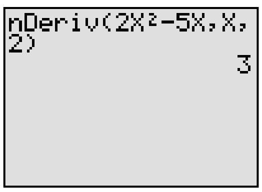
Example 2 provides a way of finding the derivative at 2 analytically. Graphing utilities have built-in procedures to approximate the derivative of a function at any number c. Consult your owner's manual for the appropriate keystrokes.

EXAMPLE 3 Finding the Derivative of a Function Using a Graphing Utility

Use a graphing utility to find the derivative of $f(x) = 2x^2 - 5x$ at 2. That is, find $f'(2)$.

SOLUTION Figure 7 shows the solution using a TI-83 graphing calculator.

FIGURE 7



So $f'(2) = 3$.

NOW WORK PROBLEM 45.

EXAMPLE 4 **Finding the Derivative of a Function at *c***

Find the derivative of $f(x) = x^2$ at c. That is, find $f'(c)$.

SOLUTION Since $f(c) = c^2$, we have

$$\frac{f(x) - f(c)}{x - c} = \frac{x^2 - c^2}{x - c} = \frac{(x + c)(x - c)}{x - c}$$

The derivative of f at c is

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c} \frac{(x + c)\cancel{(x - c)}}{\cancel{x - c}} = 2c$$

As Example 4 illustrates, the derivative of $f(x) = x^2$ exists and equals $2c$ for any number c. In other words, the derivative is itself a function and, using x for the independent variable, we can write $f'(x) = 2x$. The function f' is called the **derivative function of *f*** or the **derivative of *f*.** We also say that f is **differentiable.** The instruction "differentiate f" means "find the derivative of f".

It is usually easier to find the derivative function by using another form. We derive this alternate form as follows:

Formula (3) for the derivative of f at c is

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \qquad \text{Formula (3)}$$

Let $h = x - c$. Then $x = c + h$ and

$$\frac{f(x) - f(c)}{x - c} = \frac{f(c + h) - f(c)}{h}$$

Since $h = x - c$, then, as $x \to c$, it follows that $h \to 0$. As a result,

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lim_{h \to 0} \frac{f(c + h) - f(c)}{h} \tag{4}$$

Now replace c by x in (4). This gives us the following formula for finding the derivative of f at any number x.

Formula for the Derivative of a Function $y = f(x)$ at x

$$\boxed{f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}} \tag{5}$$

That is, the derivative of the function f is the limit of its difference quotient.

3 **EXAMPLE 5** **Using the Difference Quotient to Find a Derivative**

(a) Use formula (5) to find the derivative of $f(x) = x^2 + 2x$.
(b) Find $f'(0)$, $f'(-1)$, $f'(3)$.

SOLUTION **(a)** First, we find the difference quotient of $f(x) = x^2 + 2x$.

$$\frac{f(x+h) - f(x)}{h} = \frac{[(x+h)^2 + 2(x+h)] - [x^2 + 2x]}{h}$$

$$= \frac{x^2 + 2xh + h^2 + 2x + 2h - x^2 - 2x}{h}$$

$$= \frac{2xh + h^2 + 2h}{h} \qquad \text{Simplify.}$$

$$= \frac{h(2x + h + 2)}{h} \qquad \text{Factor out } h.$$

$$= 2x + h + 2 \qquad \text{Cancel the } h\text{'s.}$$

The derivative of f is the limit of the difference quotient as $h \to 0$. That is,

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0}(2x + h + 2) = 2x + 2$$

(b) Since,

$$f'(x) = 2x + 2$$

we have

$$f'(0) = 2 \cdot 0 + 2 = 2$$
$$f'(-1) = 2(-1) + 2 = 0$$
$$f'(3) = 2(3) + 2 = 8$$

NOW WORK PROBLEM 29.

Instantaneous Rate of Change

In Chapter 1 we defined the average rate of change of a function f from c to x as

$$\frac{\Delta y}{\Delta x} = \frac{f(x) - f(c)}{x - c}$$

The limit as x approaches c of the average rate of change of f, based on formula (3), is the derivative of f at c. As a result, we call the derivative of f at c the **instantaneous rate of change of f with respect to x at c.** That is,

$$\boxed{\left(\begin{array}{c}\text{Instantaneous rate of}\\ \text{change of } f \text{ with respect to } x \text{ at } c\end{array}\right) = f'(c) = \lim_{x\to c}\frac{f(x) - f(c)}{x - c}} \qquad (6)$$

4 **EXAMPLE 6 Finding the Instantaneous Rate of Change**

During a month-long advertising campaign, the total sales S of a magazine were given by the fraction

$$S(x) = 5x^2 + 100x + 10{,}000$$

where x represents the number of days of the campaign, $0 \le x \le 30$.

(a) What is the average rate of change of sales from $x = 10$ to $x = 20$ days?
(b) What is the instantaneous rate of change of sales when $x = 10$ days?

SOLUTION **(a)** Since $S(10) = 11{,}500$ and $S(20) = 14{,}000$, the average rate of change of sales from $x = 10$ to $x = 20$ is

$$\frac{\Delta S}{\Delta x} = \frac{S(20) - S(10)}{20 - 10} = \frac{14{,}000 - 11{,}500}{10} = 250 \text{ magazines per day}$$

(b) The instantaneous rate of change of sales when $x = 10$ is the derivative of S at 10.

$$S'(10) = \lim_{x \to 10} \frac{S(x) - S(10)}{x - 10} = \lim_{x \to 10} \frac{[(5x^2 + 100x + 10{,}000) - 11500]}{x - 10}$$

$$= \lim_{x \to 10} \frac{5(x^2 + 20x - 300)}{x - 10}$$

$$= 5 \lim_{x \to 0} \frac{(x + 30)(x - 10)}{x - 10} = 5 \lim_{x \to 10} (x + 30) = 5 \cdot 40 = 200$$

The instantaneous rate of change of S at 10 is 200 magazines per day.

We interpret the results of Example 6 as follows: The fact that the average rate of sales from $x = 10$ to $x = 20$ is $\Delta S/\Delta x = 250$ magazines per day indicates that on the 10th day of the campaign, we can expect to average 250 magazines per day of additional sales if we continue the campaign for 10 more days. The fact that $S'(10) = 200$ magazines per day indicates that on the 10th day of the campaign, one more day of advertising will result in additional sales of approximately 200 magazines per day.

NOW WORK PROBLEM 39.

Find marginal cost and marginal revenue 5

Application to Economics: Marginal Analysis

Economics is one of the many fields in which calculus has been used to great advantage. Economists have a special name for the application of derivatives to problems in economics—it is **marginal analysis.** Whenever the term *marginal* appears in a discussion, involving cost functions or revenue functions, it signals the presence of derivatives in the background.

Marginal Cost

Suppose $C = C(x)$ is the cost of producing x units. Then the derivative $C'(x)$ is called the **marginal cost.**

We interpret the marginal cost as follows. Since

$$C'(x) = \lim_{h \to 0} \frac{C(x + h) - C(x)}{h}$$

it follows for small values of h that

$$C'(x) \approx \frac{C(x + h) - C(x)}{h}$$

That is to say,

$$C'(x) \approx \frac{\text{cost of increasing production from } x \text{ to } x + h}{h}$$

In most practical situations x is very large. Because of this, many economists let $h = 1$, which is small compared to large x. Then, marginal cost may be interpreted as

$$\boxed{C'(x) = C(x+1) - C(x) = \text{cost of increasing production by one unit}}$$

EXAMPLE 7 Finding Marginal Cost

Suppose that the cost in dollars for a weekly production of x tons of steel is given by the function:

$$C(x) = \frac{1}{10}x^2 + 5x + 1000$$

(a) Find the marginal cost.
(b) Find the cost and marginal cost when $x = 1000$ tons.
(c) Interpret $C'(1000)$.

SOLUTION **(a)** The marginal cost is the derivative $C'(x)$. We use the difference quotient of $C(x)$ to find $C'(x)$.

$$C'(x) = \lim_{h\to 0}\frac{C(x+h) - C(x)}{h} = \lim_{h\to 0}\frac{\left[\frac{1}{10}(x+h)^2 + 5(x+h) + 1000\right] - \left[\frac{1}{10}x^2 + 5x + 1000\right]}{h}$$

$$= \lim_{h\to 0}\frac{\frac{1}{10}(x^2 + 2xh + h^2) + 5x + 5h - \frac{1}{10}x^2 - 5x}{h}$$

$$= \lim_{h\to 0}\frac{\frac{1}{5}xh + \frac{1}{10}h^2 + 5h}{h} = \lim_{h\to 0}\left(\frac{1}{5}x + \frac{1}{10}h + 5\right) = \frac{1}{5}x + 5$$

(b) We evaluate $C(x)$ and $C'(x)$ at $x = 1000$

The cost when $x = 1000$ tons is $C(1000) = \frac{1}{10}(1000)^2 + 5 \cdot 1000 + 1000 = \$106{,}000$

The marginal cost when $x = 1000$ tons is $C'(1000) = \frac{1}{5} \cdot 1000 + 5 = \$205/\text{ton}$

(c) $C'(1000) = \$205$ per ton means that the cost of producing one additional ton of steel after 1000 tons have been produced is approximately \$205.

Note that the average cost of producing one more ton of steel after the 1000th ton is

$$\frac{\Delta C}{\Delta x} = \frac{C(1001) - C(1000)}{1001 - 1000}$$

$$= \left(\frac{1}{10} \cdot 1001^2 + 5 \cdot 1001 + 1000\right) - \left(\frac{1}{10} \cdot 1000^2 + 5 \cdot 1000 + 1000\right)$$

$$= \$205.10/\text{ton}$$

We observe that the average cost differs from the marginal cost by only 0.1 dollar/ton, which is less than $\frac{1}{20}$th of 1%. Note, too, that the marginal cost is easier to compute than the actual average cost.

The money received by our hypothetical steel producer when he sells his product is the revenue. Specifically, let $R = R(x)$ be the total revenue received from selling x tons. Then the derivative $R'(x)$ is called the **marginal revenue.** For this example, marginal revenue, like marginal cost, is measured in dollars per ton. An approximate value for $R'(x)$ is obtained by noting again that

$$R'(x) \approx \frac{R(x+h) - R(x)}{h}$$

When x is large, then $h = 1$ is small by comparison, so that

$$\boxed{R'(x) = R(x+1) - R(x) = \text{revenue resulting from the sale of one additional unit}}$$

This is the interpretation many economists give to marginal revenue.

EXAMPLE 8 **Suppose that the revenue R for a weekly sale of x tons of steel is given by the formula**

$$R = x^2 + 5x$$

(a) Find the marginal revenue.
(b) Find the revenue and marginal revenue when $x = 1000$ tons.
(c) Interpret $R'(1000)$.

SOLUTION **(a)** The marginal revenue is the derivative $R'(x)$. We use the difference quotient of $R(x)$ to find $R'(x)$.

$$\begin{aligned} R'(x) = \lim_{h\to 0}\frac{R(x+h) - R(x)}{h} &= \lim_{h\to 0}\frac{\{[(x+h)^2 + 5(x+h)] - [x^2 + 5x]\}}{h} \\ &= \lim_{h\to 0}\frac{(x^2 + 2xh + h^2 + 5x + 5h - x^2 - 5x)}{h} \\ &= \lim_{h\to 0}\frac{(2xh + h^2 + 5h)}{h} = \lim_{h\to 0}(2x + h + 5) = 2x + 5 \end{aligned}$$

(b) The revenue when $x = 1000$ tons is

$$R(1000) = (1000)^2 + 5(1000) = \$1{,}005{,}000$$

The marginal revenue when $x = 1000$ tons is

$$R'(1000) = 2(1000) + 5 = \$2005/\text{ton}$$

(c) $R'(1000) = \$2005/\text{ton}$ means that the revenue due to selling one additional ton of steel after 1000 tons have been sold is approximately \$2005.

Note that the average revenue derived from selling one additional ton after 1000 tons have been sold is

$$\frac{\Delta R}{\Delta x} = \frac{R(1001) - R(1000)}{1001 - 1000} = 1{,}007{,}006 - 1{,}005{,}000 = \$2006/\text{ton}$$

Observe that the actual average revenue differs from the marginal revenue by only \$1/ton, or 0.05%.

NOW WORK PROBLEM 63.

SUMMARY The derivative of a function $y = f(x)$ at c is defined as

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

The derivative $f'(x)$ of a function $y = f(x)$ is $f'(x) = \lim_{h \to 0} \dfrac{f(x+h) - f(x)}{h}$

In geometry, $f'(c)$ equals the slope of the tangent line to the graph of f at the point $(c, f(c))$.

In applications, if two variables are related by the function $y = f(x)$, then $f'(c)$ equals the instantaneous rate of change of f with respect to x at c.

In economics, the derivative of a cost function is the marginal cost and the derivative of a revenue function is the marginal revenue.

EXERCISE 4.1 Answers to Odd-Numbered Problems Begin on Page AN-XX.

In Problems 1–12, find the slope of the tangent line to the graph of f at the given point. What is an equation of the tangent line? Graph f and the tangent line.

1. $f(x) = 3x + 5$ at $(1, 8)$
2. $f(x) = -2x + 1$ at $(-1, 3)$
3. $f(x) = x^2 + 2$ at $(-1, 3)$
4. $f(x) = 3 - x^2$ at $(1, 2)$
5. $f(x) = 3x^2$ at $(2, 12)$
6. $f(x) = -4x^2$ at $(-2, -16)$
7. $f(x) = 2x^2 + x$ at $(1, 3)$
8. $f(x) = 3x^2 - x$ at $(0, 0)$
9. $f(x) = x^2 - 2x + 3$ at $(-1, 6)$
10. $f(x) = -2x^2 + x - 3$ at $(1, -4)$
11. $f(x) = x^3 + x$ at $(2, 10)$
12. $f(x) = x^3 - x^2$ at $(1, 0)$

In Problems 13–24, find the derivative of each function at the given number.

13. $f(x) = -4x + 5$ at 3
14. $f(x) = -4 + 3x$ at 1
15. $f(x) = x^2 - 3$ at 0
16. $f(x) = 2x^2 + 1$ at -1
17. $f(x) = 2x^2 + 3x$ at 1
18. $f(x) = 3x^2 - 4x$ at 2
19. $f(x) = x^3 + 4x$ at -1
20. $f(x) = 2x^3 - x^2$ at 2
21. $f(x) = x^3 + x^2 - 2x$ at 1
22. $f(x) = x^3 - 2x^2 + x$ at -1
23. $f(x) = \dfrac{1}{x}$ at 1
24. $f(x) = \dfrac{1}{x}$ at 1

In Problems 25–36, find the derivative of f using difference quotients.

25. $f(x) = 2x$
26. $f(x) = 3x$
27. $f(x) = 1 - 2x$
28. $f(x) = 5 - 3x$
29. $f(x) = x^2 + 2$
30. $f(x) = 2x^2 - 3$
31. $f(x) = 3x^2 - 2x + 1$
32. $f(x) = 2x^2 + x + 1$
33. $f(x) = x^3$
34. $f(x) = \dfrac{1}{x}$
35. $f(x) = mx + b$
36. $f(x) = ax^2 + bx + c$

In Problems 37–44, find

(a) The average rate of change as x changes from 1 to 3.
(b) The (instantaneous) rate of change at 1.

37. $f(x) = 3x + 4$ **38.** $f(x) = 2x - 6$ **39.** $f(x) = 3x^2 + 1$ **40.** $f(x) = 2x^2 + 1$

41. $f(x) = x^2 + 2x$ **42.** $f(x) = x^2 - 4x$ **43.** $f(x) = 2x^2 - x + 1$ **44.** $f(x) = 2x^2 + 3x - 2$

In Problems 45–54, find the derivative of each function at the given number using a graphing utility.

45. $f(x) = 3x^3 - 6x^2 + 2$ at -2

46. $f(x) = -5x^4 + 6x^2 - 10$ at 5

47. $f(x) = \dfrac{-x^3 + 1}{x^2 + 5x + 7}$ at 8

48. $f(x) = \dfrac{-5x^4 + 9x + 3}{x^2 + 5x^2 - 6}$ at -3

49. $f(x) = xe^x$ at 0 **50.** xe^x at 1 **51.** $f(x) = x^2 e^x$ at 1

52. $f(x) = x^2 e^x$ at 0 **53.** $f(x) = xe^{-x}$ at 1 **54.** $f(x) = x^2e^{-x}$ at 2

---

55. Does the tangent line to the graph of $y = x^2$ at $(1, 1)$ pass through the point $(2, 5)$?

56. Does the tangent line to the graph of $y = x^3$ at $(1, 1)$ pass through the point $(2, 5)$?

57. A dive bomber is flying from right to left along the graph of $y = x^2$. When a rocket bomb is released, it follows a path that approximately follows the tangent line. Where should the pilot release the bomb if the target is at $(1, 0)$?

58. Answer the question in Problem 57 if the plane is flying from right to left along the graph of $y = x^3$.

59. Ticket Sales The cumulative ticket sales for the 10 days preceding a popular concert is given by

$$S = 4x^2 + 50x + 5000$$

where x represents the 10 days leading up to the concert, $1 \leq x \leq 10$.

(a) What is the average rate of change in sales from day 1 to day 5?
(b) What is the average rate of change in sales from day 1 to day 10?
(c) What is the average rate of change in sales from day 5 to day 10?
(d) What is the instantaneous rate of change in sales on day 5?
(e) What is it on day 10?

60. Computer Sales The weekly revenue R, in dollars, due to selling x computers is

$$R(x) = -20x^2 + 1000x$$

(a) Find the average rate of change in revenue due to selling 5 additional computers after the 20th has been sold.
(b) Find the marginal revenue.
(c) Find the marginal revenue at $x = 20$.
(d) Interpret the answers found in (a) and (c).
(e) For what value of x is $R'(x) = 0$?

61. Supply and Demand Suppose $S(x) = 50x^2 - 50x$ is the supply function describing the number of crates of grapefruit a farmer is willing to supply to the market for x dollars per crate.

(a) How many crates is the farmer willing to supply for \$10 per crate?
(b) How many crates is the farmer willing to supply for \$13 per crate?
(c) Find the average rate of change in supply from \$10 per crate to \$13 per crate.
(d) Find the instantaneous rate of change in supply at $x = 10$.
(e) Interpret the answers found in (c) and (d).

62. Glucose Conversion In a metabolic experiment, the mass M of glucose decreases over time t according to the formula

$$M = 4.5 - 0.03t^2$$

(a) Find the average rate of change of the mass from $t = 0$ to $t = 2$.
(b) Find the instantaneous rate of change of mass at $t = 0$.
(c) Interpret the answers found in (a) and (b).

63. Cost and Revenue Functions For a certain production facility, the cost function is

$$C(x) = 2x + 5$$

and the revenue function is

$$R(x) = 8x - x^2$$

where x is the number of units produced (in thousands) and R and C are measured in millions of dollars. Find:

(a) The marginal revenue.
(b) The marginal cost.
(c) The break-even point(s) [the number(s) x for which $R(x) = C(x)$].
(d) The number x for which marginal revenue equals marginal cost.
(e) Graph $C(x)$ and $R(x)$ on the same set of axes.

64. Cost and Revenue Functions For a certain production facility, the cost function is

$$C(x) = x + 5$$

and the revenue function is

$$R(x) = 12x - 2x^2$$

where x is the number of units produced (in thousands) and R and C are measured in millions of dollars. Find:

(a) The marginal revenue.
(b) The marginal cost.
(c) The break-even point(s) [the number(s) x for which $R(x) = C(x)$].
(d) The number x for which marginal revenue equals marginal cost.
(e) Graph $C(x)$ and $R(x)$ on the same set of axes.

65. Demand Equation The price p per ton of cement when x tons of cement are demanded is given by the equation

$$p = -10x + 2000$$

dollars. Find:

(a) The revenue function $R = R(x)$ (*Hint:* $R = xp$, where p is the unit price.)
(b) The marginal revenue.
(c) The marginal revenue at $x = 100$ tons.
(d) The average rate of change in revenue from $x = 100$ to $x = 101$ tons.

66. Demand Equation The cost function and demand equation for a certain product are

$$C(x) = 50x + 40{,}000 \quad \text{and} \quad p = 100 - 0.01x$$

Find:

(a) The revenue function.
(b) The marginal revenue.
(c) The marginal cost.
(d) The break-even point(s).
(e) The number x for which marginal revenue equals marginal cost.

67. Demand Equation A certain item can be produced at a cost of \$10 per unit. The demand equation for this item is

$$p = 90 - 0.02x$$

where p is the price in dollars and x is the number of units. Find:

(a) The revenue function.
(b) The marginal revenue.
(c) The marginal cost.
(d) The break-even point(s).
(e) The number x for which marginal revenue equals marginal cost.

68. Instantaneous Rate of Change A circle of radius r has area $A = \pi r^2$ and circumference $C = 2\pi r$. If the radius changes from r to $(r + h)$, find the:

(a) Change in area.
(b) Change in circumference.
(c) Average rate of change of area with respect of the radius.
(d) Average rate of change of the circumference with respect to the radius.
(e) Instantaneous rate of change of area with respect to the radius.
(f) Instantaneous rate of change of the circumference with respect to the radius.

69. Instantaneous Rate of Change The volume V of a right circular cylinder of height 3 feet and radius r feet is $V = V(r) = 3\pi r^2$. Find the instantaneous rate of change of the volume with respect to the radius r at $r = 3$.

70. Instantaneous Rate of Change The surface area S of a sphere of radius r feet is $S = S(r) = 4\pi r^2$. Find the instantaneous rate of change of the surface area with respect to the radius r at $r = 2$.

4.2 The Derivative of a Power Function; Sum and Difference Formulas

OBJECTIVES
1. Find the derivative of a power function
2. Find the derivative of a constant times a function
3. Find the derivative of a polynomial function

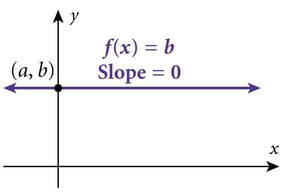
In the previous section, we found the derivative $f'(x)$ of a function $y = f(x)$ by using the difference quotient:

$$\boxed{f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}} \qquad (1)$$

We use this form for the derivative to derive formulas for finding derivatives.

We begin by considering the constant function $f(x) = b$, where b is a real number. Since the graph of the constant function f is a horizontal line (see Figure 8), the tangent line to f at any point is also a horizontal line. Since the derivative equals the slope of the tangent line to the graph of a function f at a point, then the derivative of f should be 0.

FIGURE 8



Algebraically, the derivative is obtained by using formula (1). The difference quotient of $f(x) = b$ is

$$\frac{f(x+h) - f(x)}{h} = \frac{b - b}{h} = \frac{0}{h} = 0$$

The derivative of $f(x) = b$ is

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} 0 = 0$$

Derivative of the Constant Function

For the constant function $f(x) = b$, the derivative is $f'(x) = 0$. In other words, the derivative of a constant is 0.

Besides the **prime notation** f', there are several other ways to denote the derivative of a function $y = f(x)$. The most common ones are

$$y' \qquad \text{and} \qquad \frac{dy}{dx}$$

The notation $\frac{dy}{dx}$, often referred to as the **Leibniz notation,** may also be written as

$$\frac{dy}{dx} = \frac{d}{dx}(y) = \frac{d}{dx}f(x)$$

where $\frac{d}{dx}f(x)$ is an instruction to compute the derivative of the function f with respect to its independent variable x. A change in the symbol used for the independent variable does not affect the meaning. If $s = f(t)$ is a function of t, then $\frac{ds}{dt}$ is an instruction to differentiate f with respect to t.

In terms of the Leibniz notation, if b is a constant, then

$$\boxed{\frac{d}{dx} b = 0} \tag{2}$$

EXAMPLE 1 Finding the Derivative of *a* Constant Function

(a) If $f(x) = 5$, then $f'(x) = 0$.
(b) If $y = -1.7$, then $y' = 0$.
(c) If $y = \dfrac{2}{3}$, then $\dfrac{dy}{dx} = 0$.
(d) If $s = f(t) = \sqrt{5}$, then $\dfrac{ds}{dt} = f'(t) = 0$.

In subsequent work with derivatives we shall use the prime notation or the Leibniz notation, or sometimes a mixture of the two, depending on which is more convenient.

NOW WORK PROBLEM (1)

Derivative of a Power Function

We now investigate the derivative of the power function $f(x) = x^n$, where n is a positive integer, to see if a pattern appears.

For $f(x) = x$, $n = 1$, we have

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{(x+h) - x}{h} = \lim_{h\to 0} \frac{h}{h} = \lim_{h\to 0} 1 = 1$$

For $f(x) = x^2$, $n = 2$, we have

$$\begin{aligned} f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} &= \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h\to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} \\ &= \lim_{h\to 0} \frac{2xh + h^2}{h} = \lim_{h\to 0} \frac{h(2x+h)}{h} \\ &= \lim_{h\to 0} (2x + h) = 2x \end{aligned}$$

For $f(x) = x^3$, $n = 3$, we have

$$\begin{aligned} f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} &= \lim_{h\to 0} \frac{(x+h)^3 - x^3}{h} \\ &= \lim_{h\to 0} \frac{x^3 + 3x^2h + 3xh^2 + h^3 - x^3}{h} \\ &= \lim_{h\to 0} \frac{3x^2h + 3xh^2 + h^3}{h} \\ &= \lim_{h\to 0} \frac{h(3x^2 + 3xh + h^2)}{h} \\ &= \lim_{h\to 0} (3x^2 + 3xh + h^2) = 3x^2 \end{aligned}$$

In the Leibniz notation, these results take the form

$$\frac{d}{dx}x = 1 \qquad \frac{d}{dx}x^2 = 2x \qquad \frac{d}{dx}x^3 = 3x^2$$

This pattern suggests the following formula:

Derivative of $f(x) = x^n$

For the power function $f(x) = x^n$, n a positive integer, the derivative is $f'(x) = nx^{n-1}$. That is,

$$\boxed{\frac{d}{dx}x^n = nx^{n-1}} \tag{3}$$

Formula (3) may be stated in words as

The derivative with respect to x of x raised to the power n, where n is a positive integer, is n times x raised to power $n - 1$.

Problems 68 and 69 outline proofs of Formula (3).

1 **EXAMPLE 2 Finding the Derivative of a Power Function**

(a) If $f(x) = x^6$, then $f'(x) = 6x^{6-1} = 6x^5$

(b) $\frac{d}{dt}t^5 = 5t^4$ **(c)** $\frac{d}{dx}x = 1$

NOW WORK PROBLEM 3.

EXAMPLE 3 Finding the Derivative of a Power Function at a Number

Find $f'(4)$ if $f(x) = x^3$

SOLUTION We use formula (3)

$$f'(x) = 3x^2 \qquad \text{Formula (3)}$$
$$f'(4) = 3(4)^2 = 48 \qquad \text{Substitute 4 for } x.$$

Formula (3) allows us to compute some derivatives with ease. However, do not forget that a derivative is, in actuality, the limit of a difference quotient.

The next formula is used often.

Derivative of a Constant Times a Function

The derivative of a constant times a function equals the constant times the derivative of the function. That is, if C is a constant and f is a differentiable function, then

$$\boxed{\frac{d}{dx}[Cf(x)] = C\frac{d}{dx}f(x)} \tag{4}$$

Proof We prove formula (4) as follows.

$$\begin{aligned}\frac{d}{dx}Cf(x) &= \lim_{h\to 0}\frac{Cf(x+h)-Cf(x)}{h}\\ &= \lim_{h\to 0}C\frac{f(x+h)-f(x)}{h}\\ &= C\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\\ &= C\frac{d}{dx}f(x)\end{aligned}$$

The usefulness and versatility of this formula are often overlooked, especially when the constant appears in the denominator. Note that

$$\frac{d}{dx}\left[\frac{f(x)}{C}\right] = \frac{d}{dx}\left[\frac{1}{C}f(x)\right] = \frac{1}{C}\frac{d}{dx}[f(x)]$$

Always be on the lookout for constant factors *before* differentiating.

2 **EXAMPLE 4** **Finding the Derivative of a Constant Times a Function**

(a) If $f(x) = 10x^3$, then

$$f'(x) = \frac{d}{dx}10x^3 = 10\frac{d}{dx}x^3 = 10\cdot 3x^2 = 30x^2$$

(b) $\dfrac{d}{dx}\dfrac{x^5}{10} = \dfrac{1}{10}\dfrac{d}{dx}x^5 = \dfrac{1}{10}\cdot 5x^4 = \dfrac{1}{2}x^4$

(c) $\dfrac{d}{dt}6t = 6\dfrac{d}{dt}t = 6\cdot 1 = 6$

(d) $\dfrac{d}{dx}\dfrac{2\sqrt{3}}{3}x^3 = \dfrac{2\sqrt{3}}{3}\dfrac{d}{dx}x^3 = \dfrac{2\sqrt{3}}{3}\cdot 3x^2 = 2\sqrt{3}\,x^2$

NOW WORK PROBLEM 7.

Sum and Difference Formulas

Derivative of a Sum

The derivative of the sum of two differentiable functions equals the sum of their derivatives. That is,

$$\boxed{\frac{d}{dx}[f(x)+g(x)] = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)} \tag{5}$$

A proof is given at the end of this section.

This formula for differentiating states that functions that are sums can be differentiated "term by term."

EXAMPLE 5 Finding the Derivative of a Function

Find the derivative of: $f(x) = x^2 + 4x$

SOLUTION The function f is the sum of the two power functions x^2 and $4x$. We can differentiate term by term.

$$\frac{d}{dx}f(x) = \frac{d}{dx}(x^2 + 4x) \underset{\text{Formula (5)}}{=} \frac{d}{dx}x^2 + \frac{d}{dx}(4x) \underset{\text{Formulas (3) and (4)}}{=} 2x + 4\frac{d}{dx}x \underset{\frac{d}{dx}x = 1}{=} 2x + 4$$

Derivative of a Difference

The derivative of the difference of two differentiable functions equals the difference of their derivatives. That is,

$$\boxed{\frac{d}{dx}[f(x) - g(x)] = \frac{d}{dx}f(x) - \frac{d}{dx}g(x)} \qquad (6)$$

Formulas (5) and (6) extend to sums and differences of more than two functions. Since a polynomial function is a sum (or difference) of power functions, we can find the derivative of any polynomial function by using a combination of Formulas (3), (4), (5), and (6).

3

EXAMPLE 6 Finding the Derivative of a Polynomial Function

Find the derivative of: $f(x) = 6x^4 - 3x^2 + 10x - 8$

SOLUTION

$$\begin{aligned} f'(x) &= \frac{d}{dx}(6x^4 - 3x^2 + 10x - 8) \\ &= \frac{d}{dx}(6x^4) - \frac{d}{dx}(3x^2) + \frac{d}{dx}(10x) - \frac{d}{dx}8 && \text{Use Formulas (5) and (6).} \\ &= 6\frac{d}{dx}x^4 - 3\frac{d}{dx}x^2 + 10\frac{d}{dx}x - 0 && \text{Use Formulas (2) and (4).} \\ &= 24x^3 - 6x + 10 && \text{Use Formula (3); Simplify.} \end{aligned}$$

NOW WORK PROBLEM 21.

EXAMPLE 7 Finding the Derivative of a Polynomial Function

If $f(x) = -\frac{x^4}{2} - 2x + 3$, find

(a) $f'(x)$ **(b)** $f'(-1)$

SOLUTION

(a) $f'(x) = -\frac{4x^3}{2} - 2 + 0 = -2x^3 - 2$

(b) $f'(-1) = -2(-1)^3 - 2 = 0$

Proof of the Sum Formula We verify Formula (5) as follows. To compute

$$\frac{d}{dx}[f(x) + g(x)]$$

we need to find the limit of the difference quotient of $f(x) + g(x)$.

$$\begin{aligned}\frac{d}{dx}[f(x) + g(x)] &= \lim_{h\to 0}\frac{[f(x+h) + g(x+h)] - [f(x) + g(x)]}{h}\\ &= \lim_{h\to 0}\frac{[f(x+h) - f(x)] + [g(x+h) - g(x)]}{h}\\ &= \lim_{h\to 0}\left[\frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h}\right]\\ &= \lim_{h\to 0}\left[\frac{f(x+h) - f(x)}{h}\right] + \lim_{h\to 0}\left[\frac{g(x+h) - g(x)}{h}\right]\\ &= \frac{d}{dx}f(x) + \frac{d}{dx}g(x)\end{aligned}$$

Proof of the Difference Formula The proof uses Formulas (4) and (5).

$$\begin{aligned}\frac{d}{dx}[f(x) - g(x)] &= \frac{d}{dx}[f(x) + (-1)g(x)]\\ &= \frac{d}{dx}f(x) + \frac{d}{dx}[(-1)g(x)]\\ &= \frac{d}{dx}f(x) + (-1)\frac{d}{dx}g(x)\\ &= \frac{d}{dx}f(x) - \frac{d}{dx}g(x)\end{aligned}$$

EXAMPLE 8 Analyzing a Cost Function

The total daily cost C, in dollars, of producing dishwashers is

$$C(x) = 1000 + 72x - 0.06x^2 \qquad 0 \le x \le 60$$

where x represents the number of dishwashers produced.

(a) Find the total daily cost of producing 50 dishwashers.
(b) Determine the marginal cost function.
(c) Find $C'(50)$ and interpret its meaning.
(d) Use the marginal cost to estimate the cost of producing 51 dishwashers.
(e) Find the actual cost of producing 51 dishwashers. Compare the actual cost of making 51 dishwashers to the estimated cost of producing 51 dishwashers found in part (d).
(f) Determine the actual cost of manufacturing the 51st dishwashwer.
(g) The average cost function is defined as $\overline{C}(x) = \dfrac{C(x)}{x}$, $0 < x \le 60$. Determine the average cost function for producing x dishwashers.
(h) Find the average cost of producing 51 dishwashers.

SOLUTION **(a)** The total daily cost of producing 50 dishwashers is

$$C(50) = 1000 + 72(50) - 0.06(50)^2 = \$4450$$

(b) The marginal cost is

$$C'(x) = \frac{d}{dx}(72x - 0.06x^2) = 72 - 0.12x$$

(c) $C'(50) = 72 - 0.12(50) = \$66$
The marginal cost of producing 50 dishwashers may be interpreted as the cost to produce the 51st dishwasher.

(d) From part (a) the cost to produce 50 dishwashers is \$4450. If the 51st costs \$66, then the cost to produce 51 will be

$$\$4450 + \$66 = \$4516$$

(e) The actual cost to produce 51 dishwashers is

$$C(51) = \$1000 + 72(51) - 0.06(51)^2 = \$4515.90$$

There is a \$0.10 difference between the actual cost and the cost obtained using the marginal cost.

(f) The actual cost of producing the 51st dishwasher is

$$C(51) - C(50) = \$4515.90 - \$4450 = \$65.90$$

(g) The average cost function is

$$\overline{C}(x) = \frac{c(x)}{x} = \frac{1000 + 72x - 0.06x^2}{x} = \frac{1000}{x} + 72 - 0.06x$$

(h) The average cost of producing 51 dishwashers is

$$\overline{C}(51) = \frac{1000}{51} + 72 - 0.06(51) = \$88.55$$

NOW WORK PROBLEM 65.

EXERCISE 4.2 Answers to Odd-Numbered Problems Begin on Page AN-XX.

In Problems 1–20, find the derivative of each function.

1. $f(x) = 4$
2. $f(x) = -2$
3. $f(x) = x^3$
4. $f(x) = x^4$
5. $f(x) = 6x^2$
6. $f(x) = -8x^3$
7. $f(t) = \frac{t^4}{4}$
8. $f(t) = \frac{t^3}{6}$
9. $f(x) = x^2 + x$
10. $f(x) = x^2 - x$
11. $f(x) = x^3 - x^2 + 1$
12. $f(x) = x^4 - x^3 + x$
13. $f(t) = 2t^2 - t + 4$
14. $f(t) = 3t^3 - t^2 + t$
15. $f(x) = \frac{1}{2}x^8 + 3x + \frac{2}{3}$
16. $f(x) = \frac{2}{3}x^6 - \frac{1}{2}x^4 + 2$
17. $f(x) = \frac{1}{3}(x^5 - 8)$
18. $f(x) = \frac{x^3 + 2}{5}$
19. $f(x) = ax^2 + bx + c$
a, b, c are constants
20. $f(x) = ax^3 + bx^2 + cx + d$
a, b, c, d are constants

In Problems 21–28, find the indicated derivative.

21. $\frac{d}{dx}(-6x^2 + x + 4)$
22. $\frac{d}{dx}(8x^3 - 6x^2 + 2x)$
23. $\frac{d}{dt}(-16t^2 + 80t)$
24. $\frac{d}{dt}(-16t^2 + 64t)$
25. $\frac{dA}{dr}$ if $A = \pi r^2$
26. $\frac{dC}{dr}$ if $C = 2\pi r$
27. $\frac{dV}{dr}$ if $V = \frac{4}{3}\pi r^3$
28. $\frac{dP}{dt}$ if $P = 0.2t$

In Problems 29–38, find the value of the derivative at the indicated number.

29. $f(x) = 4x^2$ at $x = -3$

30. $f(x) = -10x^3$ at $x = -2$

31. $f(x) = 2x^2 - x$ at $x = 4$

32. $f(x) = x^4 - 2x^2$ at $x = 2$

33. $f(t) = -\frac{1}{3}t^3 + 5t$ at $t = 3$

34. $f(t) = -\frac{1}{4}t^4 + \frac{1}{2}t^2 + 4$ at $t = 1$

35. $f(x) = \frac{1}{2}(x^6 - x^4)$ at $x = 1$

36. $f(x) = \frac{1}{3}(x^6 + x^3 + 1)$ at $x = -1$

37. $f(x) = ax^2 + bx + c$ at $x = -b/2a$
a, b, c are constants

38. $f(x) = ax^3 + bx^2 + cx + d$ at $x = 0$
a, b, c, d are constants

In Problems 39–48, find the value of $\frac{dy}{dx}$ *at the indicated point.*

39. $y = x^4$ at $(1, 1)$

40. $y = x^4$ at $(2, 16)$

41. $y = x^2 - 14$ at $(4, 2)$

42. $y = x^3 + 1$ at $(3, 28)$

43. $y = 3x^2 - x$ at $(-1, 4)$

44. $y = x^2 - 3x$ at $(-1, 4)$

45. $y = \frac{1}{2}x^2$ at $\left(1, \frac{1}{2}\right)$

46. $y = x^3 - x^2$ at $(1, 0)$

47. $y = 2 - 2x + x^3$ at $(2, 6)$

48. $y = 2x^2 - \frac{1}{2}x + 3$ at $(0, 3)$

In Problems 49–50, find the slope of the tangent line to the graph of the function at the indicated point. What is an equation of the tangent line?

49. $f(x) = x^3 + 3x - 1$ at $(0, -1)$

50. $f(x) = x^4 + 2x - 1$ at $(1, 2)$

In Problems 51–56, find those x, if any, at which $f'(x) = 0$.

51. $f(x) = 3x^2 - 12x + 4$

52. $f(x) = x^2 + 4x - 3$

53. $f(x) = x^3 - 3x + 2$

54. $f(x) = x^4 - 4x^3$

55. $f(x) = x^3 + x$

56. $f(x) = x^5 - 5x^4 + 1$

---

57. Find the point(s), if any, on the graph of the function $y = 9x^3$ at which the tangent line is parallel to the line $3x - y + 2 = 0$.

58. Find the points(s), if any, on the graph of the function $y = 4x^2$ at which the tangent line is parallel to the line $2x - y - 6 = 0$.

59. Two lines through the point $(1, -3)$ are tangent to the graph of the function $y = 2x^2 - 4x + 1$. Find the equations of these two lines.

60. Two lines through the point $(0, 2)$ are tangent to the graph of the function $y = 1 - x^2$. Find the equations of these two lines.

61. Marginal Cost The cost per day, $C(x)$, in dollars, of producing x pairs of eyeglasses is

$$C(x) = 0.2x^2 + 3x + 1000$$

(a) Find the average cost due to producing 10 additional pairs of eyeglasses after 100 have been produced.
(b) Find the marginal cost.
(c) Find the marginal cost at $x = 100$.
(d) Interpret $C'(100)$.

62. Toy Truck Sales At Dan's Toy Store, the revenue R, in dollars, derived from selling x electric trucks is

$$R(x) = -0.005x^2 + 20x$$

(a) What is the average rate of change in revenue due to selling 10 additional trucks after 1000 have been sold?
(b) What is the marginal revenue?
(c) What is the marginal revenue at $x = 1000$?
(d) Interpret $R'(1000)$.
(e) For what value of x is $R'(x) = 0$?

63. Medicine The French physician Poiseville discovered that the volume V of blood (in cubic centimeters) flowing through a clogged artery with radius R (in centimeters) can be modeled by

$$V(R) = kR^4$$

where k is a positive constant.

(a) Find the derivative $V'(R)$.
(b) Find the rate of change of volume for a radius of 0.3 cm.
(c) Find the rate of change of volume for a radius of 0.4 cm.
(d) If the radius of a clogged artery is increased from 0.3 cm to 0.4 cm, estimate the effect on the volume of blood flowing through the enlarged artery.

64. Respiration Rate A human being's respiration rate R (in breaths per minute) is given by

$$R = -10.35p + 0.59p^2$$

where p is the partial pressure of carbon dioxide in the lungs. Find the rate of change in respiration rate when $p = 50$.

65. Analyzing a Cost Function The total daily cost C of producing microwave ovens is

$$C(x) = 2000 + 50x - 0.05x^2, 0 \le x \le 50$$

where x represents the number of microwave ovens produced.

(a) Find the total daily cost of producing 40 microwave ovens.
(b) Determine the marginal cost function.
(c) Find $C'(40)$ and interpret its meaning.
(d) Use the marginal cost to estimate the cost of producing 41 microwave ovens.
(e) Find the actual cost of producing 41 microwave ovens. Compare the actual cost of making 51 microwave ovens to the estimated cost of producing 51 microwave ovens.
(f) Determine the actual cost of manufacturing the 51st microwave oven.
(g) The average cost function is defined as $\overline{C}(x) = \dfrac{C(x)}{x}$, $0 < x \le 50$. Determine the average cost function for producing x microwave ovens.
(h) Find the average cost of producing 51 microwave ovens.
(i) Compare your answers from parts (g), (h), and (j). Give explanations for the differences.

66. Analyzing a Cost Function The total daily cost C of producing small televisions is

$$C(x) = 1500 + 25x - 0.05x^2, 0 \le x \le 100$$

where x represents the number of televisions produced.

(a) Find the total daily cost of producing 70 televisions.
(b) Determine the marginal cost function.
(c) Find $C'(70)$ and interpret its meaning.
(d) Use the marginal cost to estimate the cost of producing 71 televisions.
(e) Find the actual cost of producing 71 televisions. Compare the actual cost of making 71 televisions to the estimated cost of producing 71 televisions.
(f) Determine the actual cost of manufacturing the 71st televisions.
(g) The average cost function is defined as $\overline{C}(x) = \dfrac{C(x)}{x}$, $0 < x \le 100$. Determine the average cost function for producing x televisions.
(h) Find the average cost of producing 71 televisions.
(i) Compare your answers from parts (g), (h), and (j). Give explanations for the differences.

65. Price of Beans The price per unit in dollars per cwt for beans from 1993 through 2002 can be modeled by the polynomial function $p(t) = 0.007t^3 - 0.63t^2 + 0.005t + 6.123$, where t is in years, and $t = 0$ corresponds to 1993.

(a) Find the marginal price of beans for the year 1995.
(b) Find the marginal price for beans for the year 2002.
(c) How do you interpret the two marginal prices? What is the trend?

66. Price of Beans price per unit in dollars per cwt for beans from 1993 through 2002 can also be modeled by the polynomial function, $p(t) = -0.002t^4 + 0.044t^3 - 0.335t^2 + 0.750t + 5.543$, where t is in years and $t = 0$ corresponds to 1993.

(a) Find the marginal price for beans for the year 1995.
(b) Find the marginal price for beans for the year 2002.
(c) How do you interpret the two marginal prices? What is the trend?
(d) Explain why there might be two different functions that can model the price of beans.

67. Instantaneous Rate of Change The volume V of a sphere of radius r feet is $V = V(r) = \frac{4}{3}\pi r^3$. Find the instantaneous rate of change of the volume with respect to the radius r at $r = 2$.

68. Instantaneous Rate of Change The volume V of a cube of side x meters is $V = V(x) = x^3$. Find the instantaneous rate of change of the volume with respect to the side x at $x = 3$.

69. Work Output The relationship between the amount $A(t)$ of work output and the elapsed time t, $t \ge 0$, was found through empirical means to be

$$A(t) = a_3t^3 + a_2t^2 + a_1t + a_0$$

where a_0, a_1, a_2, a_3 are constants. Find the instantaneous rate of change of work output at time t.

70. Consumer Price Index The consumer price index (CPI) of an economy is described by the function

$$I(t) = -0.2t^2 + 3t + 200 \qquad 0 \le t \le 10$$

where $t = 0$ corresponds to the year 2000.

(a) What was the average rate of increase in the CPI over the period from 2000 to 2003?
(b) At what rate was the CPI of the economy changing in 2003? in 2006?

71. Use the binomial theorem to prove formula (3)

[**Hint:** $(x + h)^n - x^n = x^n + nx^{n-1}h + \dfrac{n(n-1)}{2}x^{n-2}h^2 + \cdots + h^n - x^n = nx^{n-1}h + h^2 \cdot$ (terms involving x and h). Now apply formula (1), page 000.]

72. Use the following factoring rule to prove formula (3)

$$f(x) = x^n - c^n = (x - c)(x^{n-1} + x^{n-2}c + x^{n-3}c^2 + \cdots + c^{n-1})$$

Now apply Formula (3), page 000, to find $f'(c)$.

4.3 Product and Quotient Formulas

OBJECTIVES

1. Find the derivative of a product
2. Find the derivative of a quotient
3. Find the derivative of $f(x) = x^n$, n is a negative integer

The Derivative of a Product

In the previous section we learned that the derivative of the sum or the difference of two functions is simply the sum or the difference of their derivatives. The natural inclination at this point may be to assume that differentiating a product or quotient of two functions is as simple. But this is not the case, as illustrated for the case of a product of two functions. Consider

$$F(x) = f(x) \cdot g(x) = (3x^2 - 3)(2x^3 - x) \tag{1}$$

where $f(x) = 3x^2 - 3$, and $g(x) = 2x^3 - x$. The derivative of $f(x)$ is $f'(x) = 6x$ and the derivative of $g(x)$ is $g'(x) = 6x^2 - 1$. The product of these derivatives is

$$f'(x) \cdot g'(x) = 6x(6x^2 - 1) = 36x^3 - 6x \tag{2}$$

To see if this is equal to the derivative of the product, we first multiply the right side of (1) and then differentiate using the rules of differentiation of the previous section:

$$F(x) = (3x^2 - 3)(2x^3 - x) = 6x^5 - 9x^3 + 3x$$

so that

$$F'(x) = 30x^4 - 27x^2 + 3 \tag{3}$$

Since (2) and (3) are not equal, we conclude that the derivative of a product *is not* equal to the product of the derivatives.

The formula for finding the derivative of the product of two functions is given below:

Derivative of a Product

The derivative of the product of two differentiable functions equals the first function times the derivative of the second plus the second function times the derivative of the first. That is,

$$\boxed{\frac{d}{dx}[f(x)g(x)] = f(x)\frac{d}{dx}g(x) + g(x)\frac{d}{dx}f(x)} \tag{4}$$

The following version of formula (4) may help you remember it.

$$\boxed{\frac{d}{dx}(\text{first} \cdot \text{second}) = \text{first} \cdot \frac{d}{dx}\text{second} + \text{second} \cdot \frac{d}{dx}\text{first}}$$

1 **EXAMPLE 1 Finding the Derivative of a Product**

Find the derivative of: $F(x) = (x^2 + 2x - 5)(x^3 - 1)$

SOLUTION The function F is the product of the two functions $f(x) = x^2 + 2x - 5$ and $g(x) = x^3 - 1$ so that, by (1), we have

$$\begin{aligned} F'(x) &= (x^2 + 2x - 5)\left[\frac{d}{dx}(x^3 - 1)\right] + (x^3 - 1)\left[\frac{d}{dx}(x^2 + 2x - 5)\right] && \text{Use formula (4).} \\ &= (x^2 + 2x - 5)(3x^2) + (x^3 - 1)(2x + 2) && \text{Differentiate.} \\ &= 3x^4 + 6x^3 - 15x^2 + 2x^4 + 2x^3 - 2x - 2 && \text{Simplify.} \\ &= 5x^4 + 8x^3 - 15x^2 - 2x - 2 && \text{Simplify.} \end{aligned}$$

Now that you know the formula for the derivative of a product, be careful not to use it unnecessarily. When one of the factors is a constant, you should use the formula for the derivative of a constant times a function. For example, it is easier to work

$$\frac{d}{dx}[5(x^2 + 1)] = 5\frac{d}{dx}(x^2 + 1) = (5)(2x) = 10x$$

than it is to work

$$\begin{aligned} \frac{d}{dx}[5(x^2 + 1)] &= 5\left[\frac{d}{dx}(x^2 + 1)\right] + (x^2 + 1)\left(\frac{d}{dx}5\right) \\ &= (5)(2x) + (x^2 + 1)(0) = 10x \end{aligned}$$

NOW WORK PROBLEM 1.

The Derivative of a Quotient

As in the case with a product, the derivative of a quotient is *not* the quotient of the derivatives.

Derivative of a Quotient

The derivative of the quotient of two differentiable functions is equal to the denominator times the derivative of the numerator minus the numerator times the derivative of the denominator, all divided by the square of the denominator.

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)\dfrac{d}{dx}f(x) - f(x)\dfrac{d}{dx}g(x)}{[g(x)]^2} \quad \text{where } g(x) \neq 0 \qquad (5)$$

You may want to memorize the following version of Formula (5):

$$\frac{d}{dx}\,\frac{\textbf{numerator}}{\textbf{denominator}}$$

$$= \frac{(\textbf{denominator})\dfrac{d}{dx}(\textbf{numerator}) - (\textbf{numerator})\dfrac{d}{dx}(\textbf{denominator})}{(\textbf{denominator})^2}$$

2

EXAMPLE 2 Finding the Derivative of a Quotient

Find the derivative of: $F(x) = \dfrac{x^2 + 1}{x - 3}$

SOLUTION Here, the function F is the quotient of $f(x) = x^2 + 1$ and $g(x) = x - 3$. We use Formula (5) to get

$$\begin{aligned}
\frac{d}{dx}\left(\frac{x^2+1}{x-3}\right) &= \frac{(x-3)\dfrac{d}{dx}(x^2+1) - (x^2+1)\dfrac{d}{dx}(x-3)}{(x-3)^2} && \text{Use formula (5).}\\
&= \frac{(x-3)(2x) - (x^2+1)(1)}{(x-3)^2} && \text{Differentiate.}\\
&= \frac{2x^2 - 6x - x^2 - 1}{(x-3)^2} && \text{Simplify.}\\
&= \frac{x^2 - 6x - 1}{(x-3)^2} && \text{Simplify.}
\end{aligned}$$

NOW WORK PROBLEM 9.

We shall follow the practice of leaving our answers in factored form as shown in Example 2.

EXAMPLE 3 Finding the Derivative of a Quotient

Find the derivative of: $y = \dfrac{(1 - 3x)(2x + 1)}{3x - 2}$

SOLUTION We shall solve the problem in two ways.

Method 1 Use the formula for the derivative of a quotient right away.

$$\begin{aligned}
y' &= \frac{d}{dx}\,\frac{(1-3x)(2x+1)}{3x-2}\\
&= \frac{(3x-2)\dfrac{d}{dx}[(1-3x)(2x+1)] - (1-3x)(2x+1)\dfrac{d}{dx}(3x-2)}{(3x-2)^2} && \text{Use Formula (5).}\\
&= \frac{(3x-2)[(1-3x)\dfrac{d}{dx}(2x+1) + (2x+1)\dfrac{d}{dx}(1-3x)] - (1-3x)(2x+1)\dfrac{d}{dx}(3x-2)}{(3x-2)^2} && \text{Differentiate.}
\end{aligned}$$

$$= \frac{(3x-2)[(1-3x)(2)+(2x+1)(-3)]-(1-3x)(2x+1)(3)}{(3x-2)^2} \qquad \text{Using Formula (4) Differentiate.}$$

$$= \frac{(3x-2)[2-6x-6x-3]-(-6x^2-x+1)(3)}{(3x-2)^2} \qquad \text{Simply.}$$

$$= \frac{(3x-2)(-12x-1)-(-18x^2-3x+3)}{(3x-2)^2} \qquad \text{Simplify.}$$

$$= \frac{-36x^2+21x-2+18x^2+3x-3}{(3x-2)^2} \qquad \text{Simplify.}$$

$$= \frac{-18x^2+24x-1}{(3x-2)^2}$$

Method 2 First, multiply the factors in the numerator and then apply the formula for the derivative of a quotient.

$$y = \frac{(1-3x)(2x+1)}{3x-2} = \frac{-6x^2-x+1}{3x-2}$$

Now use Formula (5):

$$y' = \frac{d}{dx}\,\frac{-6x^2-x+1}{3x-2}$$

$$= \frac{(3x-2)\dfrac{d}{dx}(-6x^2-x+1)-(-6x^2-x+1)\dfrac{d}{dx}(3x-2)}{(3x-2)^2} \qquad \text{Formula (5).}$$

$$= \frac{(3x-2)(-12x-1)-(-6x^2-x+1)(3)}{(3x-2)^2} \qquad \text{Differentiate.}$$

$$= \frac{-36x^2+21x+2+18x^2+3x-3}{(3x-2)^2} \qquad \text{Simplify.}$$

$$= \frac{-18x^2+24x-1}{(3x-2)^2} \qquad \text{Simplify.}$$

As you can see from this example, looking at alternative methods may make the differentiation easier.

The Derivative of x^n, n a Negative Integer

3 Find the derivative of $f(x) = x^n$, n a negative integer

In the previous section, we learned that the derivative of a power function $f(x) = x^n$, $n \geq 1$ an integer, is $f'(x) = nx^{n-1}$.

The formula for the derivative of x raised to a negative integer exponent follows the same form.

The derivative of $f(x) = x^n$, where n is *any* integer, is n times x to the $n - 1$ power. Thus,

$$\boxed{\frac{d}{dx}x^n = nx^{n-1} \qquad \textbf{for any integer } n} \qquad (3)$$

EXAMPLE 4 Using Formula (3)

The proof is left as an exercise. See Problem 54.

(a) $\dfrac{d}{dx}x^{-3} = -3x^{-4} = \dfrac{-3}{x^4}$

(b) $\dfrac{d}{dx}\dfrac{4}{x^2} = \dfrac{d}{dx}4x^{-2} = 4\dfrac{d}{dx}x^{-2} = 4(-2x^{-3}) = \dfrac{-8}{x^3}$

(c) $\dfrac{d}{dx}\left(x + \dfrac{2}{x}\right) = \dfrac{d}{dx}(x + 2x^{-1}) = \dfrac{d}{dx}x + 2\dfrac{d}{dx}x^{-1} = 1 + 2(-1)x^{-1} = 1 - \dfrac{2}{x^2}$

NOW WORK PROBLEM 17.

EXAMPLE 5 Finding the Derivative of a Function

Find the derivative of: $g(x) = \left(1 - \dfrac{1}{x^2}\right)(x + 1)$

SOLUTION Since $g(x)$ is the product of two simpler functions, we begin by applying the formula for the derivative of a product:

$$\begin{aligned}
g'(x) &= \left(1 - \frac{1}{x^2}\right)\frac{d}{dx}(x + 1) + (x + 1)\frac{d}{dx}\left(1 - \frac{1}{x^2}\right) && \text{Derivative of a Product.}\\
&= \left(1 - \frac{1}{x^2}\right)(1) + (x + 1)\frac{d}{dx}(1 - x^{-2}) && \text{Differentiate; } \frac{1}{x^2} = x^{-2}.\\
&= 1 - \frac{1}{x^2} + (x + 1)(2x^{-3}) && \text{Differentiate.}\\
&= 1 - \frac{1}{x^2} + \frac{2(x + 1)}{x^3} && \text{Simplify.}\\
&= 1 - \frac{1}{x^2} + \frac{2x}{x^3} + \frac{2}{x^3} && \text{Simplify.}\\
&= 1 + \frac{1}{x^2} + \frac{2}{x^3}
\end{aligned}$$

Alternatively, we could have solved Example 5 by multiplying the factors first. Then

$$g(x) = \left(1 - \frac{1}{x^2}\right)(x + 1) = x + 1 - \frac{1}{x} - \frac{1}{x^2}$$

so

$$\begin{aligned}
g'(x) &= \frac{d}{dx}\left(x + 1 - \frac{1}{x} - \frac{1}{x^2}\right) = \frac{d}{dx}x + \frac{d}{dx}(1) - \frac{d}{dx}\frac{1}{x} - \frac{d}{dx}\frac{1}{x^2}\\
&= 1 + 0 - \frac{d}{dx}x^{-1} - \frac{d}{dx}x^{-2} = 1 - (-1)x^{-2} - (-2)x^{-3} = 1 + \frac{1}{x^2} + \frac{2}{x^3}
\end{aligned}$$

EXAMPLE 4 Application

The value $V(t)$, in dollars, of a car t years after its purchase is given by the equation

$$V(t) = \frac{8000}{t} + 5000 \qquad 1 \le t \le 5$$

Graph the function $V = V(t)$. Then find:

Art to come

(a) The average rate of change in value from $t = 1$ to $t = 4$.
(b) The instantaneous rate of change in value.
(c) The instantaneous rate of change in value after 1 year.
(d) The instantaneous rate of change in value after 3 years.
(e) Interpret the answers to (c) and (d).

SOLUTION The graph of $V = V(t)$ is given in Figure 9.

(a) The average rate of change in value from $t = 1$ to $t = 4$ is given by

$$\frac{V(4) - V(1)}{4 - 1} = \frac{7000 - 13{,}000}{3} = -2000$$

So the average rate of change in value from $t = 1$ to $t = 4$ is $-\$2000$ per year, that is, it is decreasing at the rate of \$2000 per year.

(b) The derivative $V'(t)$ of $V(t)$ equals the instantaneous rate of change in the value of the car.

$$V'(t) = \frac{d}{dt}\left(\frac{8000}{t} + 5000\right) = \frac{d}{dt}\frac{8000}{t} + \frac{d}{dt}(5000)$$
$$= \frac{d}{dt}8000t^{-1} + 0 = 8000(-1)t^{-2} = -\frac{8000}{t^2}$$

Notice that $V'(t) < 0$; we interpret this to mean that the value of the car is decreasing over time.

(c) After 1 year, $V'(1) = -\dfrac{8000}{1} = -\$8000/\text{year}$

(d) After 3 years, $V'(3) = -\dfrac{8000}{9} = -\$888.89/\text{year}$

(e) $V'(1) = -\$8000$ means that the value of the car after 1 year will decline by approximately \$8000 over the next year; $V'(3) = -\$888.89$ means that the value of the car after 3 years will decline by approximately \$888.89 over the next year.

FIGURE 9

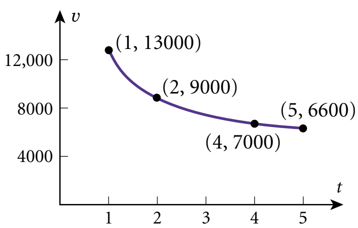


Figure 9 shows the graph of $V = V(t)$.

NOW WORK PROBLEM 41.

SUMMARY

Each of the derivative formulas given so far can be written without reference to the independent variable of the function. If f and g are differentiable functions, we have the following formulas:

Derivative of a constant times a function	$(cf)' = cf'$
Derivative of a sum	$(f + g)' = f' + g'$
Derivative of a difference	$(f - g)' = f' - g'$
Derivative of a product	$(f \cdot g)' = f \cdot g' + g \cdot f'$
Derivative of a quotient	$\left(\dfrac{f}{g}\right)' = \dfrac{g \cdot f' - f \cdot g'}{g^2}$

EXERCISE 4.3 Answers to Odd-Numbered Problems Begin on Page AN-XX.

In Problems 1–8, find the derivative of each function by using the formula for the derivative of a product.

1. $f(x) = (2x + 1)(4x - 3)$
2. $f(x) = (3x - 4)(2x + 5)$
3. $f(t) = (t^2 + 1)(t^2 - 4)$
4. $f(t) = (t^2 - 3)(t^2 + 4)$
5. $f(x) = (3x - 5)(2x^2 + 1)$
6. $f(x) = (3x^2 - 1)(4x + 1)$
7. $f(x) = (x^5 + 1)(3x^3 + 8)$
8. $f(x) = (x^6 - 2)(4x^2 + 1)$

In Problems 9–20, find the derivative of each function.

9. $f(x) = \dfrac{x}{x + 1}$
10. $f(x) = \dfrac{x + 4}{x^2}$
11. $f(x) = \dfrac{3x + 4}{2x - 1}$
12. $f(x) = \dfrac{3x - 5}{4x + 1}$
13. $f(x) = \dfrac{x^2}{x - 4}$
14. $f(x) = \dfrac{x}{x^2 - 4}$
15. $f(x) = \dfrac{2x + 1}{3x^2 + 4}$
16. $f(x) = \dfrac{2x^2 - 1}{5x + 2}$
17. $f(t) = \dfrac{-2}{t^2}$
18. $f(t) = \dfrac{4}{t^3}$
19. $f(x) = 1 + \dfrac{1}{x} + \dfrac{1}{x^2}$
20. $f(x) = 1 - \dfrac{1}{x} + \dfrac{1}{x^2}$

In Problems 21–24, find the slope of the tangent line to the graph of the function f at the indicated point. What is an equation of the tangent line?

21. $f(x) = (x^3 - 2x + 2)(x + 1)$ at $(1, 2)$
22. $f(x) = (2x^2 - 5x + 1)(x - 3)$ at $(1, 4)$
23. $f(x) = \dfrac{x^3}{x + 1}$ at $(1, \frac{1}{2})$
24. $f(x) = \dfrac{x^2}{x - 1}$ at $(-1, -\frac{1}{2})$

In Problems 25–28, find those x, if any, at which $f'(x) = 0$.

25. $f(x) = (x^2 - 2)(2x - 1)$
26. $f(x) = (3x^2 - 3)(2x^3 - x)$
27. $f(x) = \dfrac{x^2}{x + 1}$
28. $f(x) = \dfrac{x^2 + 1}{x}$

In Problems 29–40, find y'.

29. $y = x^2(3x - 2)$
30. $y = (x^2 + 2)(x - 1)$
31. $y = (x^{-2} + 4)(4x^2 + 3)$
32. $y = (2x^{-1} + 3)(x^{-3} + x^{-2})$
33. $y = \dfrac{(2x + 3)(x - 4)}{3x + 5}$
34. $y = \dfrac{(3x - 2)(x^2 + 1)}{4x - 3}$
35. $y = \dfrac{3x + 1}{(x - 2)(x + 2)}$
36. $y = \dfrac{2x - 5}{(1 - x)(1 + x)}$
37. $y = \dfrac{(3x + 4)(2x - 3)}{(2x + 1)(3x - 2)}$
38. $y = \dfrac{(2 - 3x)(1 - x)}{(x + 2)(3x + 1)}$
39. $y = \dfrac{x^{-2} - x^{-1}}{x^{-2} + x^{-1}}$
40. $y = \dfrac{3x^{-4} - x^{-2}}{x^{-3} + x^{-1}}$

---

41. Value of a Car The value V of a luxury car after t years is

$$V(t) = \frac{10{,}000}{t} + 6000 \qquad 1 \le t \le 6$$

(a) What is the average rate of change in value from $t = 2$ to $t = 5$?
(b) What is the instantaneous rate of change in value?
(c) What is the instantaneous rate of change after 2 years?
(d) What is the instantaneous rate of change after 5 years?
(e) Interpret the answers found in (c) and (d).

42. Value of a Painting The value V of a famous painting t years after it is purchased is

$$V(t) = \frac{100t^2 + 50}{t} + 400 \qquad 1 \le t \le 5$$

(a) What is the average rate of change in value from $t = 1$ to $t = 3$?
(b) What is the instantaneous rate of change in value?
(c) What is the instantaneous rate of change after 1 year?
(d) What is the instantaneous rate of change after 3 years?
(e) Interpret the answers found in (c) and (d).

43. Demand Equation The demand equation for a certain commodity is

$$p = 10 + \frac{40}{x} \qquad 1 \le x \le 10$$

where p is the price in dollars when x units are demanded. Find:

(a) The revenue function.
(b) The marginal revenue.
(c) The marginal revenue for $x = 4$.
(d) The marginal revenue for $x = 6$.

44. Cost Function The cost of fuel in operating a luxury yacht is given by the equation

$$C(s) = \frac{-3s^2 + 1200}{s}$$

where s is the speed of the yacht. Find the rate at which the cost is changing when $s = 10$.

45. Price–Demand Function The price–demand function for calculators is given by

$$D(p) = \frac{100{,}000}{p^2 + 10p + 50} \qquad 5 \le p \le 20$$

where D is the quantity demanded per week and p is the unit price in dollars.

(a) Find $D'(p)$, the rate of change of demand with respect to price.
(b) Find $D'(5)$, $D'(10)$, and $D'(15)$
(c) Interpret the results found in part (b).

46. Rising Object The height, in kilometers, that a balloon will rise in t hours is given by the formula

$$s = \frac{t^2}{2 + t}$$

Find the rate at which the balloon is rising after (a) 10 minutes, (b) 20 minutes.

47. Population Growth A population of 1000 bacteria is introduced into a culture and grows in number according to the formula

$$P(t) = 1000\left(1 + \frac{4t}{100 + t^2}\right)$$

where t is measured in hours. Find the rate at which the population is growing when

(a) $t = 1$ (b) $t = 2$ (c) $t = 3$ (d) $t = 4$

48. Drug Concentration The concentration of a certain drug in a patient's bloodstream t hours after injection is given by

$$C(t) = \frac{0.4t}{2t^2 + 1}$$

Find the rate at which the concentration of the drug is changing with respect to time. At what rate is the concentration changing

(a) 10 minutes after the injection?
(b) 30 minutes after the injection?
(c) 1 hour after the injection?
(d) 3 hours after the injection?

49. Intensity of Illumination The intensity of illumination I on a surface is inversely proportional to the square of the distance r from the surface to the source of light. If the intensity is 1000 units when the distance is 1 meter, find the rate of change of the intensity with respect to the distance when the distance is 10 meters.

50. Cost Function The cost, C, in thousands of dollars, for removal of pollution from a certain lake is

$$C(x) = \frac{5x}{110 - x}$$

where x is the percent of pollutant removed. Find:

(a) $C'(x)$, the rate of change of cost with respect to the amount of pollutant removed.
(b) Compute $C'(10)$, $C'(20)$, $C'(70)$, $C'(90)$.

51. Cost Function An airplane crosses the Atlantic Ocean (3000 miles) with an airspeed of 500 miles per hour. The cost C (in dollars) per person is

$$C(x) = 100 + \frac{x}{10} + \frac{36{,}000}{x}$$

where x is the ground speed (airspeed $\pm$ wind). Find:

(a) The marginal cost.
(b) The marginal cost at a ground speed of 500 mph.
(c) The marginal cost at a ground speed of 550 mph.
(d) The marginal cost at a ground speed of 450 mph.

52. Average Cost Function If C is the total cost function then $\overline{C}(x) = \frac{C(x)}{x}$ is defined as the **average cost function,** that is, the cost per unit produced. Typically, the graph of the average cost function has a U-shape. This is so since we expect higher average costs because of plant inefficiency at low output levels and also at high output levels near plant capacity. Suppose a company estimates that the total cost of producing x units of a certain product is given by

$$C(x) = 400 + 0.02x + 0.0001x^2$$

Then the average cost is given by

$$\overline{C}(x) = \frac{C(x)}{x} = \frac{400}{x} + 0.02 + 0.0001x$$

(a) Find the marginal average cost $\overline{C}'(x)$.
(b) Find the marginal average cost at $x = 200$, 300, and 400.
(c) Interpret your results.

53. Satisfaction and Reward The relationship between satisfaction S and total reward r has been found to be

$$S(r) = \frac{ar}{g - r}$$

where $g \geq 0$ is the predetermined goal level and $a > 0$ is the perceived justice per unit of reward. Show that the instantaneous rate of change of satisfaction with respect to reward is inversely proportional to the square of the difference between the personal goal of the individual and the amount of reward received.

54. Prove Formula (3).

Hint: If $n < 0$, then $-n > 0$. Now use the fact that

$$\frac{d}{dx}x^n = \frac{d}{dx}\frac{1}{x^{-n}}$$

and use the quotient formula.

4.4 The Power Rule

OBJECTIVES

1. Find derivatives using the Power Rule
2. Find derivatives using the Power Rule and other formulas

When a function is of the form $y = [g(x)]^n$, n an integer, the formula used to find the derivative y' is called the *Power Rule*. Let's see if we can guess this formula by finding the derivative of $y = [g(x)]^n$ when $n = 2$, $n = 3$, and $n = 4$.

If $n = 2$,

$$\frac{d}{dx}[g(x)]^2 = \frac{d}{dx}[g(x)g(x)] \underset{\substack{\uparrow \\ \text{Product formula}}}{=} g'(x)g(x) + g(x)g'(x) = 2g(x)g'(x)$$

If $n = 3$,

$$\frac{d}{dx}[g(x)]^3 = \frac{d}{dx}\{[g(x)]^2 g(x)\} = [g(x)]^2 g'(x) + g(x)\left\{\frac{d}{dx}[g(x)]^2\right\}$$
$$= [g(x)]^2 g'(x) + g(x)[2g(x)g'(x)] = 3[g(x)]^2 g'(x)$$

If $n = 4$,

$$\frac{d}{dx}[g(x)]^4 = \frac{d}{dx}\{[g(x)]^3 g(x)\} = [g(x)]^3 g'(x) + g(x)\left\{\frac{d}{dx}[g(x)]^3\right\}$$
$$= [g(x)]^3 g'(x) + g(x)\{3[g(x)]^2 g'(x)\} = 4[g(x)]^3 g'(x)$$

Let's summarize what we've found:

$$\frac{d}{dx}[g(x)]^2 = 2g(x)g'(x)$$

$$\frac{d}{dx}[g(x)]^3 = 3[g(x)]^2 g'(x)$$

$$\frac{d}{dx}[g(x)]^4 = 4[g(x)]^3 g'(x)$$

These results suggest the following formula:

The Power Rule

If g is a differentiable function and n is any integer, then

$$\boxed{\frac{d}{dx}[g(x)]^n = n[g(x)]^{n-1}g'(x)} \tag{1}$$

Note the similarity between the Power Rule and the formula for the derivative of a power function:

$$\frac{d}{dx}x^n = nx^{n-1}$$

The main difference between these formulas is the factor $g'(x)$. Be sure to remember to include $g'(x)$ when using formula (1).

1 EXAMPLE 1 Using the Power Rule to Find a Derivative

Find the derivative of the function: $f(x) = (x^2 + 1)^3$

SOLUTION We could, of course, expand the right-hand side and proceed according to techniques discussed earlier. However, the usefulness of the Power Rule is that it enables us to find derivatives of functions like this without resorting to tedious (and sometimes impossible) computation.

The function $f(x) = (x^2 + 1)^3$ is the function $g(x) = x^2 + 1$ raised to the power 3. Using the Power Rule,

$$\frac{d}{dx}f(x) = \frac{d}{dx}(x^2+1)^3 \underset{\substack{\uparrow \\ \text{Use the Power Rule}}}{=} 3(x^2+1)^2\frac{d}{dx}(x^2+1)$$

$$= 3(x^2+1)^2(2x) = 6x(x^2+1)^2$$

NOW WORK PROBLEM 1.

EXAMPLE 2 Using the Power Rule

Find $f'(x)$

(a) $f(x) = \dfrac{1}{(x^3+4)^5}$ **(b)** $f(x) = \dfrac{1}{(x^2+4)^3}$

SOLUTION **(a)** We write $f(x)$ as $f(x) = (x^3 + 4)^{-5}$. Then we use the Power Rule:

$$f'(x) = \frac{d}{dx}(x^3+4)^{-5} \underset{\substack{\uparrow \\ \text{Use the Power Rule}}}{=} -5(x^3+4)^{-6}\frac{d}{dx}(x^3+4)$$

$$= -5(x^3+4)^{-6}(3x^2) = \frac{-15x^2}{(x^3+4)^6}$$

(b) $$\frac{d}{dx}\frac{1}{(x^2+4)^3} = \frac{d}{dx}(x^2+4)^{-3} \underset{\substack{\uparrow \\ \text{Use the Power Rule}}}{=} -3(x^2+4)^{-4}\frac{d}{dx}(x^2+4)$$

$$= -3(x^2+4)^{-4}\cdot 2x = \frac{-6x}{(x^2+4)^4}$$

Often, we must use at least one other formula along with the Power Rule to differentiate a function. Here are two examples.

2 EXAMPLE 3 Using the Power Rule with Other Formulas

Find the derivative of the function: $f(x) = x(x^2+1)^3$

SOLUTION The function f is the product of x and $(x^2+1)^3$. We begin by using the formula for the derivative of a product. That is,

$$f'(x) = x\frac{d}{dx}(x^2+1)^3 + (x^2+1)^3\frac{d}{dx}x \qquad \text{Product formula.}$$

We continue by using the Power Rule:

$$\begin{aligned} f'(x) &= x\left[3(x^2+1)^2\frac{d}{dx}(x^2+1)\right] + (x^2+1)^3\cdot 1 && \text{Power Rule; } \frac{d}{dx}x = 1. \\ &= (x)(3)(x^2+1)^2(2x) + (x^2+1)^3 && \text{Differentiate.} \\ &= (x^2+1)^2(6x^2) + (x^2+1)^2(x^2+1) && \text{Simplify.} \\ &= (x^2+1)^2[6x^2+(x^2+1)] && \text{Factor.} \\ &= (x^2+1)^2(7x^2+1) && \text{Simplify.} \end{aligned}$$

NOW WORK PROBLEM 7.

EXAMPLE 4 Using the Power Rule with Other Formulas

Find the derivative of the function: $f(x) = \left(\dfrac{3x+2}{4x^2-5}\right)^5$

SOLUTION Here, f is the quotient $\dfrac{3x+2}{4x^2-5}$ raised to the power 5. We begin by using the Power Rule and then use the formula for the derivative of a quotient:

$$\begin{aligned} f'(x) &= (5)\left(\frac{3x+2}{4x^2-5}\right)^4\left[\frac{d}{dx}\left(\frac{3x+2}{4x^2-5}\right)\right] && \text{Power Rule.} \\ &= (5)\left(\frac{3x+2}{4x^2-5}\right)^4\left[\frac{(4x^2-5)\frac{d}{dx}(3x+2) - (3x+2)\frac{d}{dx}(4x^2-5)}{(4x^2-5)^2}\right] && \text{Quotient Formula.} \\ &= (5)\left(\frac{3x+2}{4x^2-5}\right)^4\left[\frac{(4x^2-5)(3)-(3x+2)(8x)}{(4x^2-5)^2}\right] && \text{Differentiate.} \\ &= \frac{5(3x+2)^4(-12x^2-16x-15)}{(4x^2-5)^6} && \text{Simplify.} \end{aligned}$$

NOW WORK PROBLEM 19.

Application

The revenue $R = R(x)$ derived from selling x units of a product at a price p per unit is

$$R = xp$$

where $p = d(x)$ is the demand equation, namely, the equation that gives the price p when the number x of units demanded is known. The marginal revenue is then the derivative of R with respect to x:

$$R'(x) = \frac{d}{dx}(xp) = p + x\frac{dp}{dx} \qquad (2)$$

It is sometimes easier to find the marginal revenue by using formula (2) instead of differentiating the revenue function directly.

EXAMPLE 5 Finding the Marginal Revenue

Suppose the price p per ton when x tons of polished aluminum are demanded is given by the equation

$$p = \frac{2000}{x + 20} - 10 \quad 0 < x < 90$$

Find:

(a) The rate of change of price with respect to x.
(b) The revenue function.
(c) The marginal revenue.
(d) The marginal revenue at $x = 20$ and $x = 80$.

SOLUTION **(a)** The rate of change of price with respect to x is the derivative $\frac{dp}{dx}$.

$$\frac{dp}{dx} = \frac{d}{dx}\left(\frac{2000}{x + 20} - 10\right) = \frac{d}{dx}2000(x + 20)^{-1} - \frac{d}{dx}10$$

$$= -2000\,(x + 20)^{-2}\frac{d}{dx}(x + 20) - 0 = \frac{-2000}{(x + 20)^2}$$

Power Rule

(b) The revenue function is

$$R(x) = xp = x\left[\frac{2000}{x + 20} - 10\right] = \frac{2000x}{x + 20} - 10x$$

(c) Using formula (2), the marginal revenue is

$$R'(x) = p + x\frac{dp}{dx} \qquad \text{Formula (2)}$$

$$= \left[\frac{2000}{x + 20} - 10\right] + x\left(\frac{-2000}{(x + 20)^2}\right) \qquad \text{Use result from (a).}$$

$$= \frac{2000}{x + 20} - 10 - \frac{2000x}{(x + 20)^2} \qquad \text{Simplify.}$$

(d) Using the result from part (*c*), we find

$$R'(20) = \frac{2000}{40} - 10 - \frac{2000(20)}{(40)^2} = \$15/\text{ton}$$

$$R'(80) = \frac{2000}{100} - 10 - \frac{2000(80)}{(100)^2} = -\$6/\text{ton}$$

NOW WORK PROBLEM 31.

EXERCISE 4.4 Answers to Odd-Numbered Problems Begin on Page AN-XX.

In Problems 1–28, find using the derivative of each function Power Rule.

1. $f(x) = (2x - 3)^4$
2. $f(x) = (5x + 4)^3$
3. $f(x) = (x^2 + 4)^3$
4. $f(x) = (x^2 - 1)^4$
5. $f(x) = (3x^2 + 4)^2$
6. $f(x) = (9x^2 + 1)^2$
7. $f(x) = x(x + 1)^3$
8. $f(x) = x(x - 4)^2$
9. $f(x) = 4x^2(2x + 1)^4$
10. $f(x) = 3x^2(x^2 + 1)^3$
11. $f(x) = [x(x - 1)]^3$
12. $f(x) = [x(x + 4)]^4$
13. $f(x) = (3x - 1)^{-2}$
14. $f(x) = (2x + 3)^{-3}$
15. $f(x) = \dfrac{4}{x^2 + 4}$
16. $f(x) = \dfrac{3}{x^2 - 9}$
17. $f(x) = \dfrac{-4}{(x^2 - 9)^3}$
18. $f(x) = \dfrac{-2}{(x^2 + 2)^4}$
19. $f(x) = \left(\dfrac{x}{x + 1}\right)^3$
20. $f(x) = \left(\dfrac{x^2}{x + 5}\right)^4$
21. $f(x) = \dfrac{(2x + 1)^4}{3x^2}$
22. $f(x) = \dfrac{(3x + 4)^3}{9x}$
23. $f(x) = \dfrac{(x^2 + 1)^3}{x}$
24. $f(x) = \dfrac{(3x^2 + 4)^2}{2x}$
25. $f(x) = \left(x + \dfrac{1}{x}\right)^3$
26. $f(x) = \left(x - \dfrac{1}{x}\right)^4$
27. $f(x) = \dfrac{3x^2}{(x^2 + 1)^2}$
28. $g(x) = \dfrac{2x^3}{(x^2 - 4)^2}$

29. Car Depreciation A certain car depreciates according to the formula

$$V(t) = \frac{29000}{1 + 0.4t + 0.1t^2}$$

where V is the value of the car at time t in years. The derivative $V'(t)$ gives the rate at which the car depreciates. Find the rate at which the car is depreciating:

(a) 1 year after purchase. (b) 2 years after purchase.
(c) 3 years after purchase. (d) 4 years after purchase.

30. Demand Function The demand function for a certain calculator is given by

$$d(x) = \frac{100}{0.02x^2 + 1} \qquad 0 \le x \le 20$$

where x (measured in units of a thousand) is the quantity demanded per week and $d(x)$ is the unit price in dollars.

(a) Find $d'(x)$.
(b) Find $d'(10)$, $d'(15)$, and $d'(20)$ and interpret your results.
(c) Find the revenue function.
(d) Find the marginal revenue.

31. Demand Equation The price p per pound when x pounds of a certain commodity are demanded is

$$p = \frac{10{,}000}{5x + 100} - 5 \qquad 0 < x < 90$$

Find:

(a) The rate of change of price with respect to x.
(b) The revenue function.
(c) The marginal revenue.
(d) The marginal revenue at $x = 10$ and at $x = 40$.
(e) Interpret the answer to (d).

32. Revenue Function The weekly revenue R in dollars resulting from the sale of x typewriters is

$$R(x) = \frac{100x^5}{(x^2 + 1)^2} \qquad 0 \le x \le 100$$

Find:

(a) The marginal revenue.
(b) The marginal revenue at $x = 40$.
(c) The marginal revenue at $x = 60$.
(d) Interpret the answers to (b) and (c).

33. Amino Acids A protein disintegrates into amino acids according to the formula

$$M = \frac{28}{t+2}$$

where M, the mass of the protein, is measured in grams and t is time measured in hours.

(a) Find the average rate of change in mass from $t = 0$ to $t = 2$ hours.
(b) Find $M'(0)$.
(c) Interpret the answers to (a) and (b).

4.5 The Derivatives of the Exponential and Logarithmic Functions; the Chain Rule

PREPARING FOR THIS SECTION *Before getting started, review the following:*

> The Exponential Function (Section 2.3, pp. xx–xx)
> The Logarithmic Function (Section 2.4, pp. xx–xx)
> Change of Base Formula (Section 2.5, pp. xx–xx)

OBJECTIVES
1 Find the derivative of functions involving e^x
2 Find a derivative using the Chain Rule
3 Find the derivative of functions involving $\ln x$
4 Find the derivative of functions including $\log a^x$ and a^x

Up to now, our discussion of finding derivatives has been focused on polynomial functions (derivative of a sum or difference), rational functions (derivative of a quotient), and these functions raised to an integer power (the Power Rule). In this section we present the formulas for finding the derivative of the exponential function and the logarithm function.

The Derivative of $f(x) = e^x$

We begin the discussion of the derivative of $f(x) = e^x$ by considering the function

$$f(x) = a^x \qquad a > 0, \qquad a \neq 1$$

To find the derivative of $f(x) = a^x$, we use the formula for finding the derivative of f at x using the difference quotient, namely:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

For $f(x) = a^x$, we have

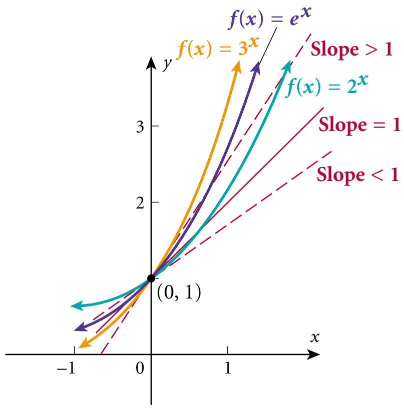
$$f'(x) = \frac{d}{dx}a^x = \lim_{h\to 0}\frac{a^{x+h} - a^x}{h} \underset{\substack{\uparrow \\ \text{Factor out } a^x}}{=} \lim_{h\to 0}\left[a^x\left(\frac{a^h - 1}{h}\right)\right] = a^x \lim_{h\to 0}\frac{a^h - 1}{h}$$

Suppose we seek $f'(0)$. Assuming the limit on the right exists and equals some number, it follows (since $a^0 = 1$) that the derivative of $f(x) = a^x$ at 0 is

$$f'(0) = \lim_{h\to 0}\frac{a^h - 1}{h}$$

This limit equals the slope of the tangent line to the graph of $f(x) = a^x$ at the point $(0, 1)$. The value of this limit depends upon the choice of a. Observe in Figure 10 that the slope of the tangent line to the graph of $f(x) = 2^x$ at $(0, 1)$ is less than 1, and that the slope of the tangent line to the graph of $f(x) = 3^x$ at $(0, 1)$ is greater than 1.

FIGURE 10



From this, we conclude there is a number a, $2 < a < 3$, for which the slope of the tangent line to the graph of $f(x) = a^x$ at $(0, 1)$ is exactly 1. The function $f(x) = a^x$ for which $f'(0) = 1$ is the function $f(x) = e^x$, whose base is the number e, we introduced in Chapter 2. A further property of the number e is that

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1$$

Using this result, we find that

$$\frac{d}{dx} e^x = \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} = \lim_{h \to 0} \frac{e^x(e^h - 1)}{h} = e^x \lim_{h \to 0} \frac{e^h - 1}{h} = e^x(1) = e^x$$

Derivative of $f(x) = e^x$

The derivative of the exponential function $f(x) = e^x$ is e^x. That is,

$$\boxed{\frac{d}{dx} e^x = e^x} \qquad (1)$$

The simple nature of formula (1) is one of the reasons the exponential function $f(x) = e^x$ appears so frequently in applications.

1 **EXAMPLE 1 Finding the Derivative of Functions Involving e^x**

Find the derivative of each function:

(a) $f(x) = x^2 + e^x$ **(b)** $f(x) = xe^x$ **(c)** $f(x) = \dfrac{e^x}{x}$

SOLUTION **(a)** Use the sum formula. Then

$$f'(x) = \frac{d}{dx}(x^2 + e^x) = \frac{d}{dx} x^2 + \frac{d}{dx} e^x = 2x + e^x$$

(b) Use the formula for the derivative of a product: Then

$$f'(x) = \frac{d}{dx}(xe^x) = x\frac{d}{dx}e^x + e^x\frac{d}{dx}x = xe^x + e^x(1) = e^x(x+1)$$

(c) Use the formula for the derivative of a quotient. Then

$$f'(x) = \frac{d}{dx}\frac{e^x}{x} \underset{\substack{\uparrow \\ \text{Quotient formula}}}{=} \frac{x\dfrac{d}{dx}e^x - e^x\dfrac{d}{dx}x}{x^2} \underset{\substack{\uparrow \\ \text{Differentiate}}}{=} \frac{xe^x - e^x \cdot 1}{x^2} \underset{\substack{\uparrow \\ \text{Factor}}}{=} \frac{(x-1)e^x}{x^2}$$

NOW WORK PROBLEM 3.

To find the derivative of other functions involving e^x and to find the derivative of the logarithmic function requires a formula called *the Chain Rule.*

The Chain Rule

The Power Rule is a special case of a more general, and more powerful formula, called the *Chain Rule.* This formula enables us to find the derivative of a *composite function.*

Consider the function $y = (2x+3)^2$. If we write $y = f(u) = u^2$ and $u = g(x) = 2x + 3$, then, by a substitution process, we can obtain the original function, namely, $y = f(u) = f(g(x)) = (2x+3)^2$. This process is called **composition** and the function $y = (2x+3)^2$ is called the **composite function** of $y = f(u) = u^2$ and $u = g(x) = 2x + 3$.

EXAMPLE 2 Finding a Composite Function

Find the composite function of

$$y = f(u) = \sqrt{u} \qquad \text{and} \qquad u = g(x) = x^2 + 4$$

SOLUTION The composite function is

$$y = f(u) = \sqrt{u} = \sqrt{g(x)} = \sqrt{x^2+4}$$

The Chain Rule will require that we find the components of a composite function.

EXAMPLE 3 Decomposing a Composite Function

(a) If $y = (5x+1)^3$, then $y = u^3$ and $u = 5x + 1$.
(b) If $y = (x^2+1)^{-2}$, then $y = u^{-2}$ and $u = x^2 + 1$.
(c) If $y = \dfrac{5}{(2x+3)^3}$, then $y = \dfrac{5}{u^3}$ and $u = 2x + 3$.

In the above examples, the composite function was "broken up" into simpler functions. The Chain Rule provides a way to use these simpler functions to find the derivative of the composite function.

The Chain Rule

Suppose f and g are differentiable functions. If $y = f(u)$ and $u = g(x)$, then, after substitution, y is a function of x. The Chain Rule states that the derivative of y with respect to x is the derivative of y with respect to u times the derivative of u with respect to x. That is,

$$\boxed{\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}} \qquad (2)$$

2 **EXAMPLE 4** **Finding a Derivative Using the Chain Rule**

Use the Chain Rule to find the derivative of: $y = (5x + 1)^3$

SOLUTION We break up y into simpler functions: If $y = (5x + 1)^3$, then $y = u^3$ and $u = 5x + 1$. To find $\frac{dy}{dx}$, we first find $\frac{dy}{du}$ and $\frac{du}{dx}$:

$$\frac{dy}{du} = \frac{d}{du}(u^3) = 3u^3 \quad \text{and} \quad \frac{du}{dx} = \frac{d}{dx}(5x + 1) = 5$$

By the Chain Rule,

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = 3u^2 \cdot 5 = 15u^2 \underset{\substack{\uparrow \\ u = 5x + 1}}{=} 15(5x + 1)^2$$

Notice that when using the Chain Rule, we must substitute for u in the expression for $\frac{dy}{du}$ so that we obtain a function of x.

NOW WORK PROBLEM 9.

EXAMPLE 5 **Finding a Derivative Using The Chain Rule**

Find the derivative of: $y = e^{x^2}$

SOLUTION We break up y into simpler functions. If $y = e^{x^2}$, then $y = e^u$ and $u = x^2$. Now use the Chain Rule to find $y' = \frac{dy}{dx}$.

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = e^u \cdot 2x \underset{\substack{\uparrow \\ u = x^2}}{=} 2xe^{x^2}$$

The result of Example 5 can be generalized.

Derivative of $y = e^{g(x)}$

The derivative of a composite function $y = e^{g(x)}$, where g is a differentiable function, is

$$\boxed{\frac{d}{dx} e^{g(x)} = e^{g(x)} \frac{d}{dx} g(x)} \qquad (3)$$

The proof is left as an exercise. See Problem 67.

EXAMPLE 6 Finding the Derivative of Functions of the Form $e^{g(x)}$

Find the derivative of each function:

(a) $f(x) = 4e^{5x}$ **(b)** $f(x) = e^{x^2+1}$

SOLUTION **(a)** Use Formula (3) with $g(x) = 5x$. Then

$$f'(x) = \frac{d}{dx}(4e^{5x}) = 4\frac{d}{dx}e^{5x} \underset{\substack{\uparrow \\ \text{Formula (3)}}}{=} 4 \cdot e^{5x}\frac{d}{dx}(5x) = 4e^{5x}(5) = 20e^{5x}$$

(b) Use Formula (3) with $g(x) = x^2 + 1$. Then

$$f'(x) = \frac{d}{dx}e^{x^2+1} \underset{\substack{\uparrow \\ \text{Formula (3)}}}{=} e^{x^2+1} = \frac{d}{dx}(x^2 + 1) = e^{x^2+1}(2x) = 2xe^{x^2+1}$$

NOW WORK PROBLEM 23.

EXAMPLE 7 Finding the Derivative of Functions Involving e^x

Find the derivative of each function:

(a) $f(x) = xe^{x^2}$ **(b)** $f(x) = \dfrac{x}{e^x}$ **(c)** $f(x) = (e^x)^2$

SOLUTION **(a)** The function f is the product of two simpler functions, so we start with the product formula.

$$f'(x) = \frac{d}{dx}(xe^{x^2}) \underset{\substack{\uparrow \\ \text{Product Formula}}}{=} x\frac{d}{dx}e^{x^2} + e^{x^2}\frac{d}{dx}x \underset{\substack{\uparrow \\ \text{Formula (3); } \frac{d}{dx}x = 1}}{=} x \cdot e^{x^2} \cdot \frac{d}{dx}x^2 + e^{x^2} \cdot 1 = xe^{x^2} \cdot 2x + e^{x^2} \underset{\substack{\uparrow \\ \text{Factor}}}{=} e^{x^2}(2x^2 + 1)$$

(b) We could use the quotient formula, but it is easier to rewrite f in the form $f(x) = xe^{-x}$ and use the product formula

$$f'(x) = \frac{d}{dx}xe^{-x} \underset{\substack{\uparrow \\ \text{Product Formula}}}{=} x\frac{d}{dx}e^{-x} + e^{-x}\frac{d}{dx}x \underset{\substack{\uparrow \\ \text{Formula (3)}}}{=} x \cdot e^{-x}\frac{d}{dx}(-x) + e^{-x} \cdot 1 = xe^{-x}(-1) + e^{-x} \underset{\substack{\uparrow \\ \text{Factor}}}{=} e^{-x}(1 - x)$$

(c) Here the function is e^x raised to the power 2. We first apply a Law of Exponents and write $f(x) = (e^x)^2 = e^{2x}$.

Then we can use Formula (3).

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

CAUTION: Notice the difference between e^{x^2} and $(e^x)^2$. In the first, one e is raised to the power x^2; in the second, the parentheses tell us e^x is raised to the power 2.

NOW WORK PROBLEM 29.

The Derivative of $f(x) = \ln x$

To find the derivative of $f(x) = \ln x$, we observe that if $y = \ln x$, then $e^y = x$. That is,

$$e^{\ln x} = x$$

If we differentiate both sides with respect to x, we obtain

$$\begin{aligned}
\frac{d}{dx} e^{\ln x} &= \frac{d}{dx} x \\
e^{\ln x} \frac{d}{dx} \ln x &= 1 && \text{Apply Formula (3) on the left.} \\
\frac{d}{dx} \ln x &= \frac{1}{e^{\ln x}} && \text{Solve for } \tfrac{d}{dx} \ln x. \\
\frac{d}{dx} \ln x &= \frac{1}{x} && e^{\ln x} = x.
\end{aligned}$$

We have proved the following formula:

Derivative of $f(x) = \ln x$

If $f(x) = \ln x$, then $f'(x) = \dfrac{1}{x}$. That is

$$\boxed{\frac{d}{dx} \ln x = \frac{1}{x}} \tag{4}$$

3 **EXAMPLE 8 Finding the Derivative of Functions Involving ln *x***

Find the derivative of each function.

(a) $f(x) = x^2 + \ln x$ **(b)** $f(x) = x \ln x$

SOLUTION **(a)** Use the sum formula. Then $f'(x) = \dfrac{d}{dx}(x^2 + \ln x) = \dfrac{d}{dx} x^2 + \dfrac{d}{dx} \ln x = 2x + \dfrac{1}{x}$

(b) Use the product formula. Then

$$\begin{aligned}
f'(x) &= \frac{d}{dx}(x \ln x) = x \frac{d}{dx} \ln x + \ln x \frac{d}{dx} x \\
&= (x)\left(\frac{1}{x}\right) + (\ln x)(1) = 1 + \ln x
\end{aligned}$$

NOW WORK PROBLEM 35.

To differentiate the natural logarithm of a function $g(x)$, namely, $\ln g(x)$, use the following formula.

Derivative of $\ln g(x)$

The formula for finding the derivative of the composite function $f(x) = \ln g(x)$, where g is a differentiable function, is

$$\boxed{\frac{d}{dx} \ln g(x) = \frac{\dfrac{d}{dx} g(x)}{g(x)}} \tag{5}$$

The proof uses the Chain Rule and is left as an exercise. See Problem 68.

EXAMPLE 9 Finding the Derivative of Functions Involving ln *x*

Finding the derivative of each function.

(a) $f(x) = \ln(x^2 + 1)$ **(b)** $f(x) = (\ln x)^2$

SOLUTION **(a)** The function $f(x) = \ln(x^2 + 1)$ is of the form $f(x) = \ln g(x)$. We use Formula (4) with $g(x) = x^2 + 1$. Then,

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

(b) The function $f(x)$ is $\ln x$ raised to the power 2. We use the Power Rule. Then

$$f'(x) = 2\ln x\left(\frac{d}{dx}\ln x\right) = \frac{2\ln x}{x}$$

NOW WORK PROBLEM 45.

The Derivative of $f(x) = \log_a x$ and $f(x) = a^x$

4 Find the derivative of functions including $\log a^x$ and a^x

To find the derivative of the logarithm function $f(x) = \log_a x$ for any base a, we use the Change-of-Base Formula: Then

$$f(x) = \log_a x = \frac{\log_e x}{\log_e a} = \frac{\ln x}{\ln a}$$

Since $\ln a$ is a constant, we have

$$f'(x) = \frac{d}{dx}\log_a x = \frac{d}{dx}\frac{\ln x}{\ln a} = \frac{1}{\ln a}\frac{d}{dx}\ln x = \frac{1}{\ln a}\frac{1}{x} = \frac{1}{x\ln a}$$

We have the formula

Derivative of $f(x) = \log_a x$

If $f(x) = \log_a x$, then $f'(x) = \dfrac{1}{x\ln a}$. That is,

$$\boxed{\frac{d}{dx}\log_a x = \frac{1}{x\ln a}} \tag{6}$$

EXAMPLE 10 Finding the Derivative of $\log_2 x$

Find the derivative of: $f(x) = \log_2 x$

SOLUTION Using Formula (6), we have

$$f'(x) = \frac{d}{dx}\log_2 x = \frac{1}{x\ln 2}$$

NOW WORK PROBLEM 47.

To find the derivative of $f(x) = a^x$, where $a > 0$, $a \neq 1$, is any real constant, we use the definition of a logarithm and the change-of-base formula. If $y = a^x$, we have

$$x = \log_a y$$
$$x = \frac{\ln y}{\ln a}$$
$$x = \frac{\ln a^x}{\ln a}$$

Now, we differentiate both sides with respect to x:

$$\frac{d}{dx}x = \frac{d}{dx}\frac{\ln a^x}{\ln a}$$

$$1 = \frac{1}{\ln a}\frac{d}{dx}\ln a^x \qquad \text{ln } a \text{ is a constant.}$$

$$1 = \frac{1}{\ln a}\frac{\frac{d}{dx}a^x}{a^x} \qquad \text{Use Formula (5).}$$

$$1 = \frac{\frac{d}{dx}a^x}{a^x \ln a} \qquad \text{Simplify.}$$

$$\frac{d}{dx}a^x = a^x \ln a \qquad \text{Solve for } \frac{d}{dx}a^x.$$

We have derived the formula:

Derivative of $f(x) = a^x$

The derivative of $f(x) = a^x$, $a > 0$, $a \neq 1$, is $f'(x) = a^x \ln a$. That is,

$$\boxed{\frac{d}{dx}a^x = a^x \ln a} \tag{7}$$

EXAMPLE 11 Finding the Derivative of 2^x

Find the derivative of: $f(x) = 2^x$

SOLUTION Using Formula (7), we have

$$f'(x) = \frac{d}{dx}2^x = 2^x \ln 2$$

NOW WORK PROBLEM 51.

EXAMPLE 12 Maximizing Profit

At a Notre Dame football weekend, the demand for game-day t-shirts is given by

$$p = 30 - 5\ln\left(\frac{x}{100} + 1\right)$$

where p is the price of the shirt in dollars and x is the number of shirts demanded.

(a) At what price can 1000 t-shirts be sold?
(b) At what price can 5000 t-shirts be sold?
(c) Find the marginal demand for 1000 t-shirts and interpret the answer.
(d) Find the marginal demand for 5000 t-shirts and interpret the answer.
(e) Find the revenue function $R = R(x)$.
(f) Find the marginal revenue from selling 1000 t-shirts and interpret the answer.
(g) Find the marginal revenue from selling 5000 t-shirts and interpret the answer.
(h) If each t-shirt costs \$4, find the profit function $P = P(x)$.
(i) What is the profit if 1000 t-shirts are sold?
(j) What is the profit if 5000 t-shirts are sold?
(k) Use the TABLE feature of a graphing utility to find the quantity x (to the nearest hundred) that maximizes profit.
(l) What price should be charged for a t-shirt to maximize profit?

SOLUTION **(a)** For $x = 1000$, the price p is

$$p = 30 - 5\ln\left(\frac{1000}{100} + 1\right) = \$18.01$$

(b) For $x = 5000$, the price p is

$$p = 30 - 5\ln\left(\frac{5000}{100} + 1\right) = \$10.34$$

(c) The marginal demand for x shirts is

$$p'(x) = \frac{dp}{dx} = \frac{d}{dx}\left[30 - 5\ln\left(\frac{x}{100} + 1\right)\right] = -5\,\frac{\frac{1}{100}}{\frac{x}{100} + 1} = \frac{-5}{x + 100}$$

For $x = 1000$,

$$p'(1000) = \frac{-5}{1000 + 100} = -\$0.0045$$

This means that another t-shirt will be demanded if the price is reduced by \$0.0045.

(d) For $x = 5000$,

$$p'(5000) = \frac{-5}{5000 + 100} = -\$0.00098$$

This means that another t-shirt will be demanded if the price is reduced by \$0.00098.

(e) The revenue function $R = R(x)$ is

$$R = xp = x\left[30 - 5\ln\left(\frac{x}{100} + 1\right)\right]$$

(f) The marginal revenue is

$$\begin{aligned} R'(x) &= \frac{d}{dx}\,[xp(x)] = xp'(x) + p(x) \\ &= x \cdot \frac{-5}{x + 100} + 30 - 5\ln\left(\frac{x}{100} + 1\right) \\ &= \frac{-5x}{x + 100} + 30 - 5\ln\left(\frac{x}{100} + 1\right) \end{aligned}$$

If $x = 1000$,

$$R'(1000) = \frac{-5000}{5100} + 30 - 5\ln(11) = \$17.03$$

The revenue received for selling the 1001st t-shirt is \$17.03

(g) If $x = 5000$

$$R'(5000) = \frac{-25000}{5100} + 30 - 5\ln(51) = \$5.44$$

The revenue received for selling the 5001st t-shirt is \$5.44.

(h) The cost C for x t-shirts is $C = -4x$, so the product function P is

$$P = P(x) = R(x) - C(x) = x\left[30 - 5\ln\left(\frac{x}{100} + 1\right)\right] - 4x$$
$$= 26x - 5x\ln\left(\frac{x}{100} + 1\right)$$

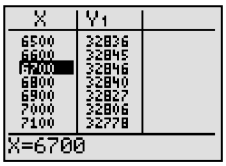
(i) If $x = 1000$, the profit is

$$P(1000) = 26(1000) - 5(1000)\ln\left(\frac{1000}{100} + 1\right) = \$14{,}010.52$$

(j) If $x = 5000$, the profit is

$$P(5000) = 26(5000) - 5(5000)\ln\left(\frac{5000}{100} + 1\right) = \$31{,}704.36$$

FIGURE 11



(k) See Figure 11. For $x = 6700$ t-shirts, the profit is largest. (\$32,846)

(l) If $x = 6700$, the price p is

$$p(6700) = 30 - 5\ln\left(\frac{6700}{100} + 1\right) = \$8.90$$

SUMMARY

$$\frac{d}{dx}e^x = e^x \qquad \frac{d}{dx}e^{g(x)} = e^{g(x)}g'(x) \qquad \frac{d}{dx}a^x = a^x \ln a$$

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

EXERCISE 4.5 Answers to Odd-Numbered Problems Begin on Page AN-XX.

In Problems 1–8, find the derivative of each function.

1. $f(x) = x^3 - e^x$ **2.** $f(x) = 2e^x - x$ **3.** $f(x) = x^2e^x$ **4.** $f(x) = x^3e^x$

5. $f(x) = \dfrac{e^x}{x^2}$ **6.** $f(x) = \dfrac{5x}{e^x}$ **7.** $f(x) = \dfrac{4x^2}{e^x}$ **8.** $f(x) = \dfrac{3x^3}{e^x}$

In Problems 9–20, find $\dfrac{dy}{dx}$ using the Chain Rule.

9. $y = u^5, \quad u = x^3 + 1$ **10.** $y = u^3, \quad u = 2x + 5$ **11.** $y = \dfrac{u}{u+1}, \quad u = x^2 + 1$

12. $y = \frac{u-1}{u}, \quad u = x^2 - 1$

13. $y = (u+1)^2, \quad u = \frac{1}{x}$

14. $y = (u^2 - 1)^3, \quad u = \frac{1}{x+2}$

15. $y = (u^3 - 1)^5, \quad u = x^{-2}$

16. $y = (u^2 + 4)^4, \quad u = x^{-2}$

17. $y = u^3, \quad u = e^x$

18. $y = 4u^2, \quad u = e^x$

19. $y = e^u, \quad u = x^3$

20. $y = e^u, \quad u = \frac{1}{x}$

21. Find the derivative y' of $y = (x^3 + 1)^2$ by:

(a) Using the Chain Rule.
(b) Using the Power Rule.
(c) Expanding and then differentiating.

22. Follow the directions in Problem 21 for the function $y = (x^2 - 2)^3$.

In Problems 23–54, find the derivative of each function.

23. $f(x) = e^{5x}$

24. $f(x) = e^{-3x}$

25. $f(x) = 8e^{-x^2}$

26. $f(x) = -e^{3x^2}$

27. $f(x) = x^2 e^{x^2}$

28. $f(x) = x^3 e^{x^2}$

29. $f(x) = 5(e^x)^3$

30. $f(x) = 4(e^x)^4$

31. $f(x) = \frac{x^2}{e^x}$

32. $f(x) = \frac{8x}{e^{-x}}$

33. $f(x) = \frac{(e^x)^2}{x}$

34. $f(x) = \frac{e^{-2x}}{x^2}$

35. $f(x) = x^2 - 3 \ln x$

36. $f(x) = 5 \ln x - 2x$

37. $f(x) = x^2 \ln x$

38. $f(x) = x^3 \ln x$

39. $f(x) = 3 \ln (5x)$

40. $f(x) = -2 \ln (3x)$

41. $f(x) = x \ln (x^2 + 1)$

42. $f(x) = x^2 \ln (x^2 + 1)$

43. $f(x) = x + 8 \ln (3x)$

44. $f(x) = 3 \ln (2x) - 5x$

45. $f(x) = 8(\ln x)^3$

46. $f(x) = 2(\ln x)^4$

47. $f(x) = \log_3 x$

48. $f(x) = x + \log_4 x$

49. $f(x) = x^2 \log_2 x$

50. $f(x) = x^3 \log_3 x$

51. $f(x) = 3^x$

52. $f(x) = x + 4^x$

53. $f(x) = x^2 \cdot 2^x$

54. $f(x) = x^3 \cdot 3^x$

In Problems 55–62, find an equation of the tangent line to the graph of each function at the given point.

55. $f(x) = e^{70}$ at $(0, 1)$

56. $f(x) = e^{4x}$ at $(0, 1)$

57. $f(x) = \ln x$ at $(1, 0)$

58. $f(x) = \ln (3x)$ at $(1, 0)$

59. $f(x) = e^{3x-2}$ at $\left(\frac{2}{3}, 1\right)$

60. $f(x) = e^{-x^2}$ at $\left(1, \frac{1}{e}\right)$

61. $f(x) = x \ln x$ at $(1, 0)$

62. $f(x) = \ln x^2$ at $(1, 0)$

---

63. Find the equation of the tangent line to $y = e^x$ that is parallel to the line $y = x$.

64. Find the equation of the tangent line to $y = e^{3x}$ that is perpendicular to the line $y = -\frac{1}{2}x$.

65. **Weber–Fechner Law** When a certain drug is administered, the reaction R to the dose x is given by the **Weber–Fechner law:**

$$R = 5.5 \ln x + 10$$

(a) Find the reaction rate for a dose of 5 units.
(b) Find the reaction rate for a dose of 10 units.
(c) Interpret the results of parts (a) and (b).

66. **Marginal Cost** The cost (in dollars) of producing x units (measured in thousands) of a certain product is found to be

$$C(x) = 20 + \ln(x + 1)$$

Find the marginal cost.

67. **Atmospheric Pressure** The atmospheric pressure at a height of x meters above sea level is $P(x) = 10^4 e^{-0.00012x}$ kilograms per square meter. What is the rate of change of the pressure with respect to the height at $x = 500$ meters? At $x = 700$ meters?

68. **Revenue** Revenue sales analysis of a new toy by Toys Inc. indicates that the relationship between the unit price p and the monthly sales x of its new toy is given by the equation

$$p = 10e^{-0.04x}$$

Find

(a) The revenue function $R = R(x)$.
(b) The marginal revenue R when $x = 200$.

Calculus

The art of measuring shape and variation!

Practice worksheets for Mastery of Differentiation

Graeme Henderson

Dear Reader,

It is no secret that, to master any skill, we need to practise it!

School textbooks usually contain sufficient material for you to learn HOW to use certain skills but, due to space restrictions, usually don't contain sufficient examples for you to MASTER those skills.

This publication is intended to fill that gap … for finding derivatives, at least!

If you are a student, let me suggest that you set time aside regularly to work through a few examples from this booklet. It need not be a great deal of time, but I recommend that, on a weekly, fortnightly, or monthly basis, you spend a few minutes practising the art of finding derivatives. You may find it a useful exercise to do this with friends and to discuss the more difficult examples. To build speed, try calculating the derivatives on the first sheet mentally … and have a friend or parent check your answers.

If you are a teacher, please note that the sheets have been designed so that they may be laminated back-to-back (questions on one side and answers on the other) and used in a classroom setting.

I have invested a great deal of time in putting this material together. Although I have tried to be very careful, it is quite likely that some mistakes appear in the answers. If you find any, please let me know by visiting my Crystal Clear Mathematics website and leaving a message on the "Contact Us" page. If you give permission for me to mention you by name, I will acknowledge your contribution in future (corrected) editions.

Because this took so long to prepare (dozens of hours), and because I am making it available for free, it would be appreciated if you could consider making a donation towards the upkeep of my website … especially if you represent a school. It costs me many hours and thousands of dollars to maintain the site and add videos and other resources. Any contribution you may be able to make in return for the free use of this material would be greatly appreciated.

I hope this booklet is of help to you!

Graeme Henderson
B.Sc. (Hons), B.Th. (Hons),
Dip.Ed., Dip.Min.

16 April 2015

Speaking of copyright, thank you, Schristia (on Flickr), for allowing me to use your image for the cover of this booklet.

Derivatives in Your Head!

Differentiate with respect to x:

	A	B	C	D	E	F	G
1.	x^3	$4x^3$	$\frac{8x^2}{3}$	$x^{\frac{1}{2}}$	$3x^{\frac{1}{3}}$	x^{q+2}	$\sqrt{x}$
2.	x^{10}	$8x$	$\frac{-7x^2}{2}$	$x^{\frac{3}{4}}$	$4x^{\frac{3}{4}}$	x^{3m}	$\sqrt[5]{x}$
3.	x^4	$5x^4$	$\frac{5x^3}{4}$	$x^{\frac{2}{3}}$	$-6x^{\frac{2}{3}}$	x^{p+q}	$-\sqrt[7]{x^2}$
4.	x^5	$-3x^2$	$\frac{2x^5}{3}$	$x^{\frac{5}{4}}$	$-10x^{-\frac{3}{5}}$	$-x^{171}$	$\sqrt[4]{x^3}$
5.	x^6	$2x$	$-\frac{6x^7}{7}$	$x^{\frac{3}{2}}$	$9x^{\frac{2}{3}}$	x^{n+1}	$\sqrt[3]{x^8}$
6.	x^2	8	$\frac{2}{9}x^3$	$x^{1.8}$	$14x^{-\frac{3}{7}}$	x^{m+n+1}	$-\sqrt{x^{1.8}}$
7.	x^7	$-7x^6$	$\frac{11}{5}x^5$	$x^{6.1}$	$-2x^{-\frac{4}{5}}$	$-x^{5d-k+6}$	$\sqrt[3]{x^{2.4}}$
8.	x^9	$9x^5$	$-\frac{3}{4}x^6$	$x^{3.4}$	$4x^{-\frac{5}{3}}$	$x^{\sqrt{2}+1}$	$\sqrt[3]{x^{-5.1}}$
9.	x	$-12x^7$	$\frac{3}{7}x^4$	$x^{-\frac{1}{2}}$	$-7x^{\frac{2}{3}}$	$x^{3\pi+4}$	$\sqrt{\frac{1}{x^{7.6}}}$
10.	x^0	-14	$\frac{5}{8}x^3$	$x^{-\frac{7}{2}}$	$\frac{4x^{\frac{5}{4}}}{5}$	x^{3i-2}	$-\frac{1}{\sqrt[4]{x^3}}$
11.	x^{-2}	$-8x^0$	$\frac{-7x^{-3}}{3}$	$x^{-\frac{4}{3}}$	$-\frac{2}{3}x^{\frac{3}{2}}$	π^3	$\frac{7}{\sqrt[5]{x^2}}$
12.	x^{-1}	$11x^{-7}$	$\frac{2x^{-6}}{9}$	$x^{-\frac{5}{4}}$	$\frac{4}{21}x^{\frac{7}{2}}$	$-5x^{2n+1}$	$-\sqrt[3]{\frac{1}{x^2}}$
13.	x^{-4}	$9x^{-8}$	$-\frac{4x^{-5}}{15}$	$x^{-5.3}$	$-\frac{9}{4}x^{-\frac{4}{3}}$	$3x^{5k+4}$	$\frac{1}{\sqrt{x^3}}$
14.	x^{-3}	$-14x^{-10}$	$\frac{5x^{-14}}{7}$	$x^{-2.6}$	$\frac{-2x^{\frac{5}{4}}}{5}$	$-4x^{\sqrt{3}+7}$	$x\sqrt{x}$
15.	x^{-10}	$4x^{-3}$	$-\frac{2}{9}x^{-3}$	$x^{-8.3}$	$\frac{4}{3}x^{\frac{7}{3}}$	$\sqrt{3}x^{\pi-2}$	$-\frac{5}{3\sqrt{x}}$
16.	x^{-7}	$5x^{-1}$	$\frac{1}{5}x^{-3}$	$x^{\frac{11}{5}}$	$\frac{2}{7}x^{-\frac{2}{5}}$	$-2x^{5i+1}$	$\sqrt[4]{\frac{1}{x^{-1.2}}}$
17.	x^{-16}	$-2x^{-6}$	$\frac{-8}{5}x^{-4}$	$x^{\frac{5}{7}}$	$-\frac{9x^{-\frac{8}{5}}}{4}$	$2.5x^{2e+3}$	$-\frac{3}{7\sqrt{x^{1.4}}}$
18.	x^{-9}	$3x^{-5}$	$\frac{4}{7}x^{-6}$	$x^{-\frac{9}{2}}$	$-\frac{3}{14}x^{\frac{7}{3}}$	$0.4x^{6f}$	$\frac{14}{3\sqrt[7]{x^3}}$
19.	x^{-6}	$-7x^{-4}$	$\frac{5}{8}x^3$	$-7x^{-4}$	$\frac{10}{9}x^{-\frac{3}{5}}$	$2x^{a+b}$	$x^2\sqrt[3]{x}$
20.	x^{-11}	$6x^{-2}$	$-\frac{5}{14}x^{-7}$	$6x^{-2}$	$\frac{7x^{-\frac{5}{3}}}{2}$	$3.5x^{4w+2}$	$-\frac{2x^3}{\sqrt[5]{x}}$

Derivatives in Your Head (Answers)!

Solutions:

	A	B	C	D	E	F	G
1.	$3x^2$	$12x^2$	$\frac{16x}{3}$	$\frac{x^{-\frac{1}{2}}}{2}$	$x^{-\frac{2}{3}}$	$(q+2)x^{q+1}$	$\frac{x^{-\frac{1}{2}}}{2}=\frac{1}{2\sqrt{x}}$
2.	$10x^9$	8	$-7x$	$\frac{3x^{-\frac{1}{4}}}{4}$	$3x^{-\frac{1}{4}}$	$3mx^{3m-1}$	$\frac{x^{-\frac{4}{5}}}{5}=\frac{1}{5\sqrt[5]{x^4}}$
3.	$4x^3$	$20x^3$	$\frac{15x^2}{4}$	$\frac{2x^{-\frac{1}{3}}}{3}$	$-4x^{-\frac{1}{3}}$	$(p+q)x^{p+q-1}$	$-\frac{2x^{-\frac{5}{7}}}{7}=-\frac{2}{7\sqrt[7]{x^5}}$
4.	$5x^4$	$-6x$	$\frac{10x^4}{3}$	$\frac{5x^{\frac{1}{4}}}{4}$	$6x^{-\frac{8}{5}}$	$-171x^{170}$	$\frac{3x^{-\frac{1}{4}}}{4}=\frac{3}{4\sqrt[4]{x}}$
5.	$6x^5$	2	$-6x^6$	$\frac{3x^{\frac{1}{2}}}{2}$	$6x^{-\frac{1}{3}}$	$(n+1)x^n$	$\frac{8x^{\frac{5}{3}}}{3}=\frac{8\sqrt[3]{x^5}}{3}$
6.	$2x$	0	$\frac{2}{3}x^2$	$1.8x^{0.8}$	$-6x^{-\frac{10}{7}}$	$(m+n+1)x^{m+n}$	$-0.9x^{-0.1}$
7.	$7x^6$	$-42x^5$	$11x^4$	$6.1x^{5.1}$	$\frac{8x^{-\frac{9}{5}}}{5}$	$-(5d-k+6)x^{5d-k+5}$	$0.8x^{-0.2}$
8.	$9x^8$	$45x^4$	$-\frac{9}{2}x^5$	$3.4x^{2.4}$	$-\frac{20x^{-\frac{8}{3}}}{3}$	$(\sqrt{2}+1)x^{\sqrt{2}+1}$	$-1.7x^{-2.7}$
9.	1	$-84x^6$	$\frac{12}{7}x^3$	$-\frac{x^{-\frac{3}{2}}}{2}$	$-\frac{14x^{-\frac{1}{3}}}{3}$	$(3\pi+4)x^{3\pi+3}$	$-3.8x^{-4.8}$
10.	0	0	$\frac{15}{8}x^2$	$-\frac{7x^{-\frac{9}{2}}}{2}$	$\frac{4x^{\frac{5}{4}}}{5}$	$(3i-2)x^{3i-3}$	$\frac{3x^{-\frac{7}{4}}}{4}=\frac{3}{4\sqrt[4]{x^7}}$
11.	$-2x^{-3}$	0	$7x^{-4}$	$-\frac{4x^{-\frac{7}{3}}}{3}$	$-x^{-\frac{5}{2}}$	0	$\frac{14x^{-\frac{7}{5}}}{5}=\frac{14}{5\sqrt[5]{x^7}}$
12.	$-x^{-2}$	$-77x^{-8}$	$-\frac{4x^{-7}}{3}$	$-\frac{5x^{-\frac{9}{4}}}{4}$	$\frac{2}{3}x^{\frac{5}{2}}$	$-5(2n+1)x^{2n}$	$\frac{2x^{-\frac{5}{3}}}{3}=\frac{2}{3\sqrt[3]{x^5}}$
13.	$-4x^{-5}$	$-72x^{-9}$	$\frac{4x^{-6}}{3}$	$-5.3x^{-6.3}$	$3x^{-\frac{7}{3}}$	$3(5k+4)x^{5k+3}$	$-\frac{3x^{-\frac{5}{2}}}{2}=-\frac{3}{2\sqrt{x^5}}$
14.	$-3x^{-4}$	$140x^{-11}$	$-10x^{-15}$	$-2.6x^{-3.6}$	$-\frac{x^{\frac{1}{4}}}{2}$	$-4(\sqrt{3}+7)x^{\sqrt{3}+6}$	$\frac{3x^{\frac{1}{2}}}{2}=\frac{3\sqrt{x}}{2}$
15.	$-10x^{-11}$	$-12x^{-4}$	$\frac{2}{3}x^{-4}$	$-8.3x^{-9.3}$	$\frac{28}{9}x^{\frac{4}{3}}$	$\sqrt{3}(\pi-2)x^{\pi-3}$	$\frac{5x^{-\frac{3}{2}}}{6}=\frac{5}{6\sqrt{x^3}}$
16.	$-7x^{-8}$	$-5x^{-2}$	$-\frac{3}{5}x^{-4}$	$\frac{11x^{\frac{6}{5}}}{5}$	$-\frac{4}{35}x^{-\frac{7}{5}}$	$-2(5i+1)x^{5i}$	$0.3x^{-0.7}$
17.	$-16x^{-17}$	$12x^{-7}$	$\frac{32}{5}x^{-5}$	$\frac{5x^{-\frac{2}{7}}}{7}$	$\frac{18x^{-\frac{13}{5}}}{45}$	$2.5(2e+3)x^{2e+2}$	$0.3x^{-1.7}$
18.	$-9x^{-10}$	$-15x^{-6}$	$-\frac{24}{7}x^{-7}$	$-\frac{9x^{-\frac{11}{2}}}{2}$	$-\frac{1}{2}x^{\frac{4}{3}}$	$2.4fx^{6f-1}$	$-2x^{-\frac{10}{7}}=-\frac{2}{\sqrt[7]{x^{10}}}$
19.	$-6x^{-7}$	$28x^{-5}$	$\frac{15}{8}x^2$	$28x^{-5}$	$-\frac{2}{3}x^{-\frac{8}{5}}$	$2(a+b)x^{a+b-1}$	$\frac{7x^{\frac{4}{3}}}{3}=\frac{7\sqrt[3]{x^4}}{3}$
20.	$-11x^{-12}$	$-12x^{-3}$	$\frac{5}{2}x^{-8}$	$-12x^{-3}$	$-\frac{35x^{-\frac{8}{3}}}{6}$	$(14w+7)x^{4w+1}$	$\frac{28x^{\frac{9}{5}}}{5}=\frac{28\sqrt[5]{x^9}}{5}$

Derivatives of Powers

Find $\frac{dy}{dx}$ if:

1. $y = x + \sqrt{x}$
2. $y = x^5 - 3\sqrt{x}$
3. $y = x^{\frac{5}{2}} - \frac{2}{x}$
4. $y = 3x^4 - \frac{2}{x} + \frac{6}{x^2}$
5. $y = (x+5)(x+2)$
6. $y = (3x+1)(5x-3)$
7. $y = (5x^2-3)(4x^3+x)$
8. $y = (x^3+1)(2x+3)$
9. $y = (x^5-2x)^2$
10. $y = (x-2)(x+1)(3x+1)$
11. $y = (x-a)^3$
12. $y = (2x+3)^3$
13. $y = 2x(3x^2-7x+8)$
14. $y = 3x^2(x+1)(x-2)$
15. $y = (x+\frac{1}{x})^2$
16. $y = \frac{2x+5}{x}$
17. $y = \frac{x^3-2}{x}$
18. $y = \frac{x^2-4x+7}{x}$
19. $y = \frac{x^3-4x^2+3x-2}{x^2}$
20. $y = \frac{3x^7-7x+11}{2x^3}$
21. $y = \frac{(2x+3)(2x-3)}{x}$
22. $y = \frac{x+6}{x^3}$
23. $y = \frac{2x^3+x+4}{2x^5}$
24. $y = \frac{x-3}{\sqrt{x}}$

Find the derivative if:

25. $f(x) = ax^3 + bx^2 + cx + d$
26. $k = \frac{1}{a}(x^2 + \frac{b}{x} + c)$
27. $b = -3m^{-8} + 3\sqrt{7}$
28. $f = ax^4 + bx^2 + c$
29. $r = \frac{t^4}{4} - \frac{t^3}{3} + \frac{t^2}{2} - t + 4$
30. $g = -8h^5 + 3h^{-2} + h^{1.6}$

Find the derivative with respect to the variable indicated:

31. $C = 2\pi r$ [r]
32. $y = mx + b$ [x]
33. $v = u + at$ [t]
34. $A = \pi r^2$ [r]
35. $E = \frac{1}{2}mv^2$ [v]
36. $P = \frac{V^2}{R}$ [V]
37. $s = ut + \frac{1}{2}at^2$ [t]
38. $P = RI^2$ [I]
39. $V = \frac{4\pi r^3}{3}$ [r]
40. $F = \frac{\pi r^4 P}{8\eta l}$ [r]
41. $T = 2\pi\sqrt{\frac{l}{g}}$ [l]
42. $v = \sqrt{\frac{2GM}{r}}$ [r]
43. $F = \frac{GMm}{r^2}$ [r]
44. $E = \frac{q}{4\pi\varepsilon_0 r^2}$ [r]
45. $\Lambda = \sqrt{\frac{h^2}{2\pi mkT}}$ [T]

Calculate the first, second, third, and fourth derivatives of:

46. $y = 4x^4 + 2x^3 + 3$
47. $k = x^3 + 7x - 11$
48. $b = -m^{-2} + 3m^3$

Find the gradient of the tangent to the curve at the point indicated:

49. $y = x^2 + 3x$ (2,10)
50. $y = 2x^3 - 4$ (3,50)
51. $y = -x^2 + \frac{1}{x}$ (-2,-4.5)

Find the equation of the tangent to the curve at the point indicated:

52. $y = 3x^2 - x$ (1,2)
53. $y = x^3 + 4x$ (-1,-5)
54. $y = x^2 - \frac{1}{x}$ (1,2)

Find the value of x that satisfies the equation given:

55. $y = x^4 + 3x + 1$ $\quad y^{IV} - y^{III} + \frac{2y^{II}}{x^2} = 0$
56. $y = \frac{1}{x}, x \neq 0$ $\quad x^3y^{II} + x^3y^{I} + xy = 0$

Derivatives of Powers (Answers)

Answers:

1. $\frac{dy}{dx} = 1 + \frac{1}{2\sqrt{x}}$

2. $\frac{dy}{dx} = 5x^4 - \frac{3}{2\sqrt{x}}$

3. $\frac{dy}{dx} = \frac{5x^{\frac{3}{2}}}{2} + \frac{2}{x^2}$

4. $y' = 12x^3 + \frac{2}{x^2} - \frac{12}{x^3}$

5. $y' = 2x + 7$

6. $y' = 30x - 4$

7. $y' = 100x^4 - 21x^2 - 3$

8. $y' = 8x^3 + 9x^2 + 2$

9. $y' = 10x^9 - 24x^5 + 8x$

10. $y' = 9x^2 - 4x - 7$

11. $y' = 3x^2 - 6ax + 3a^2$

12. $y' = 24x^2 + 72x + 54$

13. $\frac{dy}{dx} = 18x^2 - 28x + 16$

14. $\frac{dy}{dx} = 12x^3 - 9x^2 - 12x$

15. $\frac{dy}{dx} = 2x - \frac{2}{x^3}$

16. $\frac{dy}{dx} = -\frac{5}{x^2}$

17. $\frac{dy}{dx} = 2x + \frac{2}{x^2}$

18. $\frac{dy}{dx} = 1 - \frac{7}{x^2}$

19. $\frac{dy}{dx} = 1 - \frac{3}{x^2} + \frac{4}{x^3}$

20. $\frac{dy}{dx} = 6x^3 + \frac{7}{x^3} - \frac{33}{2x^4}$

21. $\frac{dy}{dx} = 4 + \frac{9}{x^2}$

22. $\frac{dy}{dx} = -\frac{2}{x^3} - \frac{18}{x^4}$

23. $\frac{dy}{dx} = -\frac{2}{x^3} - \frac{2}{x^5} - \frac{10}{x^6}$

24. $\frac{dy}{dx} = \frac{1}{2\sqrt{x}} + \frac{3}{2\sqrt{x^3}}$

25. $f'(x) = 3ax^2 + 2bx + c$

26. $\frac{dk}{dx} = \frac{1}{a}(2x - \frac{b}{x^2})$

27. $\frac{db}{dm} = 24m^{-9}$

28. $\frac{df}{dx} = 4ax^3 + 2bx$

29. $\frac{dr}{dt} = t^3 - t^2 + t - 1$

30. $\frac{dg}{dh} = -40h^4 - \frac{6}{h^3} + 1.6h^{0.6}$

31. $\frac{dC}{dr} = 2\pi$

32. $\frac{dy}{dx} = m$

33. $\frac{dv}{dt} = a$

34. $\frac{dA}{dr} = 2\pi r$

35. $\frac{dE}{dv} = mv$

36. $\frac{dP}{dV} = \frac{2V}{R}$

37. $\frac{ds}{dt} = u + at$

38. $\frac{dP}{dI} = 2RI$

39. $\frac{dV}{dr} = 4\pi r^2$

40. $\frac{dF}{dr} = \frac{\pi r^3 P}{2\eta l}$

41. $\frac{dT}{dl} = \frac{\pi}{\sqrt{gl}}$

42. $\frac{dv}{dr} = -\sqrt{\frac{GM}{2r^3}}$

43. $\frac{dF}{dr} = -\frac{2GMm}{r^3}$

44. $\frac{dE}{dr} = -\frac{q}{2\pi\varepsilon_0 r^3}$

45. $\frac{d\Lambda}{dT} = -\frac{h}{\sqrt{8\pi mkT^3}}$

46. $y' = 16x^3 + 6x^2$
$y'' = 48x^2 + 12x$
$y''' = 96x + 12$
$y^{IV} = 96$

47. $\frac{dk}{dx} = 3x^2 + 7$
$\frac{d^2k}{dx^2} = 6x$
$\frac{d^3k}{dx^3} = 6$
$\frac{d^4k}{dx^4} = 0$

48. $\frac{db}{dm} = 2m^{-3} + 9m^2$
$\frac{d^2b}{dm^2} = -6m^{-4} + 18m$
$\frac{d^3b}{dm^3} = 24m^{-5} + 18$
$\frac{d^4b}{dm^4} = -120m^{-6}$

49. $\frac{dy}{dx} = 2x + 3 = 7$ *at* $(2,1)$

50. $\frac{dy}{dx} = 6x^2 = 54$ at (3,-1)

51. $y' = -2x - \frac{1}{x^2} = \frac{15}{4}$ at (-2,3)

52. $y = 5x - 3$

53. $y = 7x + 2$

54. $y = 3x - 1$

55. $y' = 4x^3 + 3, y'' = 12x^2, y''' = 24x, y^{IV} = 24$ *and*
$x = 2$

56. $y' = -\frac{1}{x^2}, y'' = \frac{2}{x^3}$ *and*
$x = 3$

Chain Rule

Differentiate:

1.	$(x+3)^4$	2.	$(2x+5)^3$	3.	$(1-x)^7$
4.	$(7x-2)^6$	5.	$(x^2+1)^3$	6.	$(x^6+x^3)^{20}$
7.	$(x^2-1)^{100}$	8.	$(3x^2-2x)^2$	9.	$2(5x-3)^8$
10.	$3(x+13)^2$	11.	$(3x^2+7x)^4$	12.	$(x^2+7x-1)^8$
13.	$2(x^7+3x^2-1)^6$	14.	$3(x^5-2x)^2$	15.	$(5x^2+4)^{11}$
16.	$\frac{(x-x^2-x^4)^5}{5}$	17.	$\frac{(5-x)^{-2}}{2}$	18.	$(2x+1)^{-1}$
19.	$(9-4x)^{-3}$	20.	$(4x^2-3x^3+x)^{-2}$	21.	$5(x^2-9)^{-3}$
22.	$(x^{-1}-2x^{-2})^{-3}$	23.	$(3x+1)^{\frac{1}{2}}$	24.	$(6x+1)^{\frac{1}{3}}$
25.	$(5x+7)^{\frac{3}{2}}$	26.	$(x^3-5x^2+x)^{\frac{3}{4}}$	27.	$(x^5-5x)^{\frac{1}{5}}$
28.	$(4x^2-6x+1)^{\frac{7}{3}}$	29.	$(2x^3-9x+12)^{-\frac{2}{3}}$	30.	$(x^{-4}+7x^{-2}+8)^{-\frac{5}{2}}$

Find y':

31.	$y=\frac{1}{3x-1}$	32.	$y=\frac{2}{3x^2-x+5}$	33.	$y=\frac{1}{x^3+x^2+x+1}$
34.	$y=\frac{3}{(7x^2-3x+7)^{10}}$	35.	$y=\frac{1}{9-x^2}$	36.	$y=\frac{7}{3(5x^2+2)^3}$
37.	$y=\frac{3}{4(2x-5)^8}$	38.	$y=\frac{1}{x^4+5x^3-2x}$	39.	$y=\frac{1}{12(4x-1)^3}$
40.	$y=\sqrt{2x+5}$	41.	$y=\sqrt{x^2-3}$	42.	$y=\sqrt[3]{9x-4}$
43.	$y=\sqrt[3]{x^3-3x}$	44.	$y=\sqrt{4-x^2}$	45.	$y=\sqrt[3]{8x^3+27}$
46.	$y=\sqrt[3]{(11-3x)^2}$	47.	$y=\sqrt[3]{(4x-1)^4}$	48.	$y=\sqrt[5]{7x^3-2x^2+5}$

Find the derivative:

49.	$y=\sqrt[4]{7m^3-4m^2+2}$	50.	$c=\sqrt[3]{(4k^2+3)^2}$	51.	$r=\sqrt[4]{(4w+3)^5}$
52.	$d=\frac{3}{\sqrt{x+2}}$	53.	$f=\frac{2}{\sqrt{4e+5}}$	54.	$g(x)=\frac{1}{\sqrt{1-2x}}$
55.	$k(n)=\frac{4}{\sqrt{n^2+6}}$	56.	$p(r)=\frac{12}{\sqrt[4]{(7-r)^5}}$	57.	$q(z)=\frac{5}{\sqrt[5]{z^5-32}}$

Calculate $\frac{dy}{dx}$ for the following functions:

58.	$y=(x-3)\sqrt{(x-3)}$	59.	$y=\sqrt{4-\sqrt{x+2}}$	60.	$y=[(2x+1)^{10}+1]^{10}$
61.	$y=(x-\frac{1}{x})^4$	62.	$y=(x^2+\frac{1}{x^2})^3$	63.	$y=(3x-1)^{2k+1}$
64.	$y=(\frac{x^3}{3}+\frac{x^2}{2}+x)^{-k}$	65.	$y=[(2x+1)^2+(x+1)^2]^3$	66.	$y=\left(7x+\sqrt{x^2+3}\right)^6$
67.	$y=\sqrt{x+\sqrt{x+\sqrt{x}}}$	68.	$y=[(1+\frac{1}{x})^{-1}+1]^{-1}$	69.	$y=[x^2+(x^2+9)^{\frac{1}{2}}]^{\frac{1}{2}}$

Chain Rule (Answers)

1. $4(x+3)^3$
2. $6(2x+5)^2$
3. $-7(1-x)^6$
4. $42(7x-2)^5$
5. $6x(x^2+1)^2$
6. $60x^{59}(2x^3+1)(x^3+1)^{19}$
7. $200x(x^2-1)^{99}$
8. $4x(3x-1)(3x-2)$
9. $80(5x-3)^7$
10. $6(x+13)$
11. $4x^3(6x+7)(3x+7)^3$
12. $8(x^2+7)(x^2+7x-1)^7$
13. $12x(7x^5+6)(x^7+3x^2-1)^5$
14. $6x(x^4-2)(5x^4-2)$
15. $110x(5x^2+4)^{10}$
16. $(1-2x-4x^3)(x-x^2-x^4)^4$
17. $(5-x)^{-3}$
18. $-2(2x+1)^{-2}$
19. $12(9-4x)^{-4}$
20. $-\dfrac{2(8x-9x^2+1)}{(4x^2-3x^3+x)^3}$
21. $-30x(x^2-9)^{-4}$
22. $-3x^5(4-x)(x-2)^{-4}$
23. $\dfrac{3}{2\sqrt{3x+1}}$
24. $2(6x+1)^{-\frac{2}{3}}$
25. $\dfrac{15(5x+7)^{\frac{1}{2}}}{2}$
26. $\dfrac{3(3x^2-10x+1)}{4\sqrt[4]{x^3-5x^2+x}}$
27. $(x^4-1)(x^5-5x)^{-\frac{4}{5}}$
28. $\dfrac{14(4x-3)(4x^2-6x+1)^{\frac{4}{3}}}{3}$
29. $-\dfrac{2(2x^2-3)}{\sqrt[3]{(2x^3-9x+12)^5}}$
30. $\dfrac{5(2-7x^2)}{x^5\sqrt{(x^{-4}+7x^{-2}+8)^7}}$
31. $y'=-\frac{3}{3x-1^2}$
32. $y'=-\dfrac{2(6x-1)}{3x^2-x+5^2}$
33. $y'=-\dfrac{(3x^2+2x+1)}{x^3+x^2+x+1^2}$
34. $y'=-\dfrac{30(14x-3)}{(7x^2-3x+7)^{11}}$
35. $y'=\dfrac{2x}{(9-x^2)^2}$
36. $y'=-\dfrac{70x}{(5x^2+2)^4}$
37. $y'=-\dfrac{12}{(2x-5)^9}$
38. $y'=\dfrac{(4x^3+15x-2)}{(x^4+5x^3-2x)^2}$
39. $y'=-\dfrac{1}{(4x-1)^4}$
40. $y'=\dfrac{1}{\sqrt{2x+5}}$
41. $y'=\dfrac{x}{\sqrt{x^2-3}}$
42. $y'=\dfrac{3}{\sqrt[3]{(9x-4)^2}}$
43. $y'=\dfrac{x^2-1}{\sqrt[3]{(x^3-3x)^2}}$
44. $y'=-\dfrac{x}{\sqrt{4-x^2}}$
45. $y'=\dfrac{8x^2}{\sqrt[3]{(8x^3+27)^2}}$
46. $y'=-\dfrac{2}{\sqrt[3]{11-3x}}$
47. $y'=\dfrac{16\sqrt[3]{4x-1}}{3}$
48. $y'=\dfrac{x(21x-4)}{5\sqrt[5]{(7x^3-2x^2+5)^4}}$
49. $\dfrac{dy}{dm}=\dfrac{21m^2-8m}{4\sqrt[4]{(7m^3-4m^2+2)^3}}$
50. $\dfrac{dc}{dk}=\dfrac{16k}{3\sqrt[3]{4k^2+3}}$
51. $\dfrac{dr}{dw}=5\sqrt[4]{4w+3}$
52. $\dfrac{dd}{dx}=-\dfrac{3}{2\sqrt{(x+2)^3}}$
53. $\dfrac{df}{de}=-\dfrac{4}{\sqrt{(4e+5)^3}}$
54. $g'(x)=\dfrac{1}{\sqrt{(1-2x)^3}}$
55. $k'(n)=-\dfrac{4n}{\sqrt{(n^2+6)^3}}$
56. $p'(r)=\dfrac{15}{\sqrt[4]{(7-r)^9}}$
57. $q'(z)=-\dfrac{5z^4}{\sqrt[5]{(z^5-32)^6}}$
58. $\dfrac{dy}{dx}=\dfrac{3\sqrt{(x-3)}}{2}$
59. $\dfrac{dy}{dx}=\dfrac{1}{4\sqrt{4-\sqrt{x+2}}.\sqrt{x+2}}$
60. $\dfrac{dy}{dx}=200(2x+1)^9.[(2x+1)^{10}+1]^9$
61. $\dfrac{dy}{dx}=4(1+\dfrac{1}{x^2})(x-\dfrac{1}{x})^3$
62. $\dfrac{dy}{dx}=6(x-\dfrac{1}{x^3})(x^2+\dfrac{1}{x^2})^2$
63. $\dfrac{dy}{dx}=3(2k+1)(3x-1)^{2k}$
64. $\dfrac{dy}{dx}=-\dfrac{k(x^2+x+1)}{(\frac{x^3}{3}+\frac{x^2}{2}+x)^{k+1}}$
65. $\dfrac{dy}{dx}=6(5x+3).[(2x+1)^2+(x+1)^2]^2$
66. $\dfrac{dy}{dx}=6\left(7x+\sqrt{x^2+3}\right)^5.[7+\dfrac{x}{\sqrt{x^2+3}}]$
67. $\dfrac{dy}{dx}=\dfrac{1+\dfrac{1+\frac{1}{2\sqrt{x}}}{2\sqrt{x+\sqrt{x}}}}{2\sqrt{x+\sqrt{x+\sqrt{x}}}}$
68. $\dfrac{dy}{dx}=\dfrac{1}{x^2(1+\frac{1}{x})^2[(1+\frac{1}{x})^{-1}+1]^2}$
69. $\dfrac{dy}{dx}=\dfrac{2x+\dfrac{x}{\sqrt{x^2+9}}}{\sqrt{x^2+(x^2+9)^{\frac{1}{2}}}}$

Product Rule

Expand and differentiate, and compare by differentiating using the product rule:

1.	$(3x-1)(4x+3)$	2.	$5x(6x-1)$	3.	$x^3(3x+2)$
4.	$4x^2(2x^3-1)$	5.	$3x^4(7x-2)$	6.	$(2x^2-5)(3x^2+8x)$

Differentiate mentally without simplification (check your answers below):

7.	$(3x-2)(5x+7)$	8.	$(4-x^2)(3x+5)$	9.	$(5x-2)(x-3)$
10.	$(5x-2)(x-1)$	11.	$(x^2-1)(x+7)$	12.	$(x^4+8)(6-5x)$
13.	$(2x+9)(x^2-4)$	14.	$(3x+5)(8x-1)$	15.	$(9x^2-5)(3x-8)$
16.	$5x^2(3x-8)$	17.	$4x^5(2x^2-5x+3)$	18.	$(x^2-7x+1)(3x-4)$
19.	$(5x^3+2)(4x-x^3)$	20.	$(x^2+3x+1)(x^3-4x+5)$	21.	$(x^{100}-5)(5x^8-11x+1)$

Differentiate on paper without simplification:

22.	$(x+2)(x^2-2x+7)$	23.	$(1-x^3)(7x+4)$	24.	$(3x-5)(x^3+2x^2-8)$
25.	$(x^2-2)(5x-x^3)$	26.	$(x^2+3x-1)(x^3-4x+7)$	27.	$(x^3-2x+8)(6-5x)$
28.	$(8x^2-5x)(13x^2-4)$	29.	$(x^5-2x^3)(7x^2+x-8)$	30.	$(3-x^3)(8x+1)$
31.	$(x+1)(x+2)(x+3)$	32.	$(x+1)(x^2+2)(x^3+3)$	33.	$4x(x-1)(2x-3)$

Differentiate (using an embedded chain rule):

34.	$x^2(x+1)^3$	35.	$8x(3x-2)^5$	36.	$2x^4(3-x)^3$
37.	$(x+1)(2x-5)^4$	38.	$(x^3-4x^2+1)(x^2+3)^5$	39.	$(3x^2-2x-1)(x^2+5x-2)^2$
40.	$4x^3(2x-1)^{-3}$	41.	$2x^8(11-x)^{-7}$	42.	$(4x-3)(5x+3)^{-2}$
43.	$2x(x+3)^{\frac{1}{2}}$	44.	$(3x-1)(6-x)^{-\frac{3}{4}}$	45.	$4x^3(8x-1)^{\frac{4}{5}}$
46.	$(2x+9)\sqrt{x^2-4}$	47.	$x\sqrt{11-x}$	48.	$x^2\sqrt[3]{4x-7}$
49.	$(3x+8)^3(x+1)^4$	50.	$(2x^2-3x+1)^2(5x-1)^4$	51.	$(5x+3)^4(x-2)^{-\frac{1}{2}}$
52.	$(8x+1)^3\sqrt[4]{9-x^3}$	53.	$(2x+7)^8\sqrt[3]{(x+4)^5}$	54.	$(6x+1)^8(3x-7)^{-5}$
55.	$\dfrac{x}{(2x-9)^5}$	56.	$(x+x^{-1})(x-x^{-2})$	57.	$(x+\sqrt{x})(1+\sqrt{x})$

Find the equation of the tangent to the following curves at the point indicated):

58.	$y=x^2(x+1)^3$	(-1,0)	59.	$y=x(3x-2)^2$	(1,1)	60.	$y=(x+1)\sqrt{x+3}$	(-2,-1)

Solutions to the Mental Calculations Above

7.	$3(5x+7)+(3x-2)5$	8.	$-2x(3x+5)+(4-x^2)3$	9.	$5(x-3)+(5x-2)1$
10.	$5(x-1)+(5x-2)1$	11.	$2x(x+7)+(x^2-1)1$	12.	$4x^3(6-5x)+(x^4+8)(-5)$
13.	$2(x^2-4)+(2x+9)(2x)$	14.	$3(8x-1)+(3x+5)8$	15.	$18x(3x-8)+(9x^2-5)3$
16.	$10x(3x-8)+5x^2(3)$	17.	$20x^4(2x^2-5x+3)+4x^5(4x-5)$	18.	$(2x-7)(3x-4)+(x^2-7x+1)3$
19.	$15x^2(4x-x^3)+(5x^3+2)(4-3x^2)$	20.	$(2x+3)(x^3-4x+5)+(x^2+3x+1)(3x^2-4)$	21.	$100x^{99}(5x^8-11x+1)+(x^{100}-5)(40x^7-11)$

Product Rule (Answers)

Simplified answers:

1. $24x + 5$
2. $60x - 5$
3. $12x^3 + 6x^2$
4. $40x^4 - 8x$
5. $105x^5 - 24x^3$
6. $24x^3 + 48x^2 - 30x - 40$

Answers with structure, but no simplification:

22. $1(x^2 - 2x + 7) + (x + 2)(2x - 2)$
23. $-3x^2(7x + 4) + (1 - x^3)(7)$
24. $3(x^3 + 2x^2 - 8) + (3x - 5)(3x^2 + 4x)$
25. $2x(5x - x^3) + (x^2 - 2)(5 - 3x^2)$
26. $(2x + 3)(x^3 - 4x + 7) + (x^2 + 3x - 1)(3x^2 - 4)$
27. $(3x^2 - 2)(6 - 5x) + (x^3 - 2x + 8)(-5)$
28. $(16x - 5)(13x^2 - 4) + (8x^2 - 5x)(26x)$
29. $(5x^4 - 6x^2)(7x^2 + x - 8) + (x^5 - 2x^3)(14x + 1)$
30. $-3x^2(8x + 1) + (3 - x^3)(8)$
31. $(1)(x + 2)(x + 3) + (x + 1)(1)(x + 3) + (x + 1)(x + 2)(1)$
32. $(1)(x^2 + 2)(x^3 + 3) + (x + 1)(2x)(x^3 + 3) + (x + 1)(x^2 + 2)(3x^2)$
33. $4(x - 1)(2x - 3) + 4x(1)(2x - 3) + 4x(x - 1)(2)$

Simplified answers (equivalent expressions use or remove radicals or negative indices):

34. $x(5x + 2)(x + 1)^2$
35. $16(9x - 1)(3x - 2)^4$
36. $-2x^3(7x - 12)(3 - x)^2$
37. $(10x + 3)(2x - 5)^3$
38. $x(13x^3 - 48x^2 + 9x - 14).(x^2 + 3)^4$
39. $2(x^2 + 5x - 2).(5x^3 + 29x^2 + 10x - 8)$
40. $-12x^2(2x - 1)^{-4}$
41. $-2x^7(x - 88)(11 - x)^{-8}$
42. $-2(10x - 21)(5x + 3)^{-3}$
43. $3(x + 2)(x + 3)^{-\frac{1}{2}}$
44. $-\dfrac{3(4x^2 - 27x + 1)(6 - x)^{-\frac{7}{4}}}{4}$
45. $\dfrac{4x^2(152x - 15)(8x - 1)^{-\frac{1}{5}}}{5}$
46. $\dfrac{4x^2 + 9x - 8}{\sqrt{x^2 - 4}}$
47. $-\dfrac{(3x - 22)}{2\sqrt{11 - x}}$
48. $\dfrac{14x(2x - 3)}{3\sqrt[3]{(4x - 7)^2}}$
49. $(21x + 41)(3x + 8)^2.(x + 1)^3$
50. $2(2x^2 - 3x + 1).(40x^2 - 49x + 13)(5x - 1)^3$
51. $\dfrac{(35x - 43)(5x + 3)^3}{2\sqrt{(x - 1)^3}}$
52. $-\dfrac{(8x + 1)^2(120x^3 + 3x^2 - 864)}{4\sqrt[4]{(9 - x^3)^3}}$
53. $\dfrac{1}{3}(58x + 227)(2x + 7)^7.\sqrt[3]{(x + 4)^2}$
54. $\dfrac{3(18x - 117)(6x + 1)^7}{(3x - 7)^6}$
55. $-\dfrac{8x + 9}{(2x - 9)^6}$
56. $2x + \dfrac{1}{x^2} + \dfrac{3}{x^4}$
57. $2 + \dfrac{3\sqrt{x}}{2} + \dfrac{1}{2\sqrt{x}}$

The equations of the tangents are:

58. $y = 0$
59. $y = 7x - 6$
60. $y = \dfrac{x}{2}$

Quotient Rule

Divide each term of the numerator by the denominator before differentiating. Compare this with the result you get by using the quotient rule:

1. $\dfrac{x+6}{x}$ 2. $\dfrac{x^3-1}{x^2}$ 3. $\dfrac{3x-1}{x^2}$

4. $\dfrac{x^3-3x^2}{x}$ 5. $\dfrac{x^4+2x^3}{x^2}$ 6. $\dfrac{2x^2+5x-1}{x}$

Using negative indices, differentiate by using the product rule. Compare this with the result you get by using the quotient rule:

7. $\dfrac{x+5}{2x+1}$ 8. $\dfrac{x-3}{5x+2}$ 9. $\dfrac{3x+8}{x-5}$

10. $\dfrac{x^2-2}{x^2+9}$ 11. $\dfrac{x^3}{x+4}$ 12. $\dfrac{x+1}{3x^2-7}$

Differentiate mentally without simplification (check your answers on the following page):

13. $\dfrac{1}{2x-1}$ 14. $\dfrac{x^3}{x^2-4}$ 15. $\dfrac{x+4}{x-6}$

16. $\dfrac{2x+5}{4x-3}$ 17. $\dfrac{x}{2x^2-8}$ 18. $\dfrac{x-7}{x^2}$

19. $\dfrac{x^2+4x-1}{x+3}$ 20. $\dfrac{x^2-9x+11}{2x+5}$ 21. $\dfrac{3x-1}{x^2+12}$

22. $\dfrac{6x+7}{x^2-x+3}$ 23. $\dfrac{x^3+x}{x^2-x-1}$ 24. $\dfrac{5x^2-2x}{3x+1}$

Differentiate (using an embedded chain rule):

25. $\dfrac{2x}{(x+1)^{\frac{1}{2}}}$ 26. $\dfrac{(2x+7)^3}{4x-1}$ 27. $\dfrac{x-1}{(7x+3)^4}$

28. $\dfrac{(3x-4)^5}{(2x+1)^3}$ 29. $\dfrac{2x-5}{\sqrt{x+1}}$ 30. $\dfrac{\sqrt{x-1}}{4x+1}$

31. $\dfrac{\sqrt{x^2+1}}{(x-8)^2}$ 32. $\dfrac{x-4}{\sqrt[3]{x}}$ 33. $\dfrac{(x+3)^4}{x^2}$

Find the derivative of:

34. $y=\dfrac{x+3}{\sqrt{x}+2}$ 35. $f=\dfrac{p^{\frac{2}{3}}}{2p+1}$ 36. $b=\dfrac{\sqrt[3]{w}}{w^2+5}$

37. $m=\dfrac{h^3-1}{h^3+1}$ 38. $g=\dfrac{7t^4+11}{t+8}$ 39. $e=\left(\dfrac{4y+3}{5y-1}\right)^3$

40. $k=\dfrac{8n^2-5n+11}{n+2}$ 41. $r=\dfrac{v-6}{\sqrt{(v+1)^5}}$ 42. $z=\dfrac{4a-9}{(a+5)^{\frac{3}{4}}}$

Find the equation of the tangent to the following curves at the point indicated):

43. $y=\frac{x+3}{x-1}$ (2,5) 44. $y=\frac{x+3}{3x-2}$ (1,4) 45. $y=\frac{x}{x-2}$ (3,3)

46. $y=\frac{x^2}{x-1}$ (-1,-½) 47. $y=\frac{x^2+1}{x^2-3}$ (2,5) 48. $y=\frac{x-7}{x^2}$ (1,-6)

Quotient Rule (Answers)

You should discover that the results are the same and that it is advisable to simplify expressions first!

1. $-\dfrac{6}{x^2}$
2. $\dfrac{x^3+2}{x^3} = 1 + \dfrac{2}{x^3}$
3. $-\dfrac{3x-2}{x^3}$
4. $2x-3$
5. $2x+2$
6. $\dfrac{2x^2+1}{x^2} = 2 + \dfrac{1}{x^2}$

You should discover that the results are the same and that the quotient rule is (usually) simpler to use.

7. $-\dfrac{9}{(2x+1)^2}$
8. $\dfrac{17}{(5x+2)^2}$
9. $-\dfrac{23}{(x-5)^2}$
10. $\dfrac{22x}{(x^2+9)^2}$
11. $\dfrac{2x^2(x+6)}{(x+4)^2}$
12. $-\dfrac{3x^2+6x+7}{(3x^2-7)^2}$

You should have found the following structures:

13. $\dfrac{0(2x-1)-1(2)}{(2x-1)^2}$
14. $\dfrac{3x^2(x^2-4)-x^3(2x)}{(x^2-4)^2}$
15. $\dfrac{1(x-6)-(x+4)1}{(x-6)^2}$
16. $\dfrac{2(4x-3)-(2x+5)4}{(4x-3)^2}$
17. $\dfrac{1(2x^2-8)-x(4x)}{(2x^2-8)^2}$
18. $\dfrac{1(x^2)-(x-7)(2x)}{x^4}$
19. $\dfrac{(2x+4)(x+3)-(x^2+4x-1)1}{(x+3)^2}$
20. $\dfrac{(2x-9)(2x+5)-(x^2-9x+11)2}{(2x+5)^2}$
21. $\dfrac{3(x^2+12)-(3x-1)(2x)}{(x^2+12)^2}$
22. $\dfrac{6(x^2-x-3)-(6x+7)(2x-1)}{(x^2-x+3)^2}$
23. $\dfrac{(3x^2+1)(x^2-x-1)-(x^3+x)(2x-1)}{(x^2-x-1)^2}$
24. $\dfrac{(10x-2)(3x+1)-(5x^2-2x)3}{(3x+1)^2}$

The derivatives are:

25. $\dfrac{x+2}{(x+1)^{\frac{3}{2}}}$
26. $\dfrac{2(8x-17)(2x+7)^2}{(4x-1)^2}$
27. $-\dfrac{21x-31}{(7x+3)^5}$
28. $\dfrac{3(4x+13)(3x-4)^4}{(2x+1)^4}$
29. $\dfrac{2x+9}{2\sqrt{(x+1)^3}}$
30. $-\dfrac{4x-9}{2\sqrt{x-1}(4x+1)^2}$
31. $-\dfrac{x^2+8x+2}{\sqrt{x^2+1}(x-8)^3}$
32. $\dfrac{2(x+2)}{3\sqrt[3]{x^4}}$
33. $\dfrac{2(x-3)(x+3)^3}{x^3}$

The derivatives are:

34. $\dfrac{dy}{dx} = \dfrac{x+4\sqrt{x}-3}{2\sqrt{x}(\sqrt{x}+2)^2}$
35. $\dfrac{df}{dp} = -\dfrac{2(p-1)}{3\sqrt[3]{p}(2p+1)^2}$
36. $\dfrac{db}{dw} = -\dfrac{5(w^2-1)}{3w^{\frac{2}{3}}(w^2+5)^2}$
37. $\dfrac{dm}{dh} = \dfrac{6h^2}{(h^3+1)^2}$
38. $\dfrac{dg}{dt} = \dfrac{21t^4+224t^3-11}{(t+8)^2}$
39. $\dfrac{de}{dy} = -\dfrac{57(4y+3)^2}{(5y-1)^4}$
40. $\dfrac{dk}{dn} = \dfrac{8n^2+32n-21}{(n+2)^2}$
41. $\dfrac{dr}{dv} = -\dfrac{3v-32}{2\sqrt{(v+1)^7}}$
42. $\dfrac{dz}{da} = \dfrac{4a+107}{4(a+5)^{\frac{7}{4}}}$

The equations of the tangents (in gradient-intercept form) are:

43. $y=-4x+13$
44. $y=-11x+15$
45. $y=-2x+9$
46. $y=\dfrac{3x}{4}+\dfrac{1}{4}$
47. $y=-16x+37$
48. $y=13x-19$

Exponential Functions

Differentiate with respect to x:

1. $y = e^x$
2. $y = 3e^x$
3. $y = e^{4x}$
4. $y = 2e^{5x}$
5. $y = e^{5x} - e^{2x}$
6. $y = 2e^{3x} + e^{-x}$
7. $y = e^{3.5x} + e^{1.9x}$
8. $y = 6e^{2x} - \frac{e^{-2x}}{2}$
9. $y = e^{2x}.e^{7x}$
10. $y = e^{x^2}$
11. $y = e^{x^2-2x+7}$
12. $y = 3e^{-x^4}$
13. $y = 8e^{5x-1}$
14. $y = x^3e^{2x}$
15. $y = (2x+1)e^{-x}$

Find the derivative function:

16. $y = xe^{-2x}$
17. $y = x^3e^{-x}$
18. $y = x^3 - xe^{4x}$
19. $y = (x^2-6)e^{8x}$
20. $y = \sqrt{x}e^x$
21. $y = 4e^{2x^2}$
22. $y = xe^{x^2}$
23. $y = e^{(e^x)}$
24. $y = \frac{e^{2x+1}}{2x+7}$
25. $y = \frac{e^{3x}}{x^2}$
26. $y = e^{\sqrt{x}}$
27. $y = \frac{e^x+1}{e^x-1}$

Calculate $\frac{dy}{dx}$:

28. $y = \frac{x}{e^{-x}}$
29. $y = (e^x+2)^8$
30. $y = e^{\sqrt[4]{x}} - e^{-\frac{1}{x}}$
31. $y = e^{\sqrt{x}} + e^{\frac{5}{4}}$
32. $y = 4x^3 + 3x^2 - e^{-2x}$
33. $y = 2e^{1-x}$
34. $y = (e^x+x)^{10}$
35. $y = e^{x^3+1}$
36. $y = x^ee^x$

Find the derivative of:

37. $y = x^ee^{x-e}$
38. $p = m^2e^{-\pi m}$
39. $a = \frac{k-1}{e^{2k}-1}$
40. $v = (t^2-3t)e^{8t}$
41. $j = \sqrt{d}e^{d+4}$
42. $b = \frac{e^{2q}}{e^q+4}$
43. $f = \frac{e^g - e^{-g}}{2}$
44. $h = A + Be^{-6w}$
45. $l = \frac{1-n^2}{2e^n}$

Given the function on the left, demonstrate that the relationship on the right is true:

46. $y = e^x + e^{-x}$ — $y'' = y$
47. $y = 4e^{-x} + 5e^{-3x}$ — $y'' + 4y' + 3y = 0$
48. $y = e^{2x} + e^{8x}$ — $y'' - 10y' + 16y = 0$
49. $y = e^{2x} + e^{4x}$ — $y'' - 6y' + 8y = 0$
50. $y = (x+1)e^{5x}$ — $y'' - 10y' + 25y = 0$
51. $y = A + Be^{-4x}$ — $y'' + 4y' = 0$

Find the equation of the tangent to the following curves at the point indicated):

52. $y = e^x$ (0,1)
53. $y = e^{-x}$ (0,1)
54. $y = e^{x-2}$ (3,e)

Find the minimum value of each function (and its location) given:

55. $y = (x-2)e^x$
56. $y = -e^{-x^2}$
57. $y = xe^{x-1}$

Exponential Functions (Answers)

The required derivatives are:

1. $y' = e^x$
2. $y' = 3e^x$
3. $y' = 4e^{4x}$
4. $y' = 10e^{5x}$
5. $y' = 5e^{5x} - 2e^{2x}$
6. $y' = 6e^{3x} - e^{-x}$
7. $y' = 3.5e^{3.5x} + 1.9e^{1.9x}$
8. $y' = 12e^{2x} + e^{-2x}$
9. $y' = 9e^{9x}$
10. $y' = 2xe^{x^2}$
11. $y' = (2x - 2)e^{x^2-2x+7}$
12. $y' = -12x^3e^{-x^4}$
13. $y' = 40e^{5x-1}$
14. $y' = 3x^2e^{2x} + 2x^3e^{2x}$
15. $y' = e^{-x} - 2xe^{-x}$
16. $y' = e^{-2x} - 2xe^{-2x}$
17. $y' = 3x^2e^{-x} - x^3e^{-x}$
18. $y' = 3x^2 - e^{4x} - 4xe^{4x}$
19. $y' = 2(4x^2 + x - 24)e^{8x}$
20. $y' = \dfrac{e^x}{2\sqrt{x}} + \sqrt{x}e^x$
21. $y' = 16xe^{2x^2}$
22. $y' = e^{x^2} + 2x^2e^{x^2}$
23. $y' = e^x e^{(e^x)}$
24. $y' = \dfrac{4(x+3)e^{2x+1}}{2x+7^2}$
25. $y' = \dfrac{(3x-2)e^{3x}}{x^3}$
26. $y' = \dfrac{e^{\sqrt{x}}}{2\sqrt{x}}$
27. $y' = -\dfrac{2e^x}{(e^x - 1)^2}$
28. $\dfrac{dy}{dx} = e^x + xe^x$
29. $\dfrac{dy}{dx} = 8e^x(e^x + 2)^7$
30. $\dfrac{dy}{dx} = \dfrac{e^{\sqrt[4]{x}}}{4\sqrt[4]{x^3}} - \dfrac{e^{-\frac{1}{x}}}{x^2}$
31. $\dfrac{dy}{dx} = \dfrac{e^{\sqrt{x}}}{2\sqrt{x}}$
32. $\dfrac{dy}{dx} = 12x^2 + 6x + 2e^{-2x}$
33. $\dfrac{dy}{dx} = -2e^{1-x}$
34. $\dfrac{dy}{dx} = 10(e^x + 1)(e^x + x)^9$
35. $\dfrac{dy}{dx} = 3x^2e^{x^3+1}$
36. $\dfrac{dy}{dx} = (x + e)x^{e-1}e^x$
37. $\dfrac{dy}{dx} = (x + e)x^{e-1}e^{x-e}$
38. $\dfrac{dp}{dm} = (2 - \pi m)me^{-\pi m}$
39. $\dfrac{da}{dk} = \dfrac{3e^{2k} - 2ke^{2k} - 1}{(e^{2k} - 1)^2}$
40. $\dfrac{dv}{dt} = (8t^2 - 22t - 3)e^{8t}$
41. $\dfrac{dj}{dd} = \dfrac{(2d+1)e^{d+4}}{2\sqrt{d}}$
42. $\dfrac{db}{dq} = \dfrac{e^{3q} + 8e^{2q}}{(e^q + 4)^2}$
43. $\dfrac{df}{dg} = \dfrac{e^g + e^{-g}}{2}$
44. $\dfrac{dh}{dw} = -6Be^{-6w}$
45. $\dfrac{dl}{dn} = \dfrac{n^2 - 2n - 1}{2e^n}$

The derivatives required in order to demonstrate the relationships are:

46. $y = e^x + e^{-x}$ $\quad y' = e^x - e^{-x}$ $\quad y'' = e^x + e^{-x}$
47. $y = 4e^{-x} + 5e^{-3x}$ $\quad y' = -4e^{-x} - 15e^{-3x}$ $\quad y'' = 4e^{-x} + 45e^{-3x}$
48. $y = e^{2x} + e^{8x}$ $\quad y' = 2e^{2x} + 8e^{8x}$ $\quad y'' = 4e^{2x} + 64e^{8x}$
49. $y = e^{2x} + e^{4x}$ $\quad y' = 2e^{2x} + 4e^{4x}$ $\quad y'' = 4e^{2x} + 16e^{4x}$
50. $y = (x+1)e^{5x} = e^{5x} + xe^{5x}$ $\quad y' = 6e^{5x} + 5xe^{5x}$ $\quad y'' = 35e^{5x} + 25xe^{5x}$
51. $y = A + Be^{-4x}$ $\quad y' = -4Be^{-4x}$ $\quad y'' = 16Be^{-4x}$

The equations of the tangents (in gradient-intercept form) are:

52. $y = x + 1$
53. $y = -x + 1$
54. $y = ex - 2e$

The minimum values of the functions are:

55. $y = -e \quad (at\ x = 1)$
56. $y = -1 \quad (at\ x = 0)$
57. $y = -\dfrac{1}{e^2} \quad (at\ x = -1)$

Logarithmic Functions

Differentiate with respect to x:

1. $y = log_e x$
2. $y = log_e(3x)$
3. $y = 2log_e x$
4. $y = ln(3x+1)$
5. $y = ln\,3x + 1$
6. $y = lnx + 2x$
7. $y = 5ln(3x)$
8. $y = ln(x^2)$
9. $y = ln(x^5)$
10. $y = ln(x^2 - 5)$
11. $y = 2lnx + 5ln(2x)$
12. $y = 7x - ln(4x^3)$

Simplify, using logarithmic laws, before finding the derivative function:

13. $y = ln\sqrt{x+9}$
14. $y = ln\frac{x+1}{x+3}$
15. $y = ln(x-5)(x+8)$

Calculate the derivative:

16. $y = ln(2x-1)(x+8)$
17. $y = ln(x+6)^4$
18. $y = ln\frac{(x+1)(x+2)}{(x+3)}$
19. $y = ln\frac{x}{x-2}$
20. $y = ln\sqrt{x+4}$
21. $y = ln\sqrt{(x+1)^3}$
22. $y = ln\frac{1}{x}$
23. $y = ln\frac{1}{3x+2}$
24. $y = ln\frac{1}{2x^5}$

Calculate $\frac{dy}{dx}$:

25. $y = xlnx$
26. $y = 2x^3 ln\,(x+4)$
27. $y = xlnx - 3x$
28. $y = ln(x^2)$
29. $y = (lnx)^2$
30. $y = ln(lnx)$
31. $y = (1 + lnx)^5$
32. $y = (lnx - x)^9$
33. $y = (x^2 + lnx)^6$
34. $y = \frac{lnx}{x-2}$
35. $y = (2x+1)lnx$
36. $y = x^3 ln(x+1)$
37. $y = \log(x)$
38. $y = \log_7(5x)$
39. $y = log\left(\frac{2x^2-1}{\sqrt{x}}\right)$
40. $y = \frac{e^{2x}}{lnx}$
41. $y = e^x lnx$
42. $y = ln\frac{e^x+1}{e^x-1}$

Given the function on the left, solve the equation on the right:

43. $y = lnx$ — $xy'' + (y')^2 = 2$
44. $y = (lnx)^2$ — $xy'' + y' = 1$
45. $y = xlnx$ — $y'' + xy' - y = 2$
46. $y = x^2 lnx$ — $xy'' - y' = 8$
47. $y = (x+3)lnx$ — $xy'' + y' = 3$
48. $y = ln(x^2 - 1)$ — $2(x^2-1)y'' + 5y' = 0$
49. $y = e^x lnx$ — $y'' - y' = 0$

Find the equation of the tangent to the following curves at the point indicated):

50. $y = ln\sqrt{2-x}$ (1,0)
51. $y = xlnx$ (e,e)
52. $y = e^x + lnx$ (1,e)

Find the minimum value of each function (and its location) given:

53. $y = (lnx)^2$
54. $y = x^2 - lnx$
55. $y = \frac{1}{x} + lnx$

Logarithmic Functions (Answers)

The derivatives are:

1. $y' = \frac{1}{x}$
2. $y' = \frac{1}{x}$
3. $y' = \frac{2}{x}$
4. $y' = \frac{3}{3x+1}$
5. $y' = \frac{1}{x} + 1$
6. $y' = \frac{1}{x} + 2$
7. $y' = \frac{5}{x}$
8. $y' = \frac{2}{x}$
9. $y' = \frac{5}{x}$
10. $y' = \frac{2x}{x^2-5}$
11. $y' = \frac{7}{x}$
12. $y' = 7 - \frac{3}{x}$
13. $y' = \frac{1}{2(x+9)}$
14. $y' = \frac{1}{x+1} - \frac{1}{x+3}$
15. $y' = \frac{1}{x-5} + \frac{1}{x+8}$
16. $y' = \frac{2}{2x-1} + \frac{1}{x+8}$
17. $y' = \frac{4}{x+6}$
18. $y' = \frac{1}{x+1} + \frac{1}{x+2} - \frac{1}{x+3}$
19. $y' = \frac{1}{x} - \frac{1}{x-2}$
20. $y' = \frac{1}{2(x+4)}$
21. $y' = \frac{3}{2(x+1)}$
22. $y' = -\frac{1}{x}$
23. $y' = -\frac{3}{3x+2}$
24. $y' = -\frac{5}{x}$
25. $\frac{dy}{dx} = lnx + 1$
26. $\frac{dy}{dx} = 6x^2 ln\,(x+4) + \frac{2x^3}{x+4}$
27. $\frac{dy}{dx} = lnx - 2$
28. $\frac{dy}{dx} = \frac{2}{x}$
29. $\frac{dy}{dx} = \frac{2lnx}{x}$
30. $\frac{dy}{dx} = \frac{1}{xlnx}$
31. $\frac{dy}{dx} = \frac{5(1+lnx)^4}{x}$
32. $\frac{dy}{dx} = 9(lnx - x)^8(\frac{1}{x} - 1)$
33. $\frac{dy}{dx} = 6(x^2 + lnx)^5(2x + \frac{1}{x})$
34. $\frac{dy}{dx} = \frac{(x-2) - xlnx}{x(x-2)^2}$
35. $\frac{dy}{dx} = 2lnx + \frac{2x+1}{x}$
36. $\frac{dy}{dx} = 3x^2 l\,n(x+1) + \frac{x^3}{x+1}$
37. $\frac{dy}{dx} = \frac{1}{xln10}$
38. $\frac{dy}{dx} = \frac{1}{xln7}$
39. $\frac{dy}{dx} = \frac{6x^2+1}{2ln10.x(2x^2-1)}$
40. $\frac{dy}{dx} = \frac{(2xlnx-1)e^{2x}}{xln^2x}$
41. $\frac{dy}{dx} = e^x lnx + \frac{e^x}{x}$
42. $\frac{dy}{dx} = -\frac{2e^x}{e^{2x}-1}$

The solutions to the equations are:

43.	$y = lnx$	$xy'' + (y')^2 = 2$	(½, -*ln*2)
44.	$y = (lnx)^2 = ln^2x$	$xy'' + y' = 1$	(2, $ln^2 2$)
45.	$y = xlnx$	$y'' + xy' - y = 2$	(1, 0)
46.	$y = x^2 lnx$	$xy'' - y' = 8$	(4, 16*ln*4)
47.	$y = (x+3)lnx$	$xy'' + y' = 3$	(e, e+3)
48.	$y = ln(x^2-1)$	$2(x^2-1)y'' + 5y' = 0$	(2, *ln*3)
49.	$y = e^x lnx$	$y'' - y' = 0$	(1, 0)

The equations of the tangents (in gradient-intercept form) are:

50. $y = -\frac{x}{2} + \frac{1}{2}$
51. $y = 2x - e$
52. $y = (e+1)x - 1$

The minimum values of the functions are:

53. $y = 0 \quad (at\ x = 1)$
54. $y = \frac{1}{2}(1 + ln2) \quad (at\ x = \frac{1}{\sqrt{2}})$
55. $y = 1 \quad (at\ x = 1)$

Trigonometric Functions

Differentiate with respect to x:

1. $y = \tan x^2$
2. $y = \tan^2 x$
3. $y = \sin 3x$
4. $y = 2\cos x$
5. $y = \sin x \cos x$
6. $y = \sin x + 5\cos x$
7. $y = \tan 2x$
8. $y = x \sin x$
9. $y = \sin 2x \tan 3x$
10. $y = \dfrac{\sin x}{x}$
11. $y = \dfrac{x}{\cos x}$
12. $y = \dfrac{\sin x}{x^2}$
13. $y = \sin\left(x + \dfrac{\pi}{4}\right)$
14. $y = x^2 + \tan\dfrac{x}{2}$
15. $y = \cos\dfrac{\pi x}{3}$
16. $y = \dfrac{\cos x}{1 + \sin x}$
17. $y = \dfrac{3x + 4}{\sin 5x}$
18. $y = x^2 \cos x$
19. $y = \sin x^3$
20. $y = \cos^3 x$
21. $y = x^3 \tan 8x$
22. $y = \sin^2 4x$
23. $y = \cos(x^2 + 1)$
24. $y = \tan \pi x$
25. $y = \cos\sqrt{x}$
26. $y = \sqrt{\cos 2x}$
27. $y = \sin^4 \sqrt{x}$
28. $y = \tan^3 x$
29. $y = \sin^2 x + \cos^2 x$
30. $y = (\sin x + \cos x)^2$
31. $y = \tan(x^2 - 1)$
32. $y = \tan(\pi - x)$
33. $y = (2x + \tan 7x)^9$
34. $y = 7\tan(x^2 + 5)$
35. $y = \sin(\cos x)$
36. $y = \cos(\sin x)$
37. $y = \sin x\,(1 + \cos x)$
38. $y = \sqrt[3]{\dfrac{\tan x}{x}}$
39. $y = \tan\left(\sqrt{\cos x}\right)$

Calculate $\frac{dy}{dx}$:

40. $y = \cot x$
41. $y = \sec x$
42. $y = \operatorname{cosec} x$
43. $y = \sec^2 4x$
44. $y = \tan x°$
45. $y = \sin 3x°$
46. $y = \sin bx$
47. $y = \cos(bx + c)$
48. $y = \tan(ax^2 + c)$

Find the derivative of:

49. $y = e^x \sin x$
50. $m = e^{\sin w}$
51. $j = e^{-a} \cos a$
52. $p = e^{4r} \sin 2r$
53. $s = \sin(e^t + t)$
54. $b = \cos(\ln s)$
55. $k = \ln(\cos q)$
56. $c = \ln(\sin 2n)$
57. $f = \ln\left(\tan h^2\right)$
58. $l = \sin(1 - \ln v)$
59. $z = \dfrac{e^{4u}}{\tan u}$
60. $q = \dfrac{\sin m^2}{e^m}$

Find the equation of the tangent to the following curves at the point indicated):

61. $y = x\cos x$ $(0, 0)$
62. $y = \sin x \tan x$ $\left(\frac{\pi}{6}, \frac{1}{2\sqrt{3}}\right)$
63. $y = \sec x$ $\left(\frac{\pi}{4}, \sqrt{2}\right)$
64. $y = \sin x$ $(0,0)$
65. $y = x + \tan x$ $(0, 0)$
66. $y = \csc x$ $\left(\frac{\pi}{3}, \frac{2}{\sqrt{3}}\right)$

Find the (relative) minimum value(s) of each function (and their locations) in the domain $0 \le x \le 2\pi$.

67. $y = x + 2\sin x$
68. $y = \sqrt{3}\sin x + \cos x$
69. $y = 2\sec x + \tan x$
70. $y = \sin^2 x$
71. $y = \tan x - 2x$
72. $y = \cos x - \sqrt{3}\sin x$

Trigonometric Functions (Answers)

The derivatives are:

1. $y' = 2x\sec^2 x^2$
2. $y' = 2\tan x\sec^2 x$
3. $y' = 3\cos 3x$
4. $y' = -2\sin x$
5. $y' = \cos^2 x - \sin^2 x$
6. $y' = \cos x - 5\sin x$
7. $y' = 2\sec^2 2x$
8. $y' = \sin x + x\cos x$
9. $y' = 2\cos 2x\tan 3x + 3\sin 2x\sec^2 3x$
10. $y' = \dfrac{x\cos x - \sin x}{x^2}$
11. $y' = \dfrac{\cos x + x\sin x}{\cos^2 x}$
12. $y' = \dfrac{x\cos x - 2\sin x}{x^3}$
13. $y' = \cos\left(x + \dfrac{\pi}{4}\right)$
14. $y' = 2x + \dfrac{1}{2}\sec^2\dfrac{x}{2}$
15. $y' = -\dfrac{\pi}{3}\sin\dfrac{\pi x}{3}$
16. $y' = -\dfrac{1}{1+\sin x}$
17. $y' = \dfrac{3\sin 5x - 5(3x+4)\cos 5x}{\sin^2 5x}$
18. $y' = 2x\cos x - x^2\sin x$
19. $y' = 3x^2\cos x^3$
20. $y' = -3\sin x\cos^2 x$
21. $y' = 3x^2\tan 8x + 8x^3\sec^2 8x$
22. $y' = 8\sin 4x\cos 4x$
23. $y' = -2x\sin(x^2+1)$
24. $y' = \pi\sec^2 \pi x$
25. $y' = -\dfrac{\sin\sqrt{x}}{2\sqrt{x}}$
26. $y' = -\dfrac{\sin 2x}{\sqrt{\cos 2x}}$
27. $y' = \dfrac{2\sin^3\sqrt{x}\cos\sqrt{x}}{\sqrt{x}}$
28. $y' = 3\tan^2 x\sec^2 x$
29. $y' = 0$
30. $y' = 2(\cos^2 x - \sin^2 x)$
31. $y' = 2x\sec^2(x^2-1)$
32. $y' = -\sec^2(\pi - x)$
33. $y' = 9(2x+\tan 7x)^8 . (2 + 7\sec^2 7x)$
34. $y' = 14x\sec^2(x^2+5)$
35. $y' = -\sin x . \cos(\cos x)$
36. $y' = -\cos x . \sin(\sin x)$
37. $y' = \cos x + \cos^2 x - \sin^2 x$
38. $y' = \dfrac{x\sec^2 x - \tan x}{3x^{\frac{4}{3}}\tan^{\frac{2}{3}} x}$
39. $y' = -\dfrac{\sin x\sec^2(\sqrt{\cos x})}{2\sqrt{\cos x}}$
40. $\dfrac{dy}{dx} = -\csc^2 x$
41. $\dfrac{dy}{dx} = \tan x\sec x$
42. $\dfrac{dy}{dx} = -\cot x\,\mathrm{cosec}\,x$
43. $\dfrac{dy}{dx} = 8\tan 4x\sec^2 4x$
44. $\dfrac{dy}{dx} = \dfrac{\pi}{180}\sec^2 x°$
45. $\dfrac{dy}{dx} = \dfrac{\pi}{180}\cos 3x°$
46. $\dfrac{dy}{dx} = b\cos bx$
47. $\dfrac{dy}{dx} = -b\sin(bx+c)$
48. $\dfrac{dy}{dx} = 2ax\sec^2(ax^2+c)$
49. $\dfrac{dy}{dx} = e^x(\sin x + \cos x)$
50. $\dfrac{dm}{dw} = \cos w\, e^{\sin w}$
51. $\dfrac{dj}{da} = -\dfrac{\sin a + \cos a}{e^a}$
52. $\dfrac{dp}{dr} = 4e^{4r}\sin 2r + 2e^{4r}\cos 2r$
53. $\dfrac{ds}{dt} = (e^t+1)\cos(e^t+t)$
54. $\dfrac{db}{ds} = -\dfrac{\sin(\ln s)}{s}$
55. $\dfrac{dk}{dq} = -\tan q$
56. $\dfrac{dc}{dn} = 2\cot 2n$
57. $\dfrac{df}{dh} = \dfrac{2h}{\cos h^2\sin h^2}$
58. $\dfrac{dl}{dv} = -\dfrac{\cos(1-\ln v)}{v}$
59. $\dfrac{dz}{du} = e^{4u}(4\cot u - \csc^2 u)$
60. $\dfrac{dq}{dm} = \dfrac{2m\cos m^2 - \sin m^2}{e^m}$

The equations of the tangents (in gradient-intercept form) are:

61. $y = x$
62. $y = \dfrac{7x}{6} + \dfrac{1}{2\sqrt{3}} - \dfrac{7\pi}{36}$
63. $y = \sqrt{2}x + \sqrt{2} - \dfrac{\sqrt{2}\pi}{4}$
64. $y = x$
65. $y = 2x$
66. $y = -\dfrac{2x}{3} + \dfrac{2}{\sqrt{3}} + \dfrac{2\pi}{9}$

The (relative) minimum values of the functions are:

67. $y = \dfrac{4\pi}{3} - \sqrt{3}\ \ (at\ x = \dfrac{4\pi}{3})$
68. $y = -2\ \ (at\ x = \dfrac{4\pi}{3})$
69. $y = \sqrt{3}\ \ (at\ x = \dfrac{11\pi}{6})$
70. $y = 0\ \ (at\ x = 0, \pi, 2\pi)$
71. $y = 1 - \dfrac{\pi}{2}\ \ (at\ x = \dfrac{\pi}{4})\ and$ $y = 1 - \dfrac{5\pi}{2}\ \ (at\ x = \dfrac{5\pi}{4})$
72. $y = -2\ \ (at\ x = \dfrac{2\pi}{3})$

$\boxed{\textbf{Partial derivatives}}$

Notice: this material must not be used as a substitute for attending the lectures

0.1 Recall: ordinary derivatives

If y is a function of x then $\frac{dy}{dx}$ is the **derivative** meaning the gradient (slope of the graph) or the rate of change with respect to x.

0.2 Functions of 2 or more variables

Functions which have more than one variable arise very commonly. Simple examples are

- formula for the area of a triangle $A = \frac{1}{2}bh$ is a function of the two variables, base b and height h
- formula for electrical resistors in parallel:

$$R = \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right)^{-1}$$

  is a function of three variables R_1, R_2 and R_3, the resistances of the individual resistors.

Let's talk about functions of two variables here. You should be used to the notation $y = f(x)$ for a function of one variable, and that the graph of $y = f(x)$ is a curve. For functions of two variables the notation simply becomes

$$z = f(x, y)$$

where the two **independent** variables are x and y, while z is the **dependent** variable. The graph of something like $z = f(x, y)$ is a **surface** in three-dimensional space. Such graphs are usually quite difficult to draw by hand.
Since $z = f(x, y)$ is a function of two variables, if we want to differentiate we have to decide whether we are differentiating with respect to x or with respect to y (the answers are different). A special notation is used. We use the symbol ∂ instead of d and introduce the **partial derivatives** of z, which are:

- $\dfrac{\partial z}{\partial x}$ is read as "partial derivative of z (or f) with respect to x", and means differentiate with respect to x holding y constant
- $\dfrac{\partial z}{\partial y}$ means differentiate with respect to y holding x constant

Another common notation is the subscript notation:

$$z_x \quad \text{means} \quad \frac{\partial z}{\partial x}$$

$$z_y \quad \text{means} \quad \frac{\partial z}{\partial y}$$

Note that we cannot use the dash $'$ symbol for partial differentiation because it would not be clear what we are differentiating with respect to.

0.3 Example

Calculate $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ when $z = x^2 + 3xy + y - 1$.

Solution. To find $\dfrac{\partial z}{\partial x}$ treat y as a constant and differentiate with respect to x. We have $z = x^2 + 3xy + y - 1$ so

$$\frac{\partial z}{\partial x} = 2x + 3y$$

Similarly

$$\frac{\partial z}{\partial y} = 3x + 1$$

0.4 Example

Calculate $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ when $z = 1 - x - \frac{1}{2}y$. Interpret your answers and draw the graph.

Solution. The graph of $z = 1 - x - \frac{1}{2}y$ is a plane passing through the points $(x, y, z) = (1, 0, 0)$, $(0, 2, 0)$ and $(0, 0, 1)$. The partial derivatives are:

$$\frac{\partial z}{\partial x} = -1, \qquad \frac{\partial z}{\partial y} = -\tfrac{1}{2}$$

Interpretation: $\dfrac{\partial z}{\partial x}$ is the slope you will notice if you walk on the surface in a direction keeping your y coordinate fixed. $\dfrac{\partial z}{\partial y}$ is the slope you will notice if you walk on the surface in such a direction that your x coordinate remains the same. There are, of course, many other directions you could walk, and the slope you will notice when walking in some other direction can be worked out knowing both $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$. It's like when you walk on a mountain, there are many directions you could walk and each one will have its own slope.

0.5 Other examples of evaluating partial derivatives

(i) $z = \ln(x^2 - y)$. Then $\dfrac{\partial z}{\partial x} = \dfrac{2x}{x^2 - y}$ and $\dfrac{\partial z}{\partial y} = \dfrac{-1}{x^2 - y}$. [To deduce these results we used the fact that if $y = \ln f(x)$ then $\dfrac{dy}{dx} = \dfrac{f'(x)}{f(x)}$].

(ii) $z = x\cos y + ye^x$. Then $\dfrac{\partial z}{\partial x} = \cos y + ye^x$ and $\dfrac{\partial z}{\partial y} = -x\sin y + e^x$.

(iii) $z = y\sin xy$. Then $\dfrac{\partial z}{\partial x} = y(y\cos xy) = y^2\cos xy$ and $\dfrac{\partial z}{\partial y} = yx\cos xy + \sin xy$. For the second result we used the product rule.

(iv) If $x^2 + y^2 + z^2 = 1$ find the rate at which z is changing with respect to y at the point $(\frac{2}{3}, \frac{1}{3}, \frac{2}{3})$. *Solution.* We have $z = (1 - x^2 - y^2)^{1/2}$. We want $\dfrac{\partial z}{\partial y}$ when

$(x, y) = (\frac{2}{3}, \frac{1}{3})$. But

$$\frac{\partial z}{\partial y} = \tfrac{1}{2}(1 - x^2 - y^2)^{-1/2}(-2y) = -\frac{y}{(1 - x^2 - y^2)^{1/2}}$$

Putting in $(x, y) = (\frac{2}{3}, \frac{1}{3})$ gives

$$\frac{\partial z}{\partial y} = -\frac{1/3}{(1 - (2/3)^2 - (1/3)^2)^{1/2}} = -\tfrac{1}{2}.$$

0.6 Functions of 3 or more variables

The general notation would be something like

$$w = f(x, y, z)$$

where x, y and z are the independent variables. For example, $w = x \sin(y + 3z)$. Partial derivatives are computed similarly to the two variable case. For example, $\partial w/\partial x$ means differentiate with respect to x holding both y and z constant and so, for this example, $\partial w/\partial x = \sin(y + 3z)$. Note that a function of three variables does not have a graph.

0.7 Second order partial derivatives

Again, let $z = f(x, y)$ be a function of x and y.

- $\frac{\partial^2 z}{\partial x^2}$ means the second derivative with respect to x holding y constant
- $\frac{\partial^2 z}{\partial y^2}$ means the second derivative with respect to y holding x constant
- $\frac{\partial^2 z}{\partial x \partial y}$ means differentiate first with respect to y and then with respect to x.

The "mixed" partial derivative $\frac{\partial^2 z}{\partial x \partial y}$ is as important in applications as the others. It is a general result that

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$$

i.e. you get the same answer whichever order the differentiation is done.

0.8 Example

Let $z = 4x^2 - 8xy^4 + 7y^5 - 3$. Find all the first and second order partial derivatives of z.

Solution.

$$\begin{aligned}
\frac{\partial z}{\partial x} &= 8x - 8y^4 \\
\frac{\partial z}{\partial y} &= -8x(4y^3) + 35y^4 = -32xy^3 + 35y^4 \\
\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = 8 \\
\frac{\partial^2 z}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) \\
&= \frac{\partial}{\partial y}(-32xy^3 + 35y^4) = -32x(3y^2) + 140y^3 \\
&= -96xy^2 + 140y^3 \\
\frac{\partial^2 z}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial x}(-32xy^3 + 35y^4) = -32y^3 \\
\frac{\partial^2 z}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}(8x - 8y^4) = -32y^3
\end{aligned}$$

0.9 Example

Find all the first and second order partial derivatives of the function $z = \sin xy$.
Solution.

$$\begin{aligned}
\frac{\partial z}{\partial x} &= y \cos xy \\
\frac{\partial z}{\partial y} &= x \cos xy \\
\frac{\partial^2 z}{\partial x^2} &= -y^2 \sin xy \\
\frac{\partial^2 z}{\partial y^2} &= -x^2 \sin xy \\
\frac{\partial^2 z}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial x}(x \cos xy) = x(-y \sin xy) + \cos xy = -xy \sin xy + \cos xy \\
\frac{\partial^2 z}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}(y \cos xy) = y(-x \sin xy) + \cos xy = -xy \sin xy + \cos xy
\end{aligned}$$

0.10 Subscript notation for second order partial derivatives

If $z = f(x, y)$ then

- z_{xx} means $\frac{\partial^2 z}{\partial x^2}$
- z_{yy} means $\frac{\partial^2 z}{\partial y^2}$

- z_{xy} means $\dfrac{\partial^2 z}{\partial x \partial y}$ or $\dfrac{\partial^2 z}{\partial y \partial x}$

0.11 Important point

Unlike ordinary derivatives, partial derivatives do **not** behave like fractions, in particular

$$\frac{\partial x}{\partial z} \neq \frac{1}{\partial z/\partial x}$$

0.12 Small changes

Let

$$z = f(x, y)$$

Imagine we change x to $x + \delta x$ and y to $y + \delta y$ with δx and δy very small. We ask: what is the corresponding change in z? The answer is that the change is δz, given by

$$\delta z \approx \frac{\partial z}{\partial x}\,\delta x + \frac{\partial z}{\partial y}\,\delta y \tag{0.1}$$

This formula requires δx and δy to be very small and even then the formula is only an approximate one. However, it becomes more and more exact as $\delta x \to 0$ and $\delta y \to 0$. This fact is sometimes expressed by saying

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy$$

where dx, dy and dz are infinitesimal increments.
Let's give some idea where formula (0.1) comes from. Let's recall the analogous result for a function of one variable and its derivation. For a function of one variable the notation would be $y = g(x)$ and the graph of this is a curve with a gradient dy/dx at each point x. If consider two points on this curve, (x, y) and a neighbouring point $(x + \delta x, y + \delta y)$ then if this neighbouring point is sufficiently close the line joining the two points, which has gradient $\delta y/\delta x$, is a good approximation to the tangent line at (x, y) which has gradient dy/dx. This means that $\delta y/\delta x \approx dy/dx$ so that $\delta y \approx (dy/dx)\delta x$.
We want to generalise this idea to a function $z = f(x, y)$ of two variables, whose graph will be a surface.
In the (x, y) plane let A be the point with coordinates (x, y), let B be the point with coordinates $(x + \delta x, y)$, and C the point with coordinates $(x + \delta x, y + \delta y)$.
The overall change in height, δz, from A to C is given by

$$\delta z = (\text{change in height } A \text{ to } B) + (\text{change in height } B \text{ to } C)$$

In calculating the change in height from A to B we are travelling across the surface from A to B along a curve in which y is held fixed, so by the result for curves,

$$\text{change in height } A \text{ to } B \approx \frac{\partial z}{\partial x}\,\delta x$$

Similarly

$$\text{change in height } B \text{ to } C \approx \frac{\partial z}{\partial y}\,\delta y$$

Therefore

$$\delta z \approx \frac{\partial z}{\partial x}\,\delta x + \frac{\partial z}{\partial y}\,\delta y$$

and we have derived formula (0.1).

0.13 Example

A cylindrical tank is 1 m high and 0.3 m radius. If height is increased by 5 cm and radius by 1 cm what is the effect on volume?
Solution. Let the radius be r and height be h. Then the volume V is given by

$$V = \pi r^2 h$$

so that $\frac{\partial V}{\partial r} = 2\pi r h$ and $\frac{\partial V}{\partial h} = \pi r^2$. Therefore in the notation of the present problem formula (0.1) becomes

$$\begin{aligned}\delta V &\approx \frac{\partial V}{\partial r}\,\delta r + \frac{\partial V}{\partial h}\,\delta h \\ &= 2\pi r h\,\delta r + \pi r^2 h\,\delta h\end{aligned}$$

In our case $r = 0.3$, $h = 1$, $\delta r = 1$ cm $= 0.01$ m, $\delta h = 5$ cm $= 0.05$ m so

$$\delta V \approx 2\pi(0.3)(1)(0.01) + \pi(0.3)^2(0.05) = 0.033 \ \ \mathrm{m}^3$$

0.14 Example

The angle of elevation of the top of a tower is found to be $30^o \pm 0.5^o$ from a point 300 ± 0.1 m from the base. Estimate the towers height.
Solution. One could imagine that this sort of problem would arise when a surveyor is unable to take completely accurate readings and wants to know the likely margin of error.
Let θ be the angle of elevation, h the towers height and x the distance from tower to observer. Then

$$h = x\tan\theta$$

so that $\frac{\partial h}{\partial x} = \tan\theta$ and $\frac{\partial h}{\partial \theta} = x\sec^2\theta$. Therefore

$$\begin{aligned}\delta h &\approx \frac{\partial h}{\partial x}\,\delta x + \frac{\partial h}{\partial \theta}\,\delta\theta \\ &= \tan\theta\,\delta x + x\sec^2\theta\,\delta\theta\end{aligned}$$

Now $\theta = 30^o = \pi/6$ radians and $\delta\theta = 0.5^o = 0.008727$ radians. Also $x = 300$ m and $\delta x = 0.1$ m. Therefore

$$\delta h \approx (\tan\pi/6)(0.1) + 300(\sec^2\pi/6)(0.008727) = 3.55 \text{ m}$$

From $h = x \tan\theta$, we get $h = 173.21$ m. Our conclusion is that the height is 173.21 ± 3.55 m.
NB: If you had not converted degrees to radians your final answer would be wrong.

0.15 Absolute, relative and percentage change

- absolute change is δz
- relative change is $\frac{\delta z}{z}$
- percentage change is $\frac{\delta z}{z} \times 100$

0.16 Example on percentage change

Length and width of a rectangle are measured with errors of at most 3% and 5% respectively. Estimate the maximum percentage error in the area.
Solution. Let x = length, y = width and A = area. Then, of course, $A = xy$. So $\frac{\partial A}{\partial x} = y$ and $\frac{\partial A}{\partial y} = x$. Therefore

$$\begin{aligned} \delta A &\approx \frac{\partial A}{\partial x}\,\delta x + \frac{\partial A}{\partial y}\,\delta y \\ &= y\,\delta x + x\,\delta y \end{aligned}$$

We want percentage change in A, which is relative change multiplied by 100 so let's work out relative change first. This is given by

$$\begin{aligned} \frac{\delta A}{A} &\approx \frac{y\delta x}{A} + \frac{x\delta y}{A} \\ &= \frac{\delta x}{x} + \frac{\delta y}{y} \end{aligned}$$

since $A = xy$. What we are told is that

$$-0.03 \le \frac{\delta x}{x} \le 0.03 \quad \text{and} \quad -0.05 \le \frac{\delta y}{y} \le 0.05$$

What we need to do now is identify the worst case scenario, i.e. the maximum possible value for $\delta A/A$ given the above constraints. This happens when $\delta x/x = 0.03$ and $\delta y/y = 0.05$, giving $\delta A/A = 0.08$. This is relative error, so the (worst) percentage error is 8%.
NB: in some problems the worst case scenario is obtained by setting one of $\delta x/x$ or $\delta y/y$ to be its most negative (rather than most positive) possible value.

0.17 Chain rule for partial derivatives

Recall the chain rule for ordinary derivatives:

$$\text{if } y = f(u) \text{ and } u = g(x) \text{ then } \frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

In the above we call u the **intermediate variable** and x the **independent variable**. For partial derivatives the chain rule is more complicated. It depends on how many intermediate variables and how many independent variables are present. Below three formulae are given which it is hoped indicate the general points. Essentially, every intermediate variable has to have a term corresponding to it in the right hand side of the chain rule formula. For example in the second one below there are three intermediate variables x, y and z and three terms in the RHS.
Formula 3 below illustrates a case when there are 2 intermediate and 2 independent variables.

(1) if $z = f(x, y)$ and x and y are functions of t ($x = x(t)$ and $y = y(t)$) then z is ultimately a function of t only and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$$

(2) if $w = f(x, y, z)$ and $x = x(t)$, $y = y(t)$, $z = z(t)$ then w is ultimately a function of t only and

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}$$

(3) if $z = f(x, y)$ and $x = x(u, v)$, $y = y(u, v)$ then z is a function of u and v and

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}$$
$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}$$

0.18 Example

Let $z = x^2y$, $x = t^2$ and $y = t^3$. Calculate dz/dt by (a) the chain rule, (b) expressing z as a function of t and finding dz/dt directly.
Solution. (a) by the chain rule

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}\\ &= (2xy)(2t) + (x^2)(3t^2)\\ &= 4xyt + 3x^2t^2\\ &= 4t^2t^3t + 3t^4t^2\\ &= 7t^6\end{aligned}$$

(b) $z = x^2y$ and $x = t^2$, $y = t^3$ so $z = t^4t^3 = t^7$. Differentiating gives $dz/dt = 7t^6$.

It might be tempting to say that approach (b) is clearly easier so why bother with the chain rule? But the fact remains that the chain rule is of fundamental importance in many applications of partial derivatives. We shall see below the use of the chain rule in studying rates of change. And the chain rule is also of importance in the derivation of the partial differential equations that govern many physical processes (eg the Navier Stokes equations of fluid dynamics); in such cases you are not simply playing around with trivial functions but dealing with *unknown* functions.

0.19 Example

Let $w = xy + z$ with $x = \cos t$, $y = \sin t$ and $z = t$. Calculate dw/dt.
Solution.

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \\ &= y(-\sin t) + x(\cos t) + (1)(1) \\ &= -\sin^2 t + \cos^2 t + 1\end{aligned}$$

0.20 Example

Let $u = x^2 - 2xy + 2y^3$ with $x = s^2 \ln t$ and $y = 2st^3$. Find $\partial u/\partial s$ and $\partial u/\partial t$.
Solution. This time u is a function of 2 variables x and y, each of which is itself a function of 2 variables s and t.

$$\begin{aligned}\frac{\partial u}{\partial s} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} \\ &= (2x - 2y)(2s \ln t) + (-2x + 6y^2)(2t^3) \\ &= (2s^2 \ln t - 4st^3)(2s \ln t) + (-2s^2 \ln t + 24s^2t^6)(2t^3) \\ \frac{\partial u}{\partial t} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} \\ &= (2x - 2y)\left(\frac{s^2}{t}\right) + (-2x + 6y^2)(6st^2) \\ &= (2s^2 \ln t - 4st^3)\left(\frac{s^2}{t}\right) + (-2s^2 \ln t + 24s^2t^6)(6st^2)\end{aligned}$$

0.21 Rates of change: an application of the chain rule

We will do some applications of the chain rule to rates of change.
Example. What rate is the area of a rectangle changing if its length is 15 m and increasing at 3 ms^{-1} while its width is 6 m and increasing at 2 ms^{-1}.
Solution. Let x be the length, y the width, A the area and $t =$ time. The information given tells us that

$$\frac{dx}{dt} = 3 \text{ ms}^{-1}, \qquad \frac{dy}{dt} = 2 \text{ ms}^{-1}$$

Obviously $A = xy$. We want dA/dt when $x = 15$ and $y = 6$. This is given by the chain rule as follows:

$$\frac{dA}{dt} = \frac{\partial A}{\partial x}\frac{dx}{dt} + \frac{\partial A}{\partial y}\frac{dy}{dt} = y\frac{dx}{dt} + x\frac{dy}{dt} = (6)(3) + (15)(2) = 48\ \mathrm{m}^2\mathrm{s}^{-1}.$$

Example. The height of a tree increases at a rate of 2 ft per year and the radius increases at 0.1 ft per year. What rate is the volume of timber increasing at when the height is 20 ft and the radius is 1.5 ft. (Assume the tree is a circular cylinder).
Solution. The volume V is given by $V = \pi r^2 h$. The chain rule gives

$$\begin{aligned}\frac{dV}{dt} &= \frac{\partial V}{\partial r}\frac{dr}{dt} + \frac{\partial V}{\partial h}\frac{dh}{dt} \\ &= 2\pi r h\,\frac{dr}{dt} + \pi r^2\,\frac{dh}{dt}\end{aligned}$$

We are told that $dh/dt = 2$ ft per year and $dr/dt = 0.1$ ft per year. So, when $h = 20$ and $r = 1.5$,

$$\frac{dV}{dt} = 2\pi(1.5)(20)(0.1) + \pi(1.5)^2(2) = 32.99\ \mathrm{ft}^3/\mathrm{year}$$

0.22 The chain rule and implicit differentiation

Suppose we cannot find y explicitly as a function of x, only implicitly through the equation $F(x, y) = 0$ (for example, $F(x, y)$ might be an awkward expression such that $F(x, y) = 0$ cannot in practice be solved to give y in terms of x). We want a formula for dy/dx.
We know that $F(x, y) = 0$ defines y as a function of x, $y = y(x)$, even if we cannot in practice find the expression for y in terms of x. This means that we could write $F(x, y) = 0$ as $F(x, y(x)) = 0$. Differentiating both sides of this, using the chain rule on the left hand side, gives

$$\frac{\partial F}{\partial x}(1) + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0$$

Hence

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y}$$

As an example of the use of this formula, let us find dy/dx for the function y defined by $x^2 + xy + y^3 - 7 = 0$. Let $F(x, y) = x^2 + xy + y^3 - 7$. Then by the above formula,

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{(2x + y)}{x + 3y^2}$$

Alternatively you could deduce this result by using implicit differentiation (a technique which you should know about from previous study). It should, of course, give the same answer.
As an extension of the above idea, let the equation $f(x, y, z) = 0$ define z as a function of x and y, so that x and y are viewed as independent variables. We want

to find $\partial z/\partial x$ and $\partial z/\partial y$. The calculation here is a somewhat subtle one, in which x actually plays the role of both an intermediate variable and an independent one. Differentiating the equation $f(x, y, z) = 0$ with respect to x using the chain rule gives

$$\frac{\partial f}{\partial x}(1) + \frac{\partial f}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0$$

Now $\partial y/\partial x$ is, in fact, zero. The reason is that y and x are independent of each other. So

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0$$

Hence

$$\frac{\partial z}{\partial x} = -\frac{\partial f/\partial x}{\partial f/\partial z}$$

and similarly

$$\frac{\partial z}{\partial y} = -\frac{\partial f/\partial y}{\partial f/\partial z}$$

0.23 Transforming to polars

Let $u = u(x, y)$ be a function of x and y. Let

$$x = r\cos\theta, \qquad y = r\sin\theta$$

Our aim is to show that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} \tag{0.2}$$

which is the expression for the Laplacian operator in plane polar coordinates. It is useful for solving, for example, the steady state heat equation in situations with circular geometry.

By the chain rule,

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r}$$

i.e.

$$\frac{\partial u}{\partial r} = \cos\theta\frac{\partial u}{\partial x} + \sin\theta\frac{\partial u}{\partial y}$$

Differentiating the above expression with respect to r gives

$$\begin{aligned}
\frac{\partial^2 u}{\partial r^2} &= \cos\theta\frac{\partial}{\partial r}\left(\frac{\partial u}{\partial x}\right) + \sin\theta\frac{\partial}{\partial r}\left(\frac{\partial u}{\partial y}\right) \\
&= \cos\theta\left(\frac{\partial^2 u}{\partial x^2}\frac{\partial x}{\partial r} + \frac{\partial^2 u}{\partial x\partial y}\frac{\partial y}{\partial r}\right) + \sin\theta\left(\frac{\partial^2 u}{\partial x\partial y}\frac{\partial x}{\partial r} + \frac{\partial^2 u}{\partial y^2}\frac{\partial y}{\partial r}\right) \\
&= \cos^2\theta\frac{\partial^2 u}{\partial x^2} + \sin\theta\cos\theta\frac{\partial^2 u}{\partial x\partial y} + \sin\theta\cos\theta\frac{\partial^2 u}{\partial x\partial y} + \sin^2\theta\frac{\partial^2 u}{\partial y^2}.
\end{aligned}$$

Also

$$
\begin{aligned}
\frac{\partial u}{\partial \theta} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \theta} \\
&= -r\sin\theta\frac{\partial u}{\partial x} + r\cos\theta\frac{\partial u}{\partial y}
\end{aligned}
$$

and, after a long calculation,

$$
\begin{gathered}
\frac{\partial^2 u}{\partial \theta^2} = r^2\sin^2\theta\frac{\partial^2 u}{\partial x^2} + r^2\cos^2\theta\frac{\partial^2 u}{\partial y^2} - 2r^2\sin\theta\cos\theta\frac{\partial^2 u}{\partial x\partial y} \\
-r\cos\theta\frac{\partial u}{\partial x} - r\sin\theta\frac{\partial u}{\partial y}
\end{gathered}
$$

It follows that

$$
\begin{gathered}
\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \tfrac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = \cos^2\theta\frac{\partial^2 u}{\partial x^2} + 2\sin\theta\cos\theta\frac{\partial^2 u}{\partial x\partial y} + \sin^2\theta\frac{\partial^2 u}{\partial y^2} \\
+\frac{1}{r}\left(\cos\theta\frac{\partial u}{\partial x} + \sin\theta\frac{\partial u}{\partial y}\right) \\
+\frac{1}{r^2}\left(r^2\sin^2\theta\frac{\partial^2 u}{\partial x^2} + r^2\cos^2\theta\frac{\partial^2 u}{\partial y^2} - 2r^2\sin\theta\cos\theta\frac{\partial^2 u}{\partial x\partial y} - r\cos\theta\frac{\partial u}{\partial x} - r\sin\theta\frac{\partial u}{\partial y}\right) \\
= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}
\end{gathered}
$$

so that (0.2) is proved.

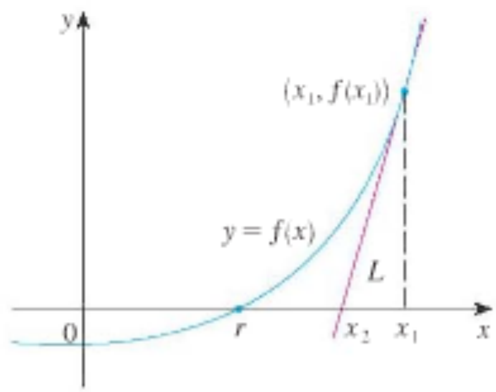
Newton's Method

In this section we will explore a method for estimating the solutions of an equation $f(x) = 0$ by a sequence of approximations that approach the solution.

Note that for a quadratic equation $ax^2+bx+c = 0$, we can solve for the solutions using the quadratic formula. There are formulas available to find the zeros of cubic and quartic polynomials, however they are more difficult (and tedious) to apply than the quadratic formula. Some notes on the history of the solutions have been included at the end of the lecture. It can be shown that the derivation of such a formula is impossible for polynomials with degree 5 or higher and there is no such systematic method to find their solutions. The methods below apply to all polynomial functions and a wide range of other functions.

Procedure for Newton's Method

To estimate the solution of an equation $f(x) = 0$, we produce a sequence of approximations that approach the solution. We find the first estimate by sketching a graph or by guessing. We choose x_1 which is close to the solution according to our estimate. The method then uses the tangent line to the curve at the point $(x_1, f(x_1))$ as an estimate of the path of the curve.



We label the point where this tangent to the graph of $f(x)$ cuts the x-axis, x_2. Usually x_2 is a better approximation to the solution of $f(x) = 0$ than x_1. We can find a formula for x_2 in terms of x_1 by using the equation of the tangent at $(x_1, f(x_1))$:

$$y = f(x_1) + f'(x_1)(x - x_1).$$

The x-intercept of this line, x_2, occurs when $y = 0$ or

$$f(x_1) + f'(x_1)(x_2 - x_1) = 0 \quad \text{or} \quad -f(x_1) = f'(x_1)(x_2 - x_1)$$

This implies that

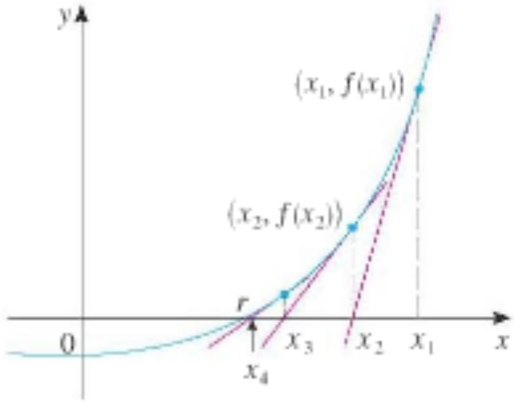
$$-\frac{f(x_1)}{f'(x_1)} = x_2 - x_1 \quad \text{or} \quad \boxed{x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}}.$$

If $f(x_2) \neq 0$, we can repeat the process to get a third approximation to the zero as

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$$

In the same way we can repeat the process to get a third and a fourth approximation to the zero using the formula:

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.}$$



Note When applying Newton's method it may happen that the sequence of approximations gets further and further away from the zero of the function (or further apart from each other) as the value of n increases. It may also happen that one of the approximations falls outside the domain of the function. In this case you should start over with a different approximation.

Example 1 Use Newton's method to find the fourth approximation, x_4, to the root of the following equation

$$x^3 - x - 1 = 0$$

starting with $x_1 = 1$.

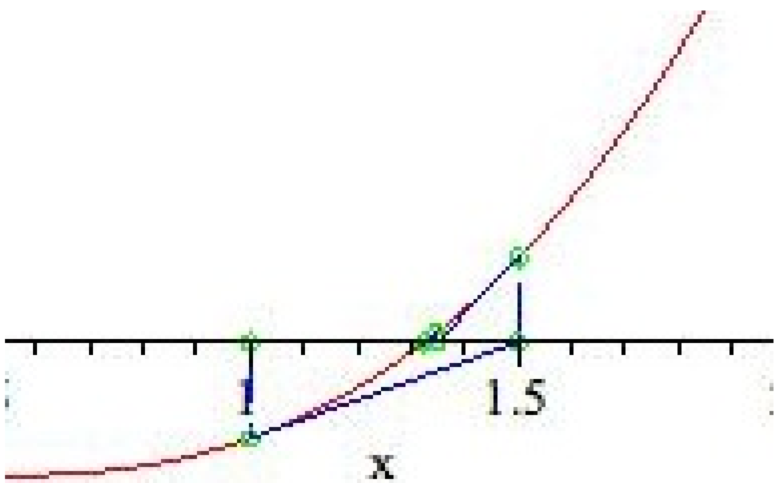
Note that if $f(x) = x^3 - x - 1$, then $f(1) = -1 < 0$ and $f(2) = 5 > 0$. Therefore by the Intermediate Value Theorem, there is a root between $x = 1$ and $x = 2$.

It is helpful to make a table for computation:

		$f(x) =$	$f'(x) =$	
n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
1	1			
2				
3				

As a **rule of thumb**, when successive approximations x_n and x_{n+1} agree up to K decimal places, then our approximation is accurate up to K decimal places. For example the table below shows the approximations $x_1, x_2, \ldots, x_6$ for the above problem. We see that x_5 and x_6 agree up to 9 decimal places. By the rule of thumb above, the solution to the equation is equal to 1.324717957 when rounded off to 9 decimal places.

		$f(x) = x^3 - x - 1$	$f'(x) = 3x^2 - 1$	
n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
1	1	-1	2	1.5
2	1.5	0.875	5.75	1.347826087
3	1.347826087	0.100682173	4.449905482	1.325200399
4	1.325200399	0.002058362	4.268468292	1.324718174
5	1.324718174	0.000000924	4.264634722	1.324717957
6	1.324717957	-1.8672×10^{-13}	4.264632999	1.324717957



Example 2 Use Newton's method to estimate $\sqrt{2}$ correct up to 5 decimal places.

Find an equation $f(x) = 0$ for which $\sqrt{2}$ is a solution.

		$f(x) =$	$f'(x) =$	
n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
1				
2				
3				
4				
5				

Example 3 Do the curves $y = \cos x$ and $y = 3x$ meet?

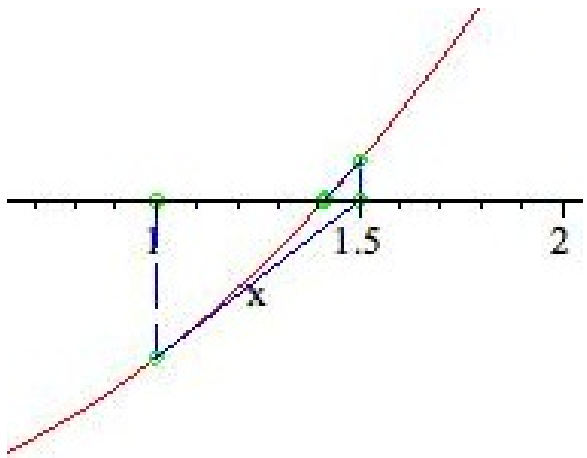
If so estimate the value of x at which they meet up to 3 decimal places.

		$f(x) =$	$f'(x) =$	
n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
1				
2				
3				
4				

We stop when successive approximations are equal up to 5 decimal places.

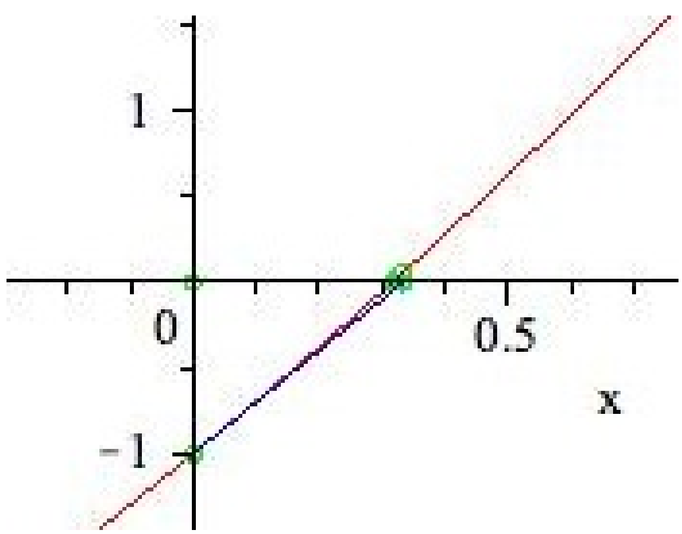
Example 2 Solution From calculator ($\sqrt{2} = 1.41421$). To use Newton's method, we look for a positive root of $f(x) = x^2 - 2$. We stop when successive approximations are equal up to 5 decimal places.

		$f(x) = x^2 - 2$	$f'(x) = 2x$	
n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
1	1	-1	2	1.5
2	1.5	.25	3	1.416666666
3	1.416666666	.006944444	2.833333333	1.414215686
4	1.414215686	.000006005	2.828431372	1.414213563



Example 3 Solution To use Newton's method, $f(x) = 3x - \cos x$. We use $x_1 = 0$ stop when successive approximations are equal up to 3 decimal places.

		$f(x) = 3x - \cos x$	$f'(x) = 3 + \sin x$	
n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
1	0	-1	3	$1/3 = .333333333333$
2	$1/3$	.05504305	3.3271946	.316789952
3	.316789952	.0001295555	3.31517854	.316750829



Niccolò Fontana Tartaglia

From Wikipedia, the free encyclopedia


Niccolò Fontana Tartaglia.

Niccolò Fontana Tartaglia (1499/1500, Brescia, Italy – 13 December 1557, Venice, Italy) was a mathematician, an engineer (designing fortifications), a surveyor (of topography, seeking the best means of defense or offense) and a bookkeeper from the then-Republic of Venice (now part of Italy). He published many books, including the first Italian translations of Archimedes and Euclid, and an acclaimed compilation of mathematics. Tartaglia was the first to apply mathematics to the investigation of the paths of cannonballs; his work was later validated by Galileo's studies on falling bodies.

Niccolò Fontana was the son of Michele Fontana, a rider and deliverer. In 1505, Michele was murdered and Niccolò, his two siblings, and his mother were impoverished. Niccolò experienced further tragedy in 1512 when the French invaded Brescia during the War of the League of Cambrai. The militia of Brescia defended their city for seven days. When the French finally broke through, they took their revenge by massacring the inhabitants of Brescia. By the end of battle, over 45,000 residents were killed. During the massacre, a French soldier sliced Niccolò's jaw and palate. This made it impossible for Niccolò to speak normally, prompting the nickname "Tartaglia" (stammerer).

There is a story that Tartaglia learned only half the alphabet from a private tutor before funds ran out, and he had to learn the rest for himself. Be that as it may, he was essentially self-taught. He and his contemporaries, working outside the academies, were responsible for the spread of classic works in modern languages among the educated middle class.

His edition of Euclid in 1543, the first translation of the *Elements* into any modern European language, was especially significant. For two centuries Euclid had been taught from two Latin translations taken from an Arabic source; these contained errors in Book V, the Eudoxian theory of proportion, which rendered it unusable. Tartaglia's edition was based on Zamberti's Latin translation of an uncorrupted Greek text, and rendered Book V correctly. He also wrote the first modern and useful commentary on the theory. Later, the theory was an essential tool for Galileo, just as it had been for Archimedes.

Solution to cubic equations

Tartaglia is perhaps best known today for his conflicts with Gerolamo Cardano. Cardano nagged Tartaglia into revealing his solution to the cubic equations, by promising not to publish them. Several years later, Cardano happened to see unpublished work by Scipione del Ferro who independently came up with the same solution as Tartaglia. As the unpublished work was dated before Tartaglia's, Cardano decided his promise could be broken and included Tartaglia's solution in his next publication. Since Cardano credited his discovery, Tartaglia was extremely upset. He responded by publicly insulting Cardano.

Gerolamo Cardano

From Wikipedia, the free encyclopedia

"Cardanus" redirects here. For the lunar crater, see Cardanus (crater). For the stag beetle genus, see Cardanus (beetle).

Gerolamo (or **Girolamo**, or **Geronimo**) **Cardano** (French **Jérôme Cardan**; Latin **Hieronymus Cardanus**) (24 September 1501 – 21 September 1576) was an Italian Renaissance mathematician, physician, astrologer and gambler.

Girolamo Cardano

Girolamo Cardano

Born	24 September 1501 Pavia
Died	21 September 1576 (aged 74) Rome
Nationality	Italian
Fields	Mathematics Physics
Alma mater	University of Pavia
Known for	Algebra

Contents

Early life and education

He was born in Pavia, Lombardy, the illegitimate child of Fazio Cardano, a mathematically gifted lawyer, who was a friend of Leonardo da Vinci. In his autobiography, Cardano claimed that his mother had attempted to abort him. Shortly before his birth, his mother had to move from Milan to Pavia to escape the plague; her three other children died from the disease.

In 1520, he entered the University of Pavia and later in Padua studied medicine. His eccentric and confrontational style did not earn him many friends and he had a difficult time finding work after his studies had ended. In 1525, Cardano repeatedly applied to the College of Physicians in Milan, but was not admitted owing to his reputation and illegitimate birth.

Eventually, he managed to develop a considerable reputation as a physician and his services were highly valued at the courts. He was the first to describe typhoid fever.

Mathematics

Today, he is best known for his achievements in algebra. He published the solutions to the cubic and quartic equations in his 1545 book *Ars Magna*. The solution to one particular case of the cubic, $x^3 + ax = b$ (in modern notation), was communicated to him by Niccolò Fontana Tartaglia (who later claimed that Cardano had sworn not to reveal it, and engaged Cardano in a decade-long fight), and the quartic was solved by Cardano's student Lodovico Ferrari. Both were acknowledged in the foreword of the book, as well as in several places within its body. In his exposition, he acknowledged the existence of what are now called imaginary numbers, although he did not understand their properties (Mathematical field theory was developed centuries later). In *Opus novum de proportionibus* he introduced the binomial coefficients and the binomial theorem.

Cardano was notoriously short of money and kept himself solvent by being

Portrait of Cardano on display at the School of Mathematics and Statistics, University of St Andrews.

Cardano was notoriously short of money and kept himself solvent by being an accomplished gambler and chess player. His book about games of chance, *Liber de ludo aleae* ("Book on Games of Chance") , written in 1526, but not published until 1663, contains the first systematic treatment of probability, as well as a section on effective cheating methods. Cardano invented several mechanical devices including the combination lock, the gimbal consisting of three concentric rings allowing a supported compass or gyroscope to rotate freely, and the Cardan shaft with universal joints, which allows the transmission of rotary motion at various angles and is used in vehicles to this day. He studied hypocycloids, published in *de proportionibus* 1570. The generating circles of these hypocycloids were later named Cardano circles or cardanic circles and were used for the construction of the first high-speed printing presses.

He made several contributions to hydrodynamics and held that perpetual motion is impossible, except in celestial bodies. He published two encyclopedias of natural science which contain a wide variety of inventions, facts, and occult superstitions. He also introduced the Cardan grille, a cryptographic tool, in 1550.

Someone also assumed to Cardano the credit for the invention of the so called *Cardano's Rings*, also called Chinese Rings, but it is very probable that they are more ancient than Cardano.

Significantly, in the history of education of the deaf, he said that deaf people were capable of using their minds, argued for the importance of teaching them, and was one of the first to state that deaf people could learn to read and write without learning how to speak first. He was familiar with a report by Rudolph Agricola about a deaf mute who had learned to write.

Family

Cardano's eldest and favorite son was executed in 1560 after he confessed to having poisoned his cuckolding wife. His other son was a gambler, who stole money from him. He allegedly cropped the ears of one of his sons. Cardano himself was accused of heresy in 1570 because he had computed and published the horoscope of Jesus in 1554. Apparently, his own son contributed to the prosecution, bribed by Tartaglia. He was arrested, had to spend several months in prison and was forced to abjure his professorship. He moved to Rome, received a lifetime annuity from Pope Gregory XIII (after first having been rejected by Pope Pius V) and finished his autobiography.

43. Vector Fields

A vector field is a function **F** that assigns to each ordered pair (x, y) in R^2 a vector of the form $\langle M(x, y), N(x, y)\rangle$. We write

$$\mathbf{F}(x, y) = \langle M(x, y), N(x, y)\rangle.$$

This can be extended into higher dimensions. For example. In R^3, we would write

$$\mathbf{F}(x, y, z) = \langle M(x, y, z), N(x, y, z), P(x, y, z)\rangle.$$

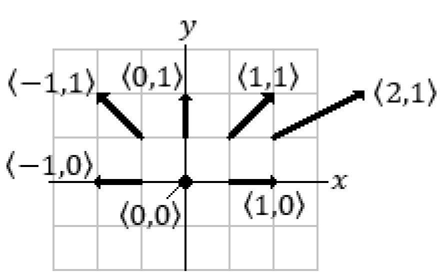
♦ • ♦ • ♦ • ♦ • ♦

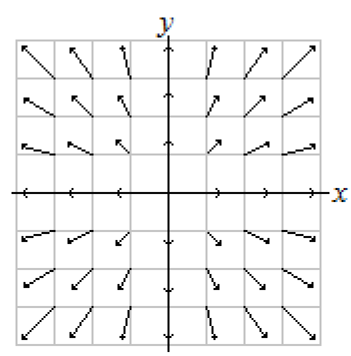
Example 43.1: Sketch $\mathbf{F}(x, y) = \langle x, y\rangle$.

Solution: Using an input-output table, we can show some of the vectors in the vector field **F**:

Ordered pair (x, y)	Vector $\langle x, y\rangle$		Ordered pair (x, y)	Vector $\langle x, y\rangle$
$(0,0)$	$\langle 0,0\rangle$		$(-1,0)$	$\langle -1,0\rangle$
$(1,0)$	$\langle 1,0\rangle$		$(-1,1)$	$\langle -1,1\rangle$
$(1,1)$	$\langle 1,1\rangle$		$(1,-1)$	$\langle 1,-1\rangle$
$(0,1)$	$\langle 0,1\rangle$		$(2,1)$	$\langle 2,1\rangle$
$(1,2)$	$\langle 1,2\rangle$		$(2,2)$	$\langle 2,2\rangle$

The vector $\langle x, y\rangle$ is drawn so that its foot is at the point described by the ordered pair (x, y). Below (left) are a sample of vectors of **F**, and at right, a slightly-more complete rendering of the vector field. In this example, the vectors point radially (along straight lines) away from the origin.





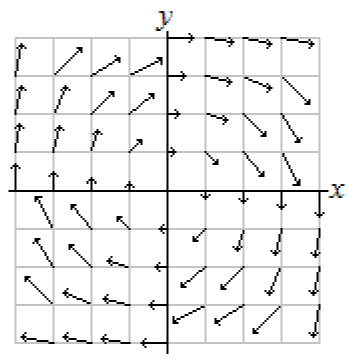
It can be time consuming to sketch a vector field. Also, the vectors themselves “cover up” other vectors, resulting in a cluttered, unreadable image. Certain artistic liberties are allowed. For example, the vectors may be scaled down in size to show relative magnitudes rather than true magnitudes. Often, it is more important to see the “flow” created by the vectors, rather than the actual magnitudes.

♦ • ♦ • ♦ • ♦ • ♦

Example 43.2: Sketch $\mathbf{F}(x, y) = \langle y, -x \rangle$.

Solution: An input-output table shows some of the vectors, followed by an image of the vector field.

Ordered pair (x, y)	Vector $\langle x, y \rangle$		Ordered pair (x, y)	Vector $\langle x, y \rangle$
$(0,0)$	$\langle 0,0 \rangle$		$(-1,0)$	$\langle 0,1 \rangle$
$(1,0)$	$\langle 0,-1 \rangle$		$(-1,1)$	$\langle 1,1 \rangle$
$(1,1)$	$\langle 1,-1 \rangle$		$(1,-1)$	$\langle -1,-1 \rangle$
$(0,1)$	$\langle 1,0 \rangle$		$(2,1)$	$\langle 1,-2 \rangle$
$(1,2)$	$\langle 2,-1 \rangle$		$(2,2)$	$\langle 2,-2 \rangle$

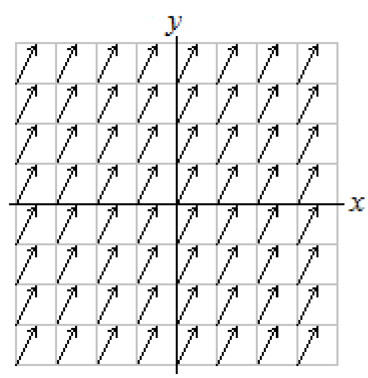


The vectors suggest a clockwise rotation around the origin.

♦ • ♦ • ♦ • ♦ • ♦

Example 43.3: Sketch $\mathbf{F}(x, y) = \langle 1,2 \rangle$.

Solution: This is a constant vector field. All vectors are identical in magnitude and orientation. In the image below, each vector is shown at half-scale so as not to clutter the image too severely.

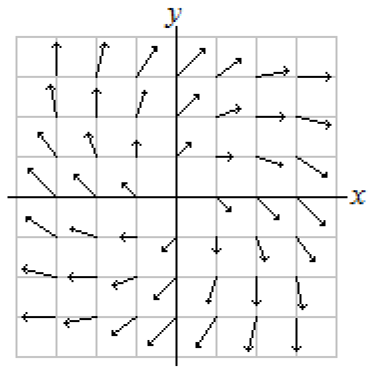


This vector field is not radial nor does it suggest any rotation.

♦ • ♦ • ♦ • ♦ • ♦

Example 43.4: Sketch $\mathbf{F}(x,y) = \langle x+y, y-x \rangle$.

Solution: The vector field is shown below:



This vector field appears to have both radial and rotational aspects in its appearance.

♦ • ♦ • ♦ • ♦ • ♦

Gradient Vector Fields

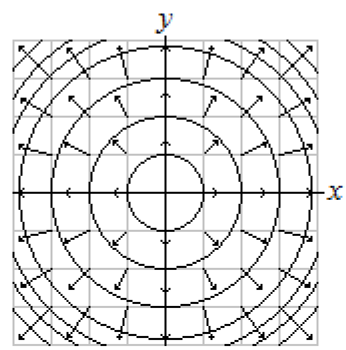
Given a function $z = f(x,y)$, its gradient is $\nabla f = \langle f_x(x,y), f_y(x,y) \rangle$. This is called a **gradient vector field** (or just **gradient field**). It is also called a **conservative vector field** and is discussed in depth in Section 47. In such a case, the vector field is written as $\mathbf{F}(x,y) = \nabla f = \langle f_x, f_y \rangle$.

Gradient vector fields have an interesting visual property: the vectors in the vector field lie orthogonal to the contours of f.

♦ • ♦ • ♦ • ♦ • ♦

Example 43.5: Given $f(x,y) = \frac{1}{2}x^2 + \frac{1}{2}y^2$, find $\mathbf{F}(x,y) = \nabla f$ and sketch it along with the contour map of f.

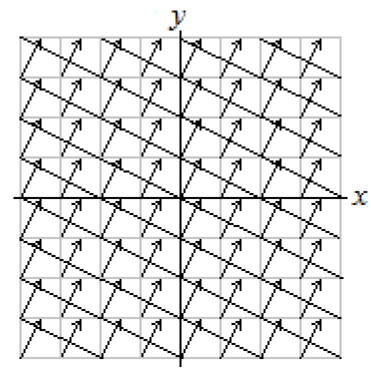
Solution: The vector field is $\mathbf{F}(x,y) = \nabla f = \langle f_x, f_y \rangle = \langle x, y \rangle$. The contours of f are concentric circles of the form $\frac{1}{2}x^2 + \frac{1}{2}y^2 = k$ centered at the origin, the surface being a paraboloid with its vertex at (0,0,0) and opening upward. Note that the vectors in **F** are orthogonal to the contours of f. This is the same vector field as seen in Example 43.1. The vectors point in the direction of increasing *z*.



♦ • ♦ • ♦ • ♦ • ♦

Example 43.6: Given $f(x, y) = x + 2y$, find $\mathbf{F}(x, y) = \nabla f$ and sketch it along with the contour map of f.

Solution: The vector field is $\mathbf{F}(x, y) = \nabla f = \langle f_x, f_y \rangle = \langle 1,2 \rangle$. The surface of f is a plane tilting "upward" as x and y both increase in value. Note that the contours of f are all lines of the form $x + 2y = k$, or $y = -\frac{1}{2}x + \frac{k}{2}$, and that the vectors in $\mathbf{F}$ are orthogonal to the contours of f, pointing in the direction of increasing z. This is the same vector field as in Example 43.3.



♦ • ♦ • ♦ • ♦ • ♦

Not all vector fields are gradient fields. Those in Examples 43.2, and 43.4 are not gradient fields. There do not exist functions $z = f(x, y)$ such that $\mathbf{F}(x, y) = \nabla f$ in these two examples.

If $\mathbf{F}$ is a gradient field, it is possible to find a function f such that $\mathbf{F}(x, y) = \nabla f$. Such a function f is called a **potential function**, and this is discussed in Section 47.

All constant vector fields $\mathbf{F}(x, y) = \langle a, b \rangle$ are gradient fields, where $f(x, y) = ax + by$ is a potential function. In R^3, we would have $\mathbf{F}(x, y, z) = \langle a, b, c \rangle$, with potential function $f(x, y, z) = ax + by + cz$.

All vector fields of the form $\mathbf{F}(x, y) = \langle M(x), N(y) \rangle$ are gradient fields, where the potential function is $f(x, y) = \int M(x)\,dx + \int N(y)\,dy$.

♦ • ♦ • ♦ • ♦ • ♦

Example 43.7: Find the potential functions for $\mathbf{F}(x, y, z) = \langle -1,4,2 \rangle$ and for $\mathbf{G}(x, y) = \langle 2x, y^4 \rangle$.

Solution: For $\mathbf{F}$, a potential function is $f(x, y, z) = -x + 4y + 2z$, and for $\mathbf{G}$, a potential function is $g(x, y) = \int 2x\,dx + \int y^4\,dy = x^2 + \frac{1}{5}y^5$.

Constants of integration are not necessary. If $\mathbf{F}$ is a gradient field, then it has infinitely-many potential functions, all equivalent up to its constant of integration. Note that for $\mathbf{F}$ above, $f(x, y, z) = -x + 4y + 2z + 7$ is also a valid potential function. Usually, we let any such constant be 0.

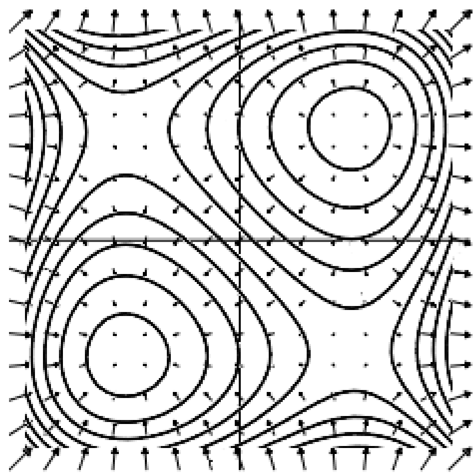
♦ • ♦ • ♦ • ♦ • ♦

Example 43.8: Given $f(x,y) = x^3 + y^3 - 3x - 3y$, find $\mathbf{F}(x,y) = \nabla f$ and sketch it along with the contour map of f.

Solution: The vector field is

$$\mathbf{F}(x,y) = \nabla f = \langle f_x, f_y \rangle = \langle 3x^2 - 3, 3y^2 - 3 \rangle.$$

The vector field **F** is shown below with the contours of f:



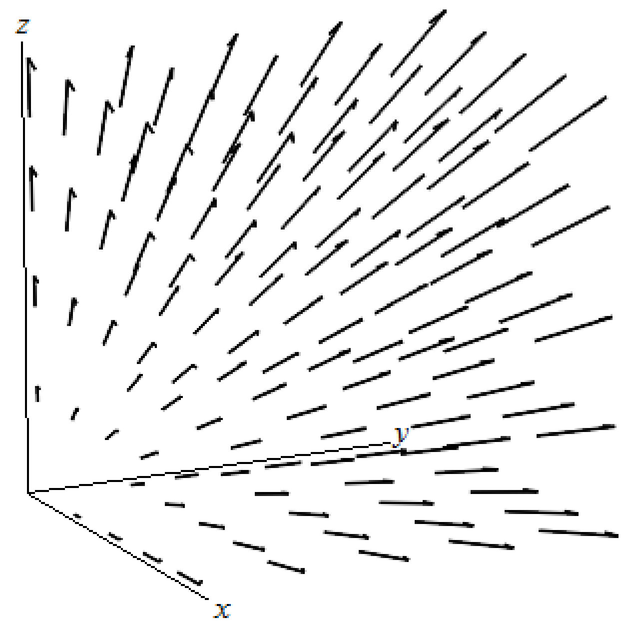
Using techniques of unconstrained optimization (Section 29), there are four critical points. They are: (1,1,–4), a minimum; (–1,–1,4), a maximum; and (1,–1,0) and (–1,1,0), both saddle points. Observe a few things:

- The vectors in **F** always point in the direction of increasing *z*, or "up".
- Note that the vectors point "up" *toward* the maximum at (–1,–1,4) and "up" *away* from the minimum at (1,1,–4).
- At each critical point, $\nabla f = \langle 0,0 \rangle$. In the image above, the vectors near these points have very small magnitudes, while the vectors at these critical points have no magnitude.

♦ • ♦ • ♦ • ♦ • ♦

Example 43.9: Sketch $\mathbf{F}(x, y, z) = \langle x, y, z \rangle$.

Solution: The vector field is sketched below, for $x > 0$, $y > 0$ and $z > 0$:



The vectors in **F** all point radially away from the origin, increasing in magnitude the farther away from the origin.

Sketching a vector field in R^3 is nearly impossible to do manually. A computer program is an essential tool to render such fields.

♦ • ♦ • ♦ • ♦ • ♦

Given a function $w = f(x, y, z)$, then a gradient field in R^3 can be defined by the gradient of f:

$$\mathbf{F}(x, y, z) = \nabla f = \langle f_x, f_y, f_z \rangle.$$

A function such as $w = f(x, y, z)$ exists in R^4 (three independent variables, one dependent variable). Its contours will be surfaces in R^3, and the vectors in the gradient field given by ∇f will be orthogonal to the contours, which in this case are surfaces.

♦ • ♦ • ♦ • ♦ • ♦

Example 43.10: Show that $f(x, y, z) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + \frac{1}{2}z^2$ is a potential function of $\mathbf{F}(x, y, z) = \langle x, y, z \rangle$. Discuss how the vectors in **F** compare to the contours of f.

Solution: The gradient of f is shown below:

$$\nabla f = \langle f_x, f_y, f_z \rangle = \langle x, y, z \rangle.$$

The contours of f are found by setting w equal to various constants. For example, when $w = 1$, then we have

$$1 = \frac{1}{2}x^2 + \frac{1}{2}y^2 + \frac{1}{2}z^2, \qquad \text{so that} \qquad x^2 + y^2 + z^2 = 2.$$

This is a sphere centered at the origin with radius $\sqrt{2}$. In fact, all contours of f are spheres. As w increases in value, its contour at $w = k$ is given by the sphere $x^2 + y^2 + z^2 = 2k$.

For example, the vector whose foot lies at $(1,2,3)$ is given by $\langle 1,2,3 \rangle$. The point $(1,2,3)$ itself lies on a sphere with radius $\sqrt{14}$. The vector $\langle 1,2,3 \rangle$ has its foot on this sphere, oriented orthogonally to this sphere, pointing directly away from the origin (in this case).

♦ • ♦ • ♦ • ♦ • ♦

A Short Note on the Gradient Vector

or, the benefit of multiple perspectives in multivariable calculus

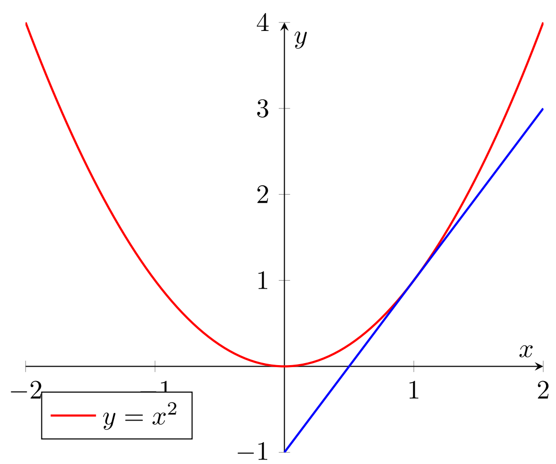
Joseph Breen

The general goal of this note is to describe how to find tangent planes to a surface via the gradient vector. Rather than jumping into an example and show you what to do, I'm going to motivate the perspective with a low-dimensional analogue. An important theme in multivariable calculus is that **every single topic is a direct extension of something familiar in one variable calculus**. As such, it can be insightful to reinterpret ideas like the gradient in a low-dimensional setting.

1 Computing a tangent line

Before computing a tangent plane, I'm going to begin with a problem that may seem a little silly:

Example 1.1. *Consider the curve* $y = x^2$*. Find the equation of the tangent line to this curve at the point* $(1, 1)$.



You should be thinking, objecting, or wondering something along the lines of: *Joe, I already know how to do this; I learned this in single variable calculus. Why are you wasting my time computing a boring tangent line?* The idea is the following: we can certainly compute the tangent line using single variable calculus, but **we can completely change the perspective of the problem and use the gradient vector of a multivariable function**. This will be the low-dimensional analogy to computing tangent planes via the gradient.

1.1 The first perspective: single variable calculus

First, we'll compute the tangent line like any normal person would and just use single variable calculus techniques. No multivariable calculus here! Recall that the equation for a tangent line to f through the point $(a, f(a))$ is given by

$$y = f(a) + f'(a)(x - a).$$

Here $f(x) = x^2$ and $a = 1$. Since $f(1) = 1$ and $f'(1) = 2x\mid_1 = 2$, the equation of the tangent line we seek is

$$y = 1 + 2(x - 1).$$

Done!

1.2 The second perspective: multivariable calculus

This is where things get a little wild, and arguably needlessly complicated. I don't disagree, but the point is for you to see how the ideas of multivariable calculus are being used in low-dimensions so that you can more easily understand the analogous situation in higher dimensions.

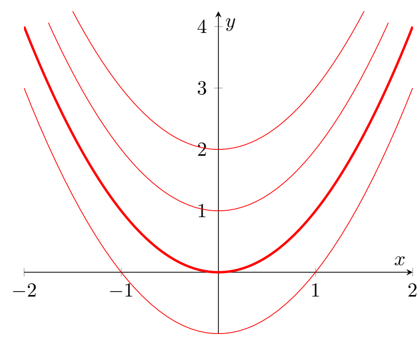
The change in perspective is that instead of viewing the curve $y = x^2$ as the graph of a single variable function $f(x)$, we will view the curve $y = x^2$ **as the level curve of a multivariable function $F(x, y)$.** In particular, let

$$F(x, y) := y - x^2.$$

Then the curve $y = x^2$, which rewritten is the curve $y - x^2 = 0$, is precisely a level curve of F with value 0:

$$F(x, y) = 0 \quad \rightsquigarrow \quad y - x^2 = 0.$$

Here is a contour plot of the function F:



I've plotted the level curves $F(x, y) = c$ for $c = -1, 0, 1, 2$. The thicker red line is the level curve $F(x, y) = 0$, and this is exactly the curve $y = x^2$ that we care about.

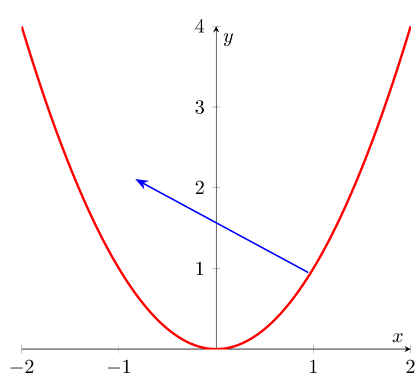
Next, **the main fact we need about the gradient vector is that it is perpendicular to level things**. In particular, the vector

$$\nabla F(x, y) = \langle F_x(x, y), F_y(x, y) \rangle = \langle -2x, 1 \rangle$$

is perpendicular to the level curve through the point (x, y). We care about the point $(1, 1)$, so let's compute the gradient vector at that point:

$$\nabla F(1, 1) = \langle -2, 1 \rangle.$$

Indeed, if I plot this vector on the contour plot above, we get:



Next, how do we use this vector to find the *tangent* line? Recall that to find the equation of a *plane*, you need a normal vector $\langle a, b, c\rangle$ and a point (x_0, y_0, z_0). With this information the plane is given by

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

It turns out that the same exact thing works for lines! You should think about why this works to see if you really understand the above equation. But given a normal vector $\langle a, b\rangle$ to the line and a point (x_0, y_0) on the line, the equation of the line is

$$a(x - x_0) + b(y - y_0) = 0.$$

In our problem, the line passes through the point $(1, 1)$ and has normal vector $\langle -2, 1\rangle$ (the gradient vector of F at that point), so the equation of the tangent line is:

$$-2(x - 1) + 1(y - 1) = 0 \qquad \rightsquigarrow \qquad y = 1 + 2(x - 1).$$

Thankfully, this is exactly the line we got in the first part!

2 Computing a tangent plane

Having done the lower dimensional example above, let's tackle a tangent plane computation using two different perspectives:

Example 2.1. *Consider the surface $z = x^2 + y^2$. Find the equation of the tangent plane at the point* $(1, 1, 2)$.

2.1 The first perspective: linearization / tangent plane formula

This solution is the analogue of the first perspective solution above. In 15.4 we learn that the tangent plane to the graph of a function $f(x, y)$ at the point (a, b) is given by

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

We can view the surface $z = x^2 + y^2$ as the graph of the function $f(x, y) = x^2 + y^2$. Since $f(1, 1) = 2$ and

$$f_x(1, 1) = (2x)\mid_{(1,1)}= 2 \quad \text{and} \quad f_y(1, 1) = (2y)\mid_{(1,1)}$$

it follows that the equation of the tangent plane we seek is

$$z = 2 + 2(x - 1) + 2(y - 1).$$

2.2 The second perspective: level surfaces

This solution is the analogue of the second perspective solution above. Instead of viewing the surface $z = x^2 + y^2$ as the graph of the function $f(x, y) = x^2 + y^2$, we can alternatively **view it as a level surface of a three variable function**. In particular, let

$$F(x, y, z) = z - x^2 - y^2.$$

Then the surface $z = x^2 + y^2$, which when rewritten is the surface $z - x^2 - y^2 = 0$, is precisely the level surface $F(x, y, z) = 0$. We know that the gradient vector ∇F is perpendicular to level surfaces, so we can find a normal vector to the tangent plane we seek by computing $\nabla F(1, 1, 2)$. We have

$$\nabla F(x, y, z) = \langle -2x, -2y, 1\rangle$$

and so a normal vector to the tangent plane is

$$\nabla F(1, 1, 2) = \langle -2, -2, 1\rangle.$$

Thus, the tangent plane equation is

$$-2(x-1) + -2(y-1) + 1(z-2) = 0 \qquad \rightsquigarrow \qquad z = 2 + 2(x-1) + 2(y-1).$$

Exactly the same as what we found above!

3 One more tangent plane example

So what's the point? Maybe it's kind of cool that we can compute the same thing using two different perspectives, but why bother? If I can always use the first perspective, why should I worry about the second perspective? **Sometimes, one perspective is much better suited for a given problem.** Being able to tackle a math problem with a variety of viewpoints is in immeasurably important skill to have! For example,

Example 3.1. *Consider the surface* $x^2 \sin z + yx + \cos(yz) = 2$. *Find the equation of the tangent plane at the point* $(1, 1, 0)$.

Solution. This is an example where you can't really treat the surface as the graph of a two variable function, so we kind of have to take the second perspective.[1] It is much more natural to view the surface defined above as the level surface of a three variable function. In particular, let

$$F(x, y, z) = x^2 \sin z + yx + \cos(yz).$$

The surface we care about is the level surface $F(x, y, z) = 2$. Thus, to get a normal vector for the tangent plane, we can compute the gradient vector $\nabla F(1, 1, 0)$. Since

$$\nabla F(x, y, z) = \big\langle 2x \sin z + y, x - z\sin(yz), x^2 \cos z - y\sin(yz)\big\rangle$$

we have

$$\nabla F(1, 1, 0) = \langle 1, 1, 1\rangle.$$

This is a normal vector for the plane we seek. Thus, the equation of the tangent plane is

$$1(x-1) + 1(y-1) + 1(z-0) = 0 \quad \rightsquigarrow \quad x + y + z = 2.$$

Easy! This would be a disaster if you tried to solve for one of the variables in terms of the others. □

[1] This is a bit of a lie, locally near the point we care about we could treat it like a function but this is not the point. It's much easier to use the second perspective.

26. Directional Derivatives & The Gradient

Given a multivariable function $z = f(x, y)$ and a point on the *xy*-plane $P_0 = (x_0, y_0)$ at which f is differentiable (*i.e.* it is smooth with no discontinuities, folds or corners), there are infinitely many directions (relative to the *xy*-plane) in which to sketch a tangent line to f at P_0. A **directional derivative** is the slope of a tangent line to f at P_0 in which a *unit* direction vector $\mathbf{u} = \langle u_1, u_2 \rangle$ has been specified, and is given by the formula

$$D_{\mathbf{u}} f(x_0, y_0) = f_x(x_0, y_0) u_1 + f_y(x_0, y_0) u_2.$$

The right side of the equation can be viewed as the result of a dot product:

$$D_{\mathbf{u}} f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle \cdot \langle u_1, u_2 \rangle.$$

The vector-valued function $\langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle$ is called the **gradient** of f at $x = x_0$ and $y = y_0$, and is written $\nabla f(x_0, y_0)$. Thus, the directional derivative of f at P_0 in the direction of $\mathbf{u}$ is written in the shortened form

$$D_{\mathbf{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \mathbf{u}.$$

♦ • ♦ • ♦ • ♦ • ♦

Example 26.1: Find $\nabla f(x, y)$, where $f(x, y) = x^2 y + 2xy^3$.

Solution: Since $\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle$, we have $\nabla f(x, y) = \langle 2xy + 2y^3, x^2 + 6xy^2 \rangle$.

♦ • ♦ • ♦ • ♦ • ♦

Example 26.2: Find the slope of the tangent line of $f(x, y) = x^2 y + 2xy^3$ at $x_0 = -1$, $y_0 = 2$ in the direction of $\mathbf{u} = \langle 4,3 \rangle$.

Solution: From the previous example, $\nabla f(x, y) = \langle 2xy + 2y^3, x^2 + 6xy^2 \rangle$. When evaluated at $x_0 = -1$ and $y_0 = 2$, we have

$$\nabla f(-1,2) = \langle 2(-1)(2) + 2(2)^3, (-1)^2 + 6(-1)(2)^2 \rangle = \langle 12, -23 \rangle.$$

The direction $\mathbf{u}$ is not a unit vector. Since $|\mathbf{u}| = \sqrt{4^2 + 3^2} = \sqrt{25} = 5$, the unit vector in the direction of $\mathbf{u}$ is $\left\langle \frac{4}{5}, \frac{3}{5} \right\rangle$. Thus,

$$D_{\mathbf{u}} f(-1,2) = \langle 12, -23 \rangle \cdot \left\langle \frac{4}{5}, \frac{3}{5} \right\rangle = 12\left(\frac{4}{5}\right) - 23\left(\frac{3}{5}\right) = -\frac{21}{5}.$$

♦ • ♦ • ♦ • ♦ • ♦

Example 26.3: Find the slope of the tangent line of $g(x,y)=\frac{x}{y^2}$ at $x_0=3$ and $y_0=5$, in the direction of the origin.

Solution: The vector from (3,5) to (0,0) is given by $\langle 0-3, 0-5\rangle=\langle -3,-5\rangle$. Its magnitude is $\sqrt{(-3)^2+(-5)^2}=\sqrt{34}$. Thus, the unit direction vector is

$$\mathbf{u}=\left\langle -\frac{3}{\sqrt{34}}, -\frac{5}{\sqrt{34}}\right\rangle.$$

The gradient of g is

$$\nabla g(x,y)=\left\langle \frac{1}{y^2}, -\frac{2x}{y^3}\right\rangle.$$

Therefore,

$$\nabla g(3,5)=\left\langle \frac{1}{(5)^2}, -\frac{2(3)}{(5)^3}\right\rangle=\left\langle \frac{1}{25}, -\frac{6}{125}\right\rangle.$$

The slope of the tangent line of g at $x_0=3$ and $y_0=5$ in the direction of $\mathbf{u}$ is

$$\begin{aligned} D_{\mathbf{u}}g(3,5) &= \left\langle \frac{1}{25}, -\frac{6}{125}\right\rangle\cdot\left\langle -\frac{3}{\sqrt{34}}, -\frac{5}{\sqrt{34}}\right\rangle \\ &= \left(\frac{1}{25}\right)\left(-\frac{3}{\sqrt{34}}\right)+\left(-\frac{6}{125}\right)\left(-\frac{5}{\sqrt{34}}\right) \\ &= -\frac{15}{125\sqrt{34}}+\frac{30}{125\sqrt{34}}=\frac{15}{125\sqrt{34}}\approx 0.0206. \end{aligned}$$

♦ • ♦ • ♦ • ♦ • ♦

Example 26.4: Find the slope of the tangent line of $h(x,y)=\sqrt{1+x^2+y^2}$ where $P_0=(1,2)$ and the direction is given by a ray from P_0 oriented at $\theta=\frac{\pi}{6}$ radians, relative to the positive x-direction.

Solution: The direction vector is given by $\mathbf{u}=\left\langle \cos\left(\frac{\pi}{6}\right), \sin\left(\frac{\pi}{6}\right)\right\rangle=\left\langle \frac{\sqrt{3}}{2}, \frac{1}{2}\right\rangle$. It is a unit vector. The gradient of h is

$$\nabla h(x,y)=\left\langle \frac{x}{\sqrt{1+x^2+y^2}}, \frac{y}{\sqrt{1+x^2+y^2}}\right\rangle,$$

So upon substitution,

$$\nabla h(1,2)=\left\langle\frac{(1)}{\sqrt{1+(1)^2+(2)^2}},\frac{(2)}{\sqrt{1+(1)^2+(2)^2}}\right\rangle=\left\langle\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}}\right\rangle.$$

The directional derivative of h at (1,2) in the direction of $\mathbf{u}$ is

$$\begin{aligned}D_{\mathbf{u}}h(1,2)&=\left\langle\frac{\sqrt{3}}{2},\frac{1}{2}\right\rangle\cdot\left\langle\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}}\right\rangle\\&=\left(\frac{\sqrt{3}}{2}\right)\left(\frac{1}{\sqrt{6}}\right)+\left(\frac{1}{2}\right)\left(\frac{2}{\sqrt{6}}\right)\\&=\frac{\sqrt{3}+2}{2\sqrt{6}}\approx 0.762.\end{aligned}$$

♦ • ♦ • ♦ • ♦ • ♦

Directional derivatives can be extended into higher dimensions.

Example 26.5: Find the slope of the tangent line of $f(x,y,z)=xy^2z^3$ at $x_0=2, y_0=1$ and $z_0=3$ in the direction of $\langle 2,4,-5\rangle$.

Solution: The gradient of f is

$$\nabla f(x,y,z)=\langle f_x,f_y,f_z\rangle=\langle y^2z^3,2xyz^3,3xy^2z^2\rangle.$$

At (2,1,3), we have

$$\nabla f(2,1,3)=\langle 27,108,54\rangle.$$

The unit direction vector is $\mathbf{u}=\left\langle\frac{2}{\sqrt{45}},\frac{4}{\sqrt{45}},-\frac{5}{\sqrt{45}}\right\rangle$. The slope of the tangent line of f at (2,1,3) in the direction of $\mathbf{u}$ is

$$\begin{aligned}D_{\mathbf{u}}f(2,1,3)&=\nabla f(2,1,3)\cdot\mathbf{u}\\&=\langle 27,108,54\rangle\cdot\left\langle\frac{2}{\sqrt{45}},\frac{4}{\sqrt{45}},-\frac{5}{\sqrt{45}}\right\rangle\\&=\frac{54}{\sqrt{45}}+\frac{432}{\sqrt{45}}-\frac{270}{\sqrt{45}}\approx 32.2.\end{aligned}$$

♦ • ♦ • ♦ • ♦ • ♦

Using the cosine form of the formula for the dot product of two vectors, $\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos\theta$, we can rewrite $D_{\mathbf{u}}f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \mathbf{u}$ as

$$D_{\mathbf{u}}f(x_0, y_0) = |\nabla f(x_0, y_0)||\mathbf{u}| \cos\theta.$$

Since $\mathbf{u}$ is a unit vector, then $|\mathbf{u}| = 1$, so that

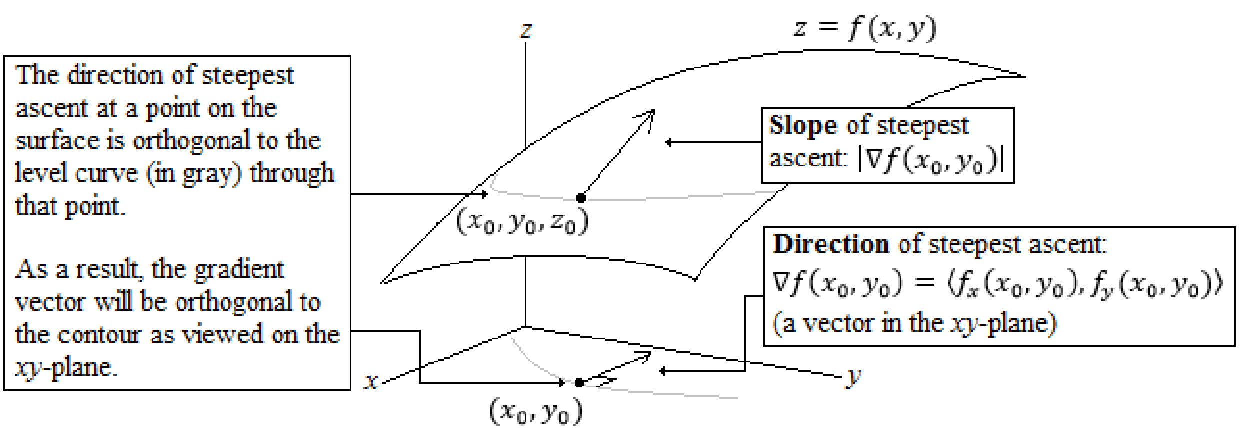
$$|\nabla f(x_0, y_0)||\mathbf{u}| \cos\theta = |\nabla f(x_0, y_0)| \cos\theta,$$

where θ is the angle between the gradient vector at (x_0, y_0), and the direction vector $\mathbf{u}$. From this, we can infer that $|\nabla f(x_0, y_0)| \cos\theta$ is maximized when $\nabla f(x_0, y_0)$ and $\mathbf{u}$ are parallel, or when $\theta = 0$ (so that $\cos\theta = 1$). This leads to a significant result in directional derivatives.

Given a function $z = f(x, y)$ and a point $P_0 = (x_0, y_0, z_0)$:

- The **direction of steepest ascent** at P_0 is given by $\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle$. In this case, it is permissible to state the direction as a non-unit vector.
- The **slope of steepest ascent** at P_0 is given by $|\nabla f(x_0, y_0)|$.
- The **direction of steepest descent** at P_0 is opposite the direction of steepest ascent, and is given by $-\nabla f(x_0, y_0) = \langle -f_x(x_0, y_0), -f_y(x_0, y_0)\rangle$.
- The **slope of steepest descent** at P_0 is $-|\nabla f(x_0, y_0)|$.

A path that follows the directions of steepest ascent is called a **gradient path** and is always orthogonal to the contours of the surface.



◆ • ◆ • ◆ • ◆ • ◆

Example 26.6: Let $f(x,y) = x^2 + 2xy^2$. State the direction(s) in which the slope of the tangent line at $x_0 = 2$ and $y_0 = 1$ is 0.

Solution: We have $\nabla f(x,y) = \langle 2x + 2y^2, 4xy\rangle$. Let $\mathbf{u} = \langle u_1, u_2\rangle$. We have

$$\begin{aligned} D_{\mathbf{u}} f(2,1) &= \nabla f(2,1) \cdot \mathbf{u} \\ &= \langle 6,8\rangle \cdot \langle u_1, u_2\rangle \\ &= 6u_1 + 8u_2. \end{aligned}$$

If the slope is to be 0, we set $6u_1 + 8u_2 = 0$. Thus, whenever $u_2 = -\frac{3}{4}u_1$, then the slope of the tangent line at $x_0 = 2$ and $y_0 = 1$ will be 0.

♦ • ♦ • ♦ • ♦ • ♦

Example 26.7: Find the direction of steepest ascent of $f(x,y) = x^2y + 2xy^3$ at $x_0 = -1$ and $y_0 = 2$, then find the slope of steepest ascent.

Solution: From an earlier example, we found that $\nabla f(x,y) = \langle 2xy + 2y^3, x^2 + 6xy^2\rangle$ and that $\nabla f(-1,2) = \langle 12, -23\rangle$. This is the *direction* of steepest ascent. The *slope* of steepest ascent is $|\langle 12, -23\rangle| = \sqrt{12^2 + (-23)^2} \approx 25.94$.

When finding a directional derivative where the direction is stated or to be determined, you *must* be sure that it is stated as a unit vector. However, when asked to find a direction of steepest ascent, it is permissible to leave it as a non-unit vector since you will likely be calculating the slope as well. While it is not incorrect to state the direction of steepest ascent as a unit vector, a common error is to then use that unit vector to find the slope, in which case the answer will be 1, which is likely incorrect.

♦ • ♦ • ♦ • ♦ • ♦

Example 26.8: Suppose the slope of the tangent line of $z = f(x,y)$ at $P_0 = (x_0, y_0)$ in the direction of $\langle 3,1\rangle$ is $\sqrt{10}$, and that the slope of the tangent line at the same point in the direction of $\langle 1,4\rangle$ is $\frac{18}{\sqrt{17}}$. What is the direction of steepest ascent of f at P_0, and what is the slope in this direction?

Solution: We don't know f, but we can treat the components in its gradient, $\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle$, as a pair of unknowns. In the direction of $\langle 3,1\rangle$, the slope of the tangent line is $\sqrt{10}$. Considering the unit direction vector $\mathbf{u} = \left\langle \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right\rangle$, we have $D_{\mathbf{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \mathbf{u} = \sqrt{10}$. Thus, we initially have

$$\langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle \cdot \left\langle \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right\rangle = \sqrt{10},$$

which gives

$$f_x(x_0, y_0)\frac{3}{\sqrt{10}} + f_y(x_0, y_0)\frac{1}{\sqrt{10}} = \sqrt{10}. \quad (\mathbf{1})$$

In a similar way, we consider the unit direction vector in the direction of $\langle 1,4\rangle$, which is $\left\langle \frac{1}{\sqrt{17}}, \frac{4}{\sqrt{17}} \right\rangle$. The slope in this direction is $\frac{18}{\sqrt{17}}$. We have

$$\langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle \cdot \left\langle \frac{1}{\sqrt{17}}, \frac{4}{\sqrt{17}} \right\rangle = \frac{18}{\sqrt{17}},$$

which gives

$$f_x(x_0, y_0)\frac{1}{\sqrt{17}} + f_y(x_0, y_0)\frac{4}{\sqrt{17}} = \frac{18}{\sqrt{17}}. \quad (\mathbf{2})$$

Taking equations **(1)** and **(2)** together, we have a system of two unknowns in two equations:

$$f_x(x_0, y_0)\frac{3}{\sqrt{10}} + f_y(x_0, y_0)\frac{1}{\sqrt{10}} = \sqrt{10}$$
$$f_x(x_0, y_0)\frac{1}{\sqrt{17}} + f_y(x_0, y_0)\frac{4}{\sqrt{17}} = \frac{18}{\sqrt{17}}.$$

The first equation is multiplied by $\sqrt{10}$, and the second by $\sqrt{17}$ to clear fractions:

$$f_x(x_0, y_0)(3) + f_y(x_0, y_0)(1) = 10$$
$$f_x(x_0, y_0)(1) + f_y(x_0, y_0)(4) = 18.$$

The bottom equation is multiplied by –3:

$$f_x(x_0, y_0)(3) + f_y(x_0, y_0)(1) = 10$$
$$f_x(x_0, y_0)(-3) + f_y(x_0, y_0)(-12) = -54.$$

Adding the second equation to the first, we have $-11f_y\,(x_0, y_0) = -44$. Thus, $f_y(x_0, y_0) = 4$. Substituting this into either of the equations **(1)** or **(2)**, we find that $f_x(x_0, y_0) = 2$. Therefore, we now know $\nabla f(x_0, y_0)$, which is $\langle 2,4\rangle$. This is the direction of steepest ascent of f. The slope at P_0 in this direction is $\sqrt{2^2 + 4^2} = \sqrt{20} = 2\sqrt{5} \approx 4.47$.

♦ • ♦ • ♦ • ♦ • ♦

Example 26.9: A plane tilts to the north at a 6% grade – that is, for every 100 feet one moves horizontally north, he or she will gain 6 feet vertically. Find the slope and the grade if someone walks to the northeast.

Solution: Assume the plane passes through the origin, assuming also that the *y*-axis is north and south, and the *x*-axis is east and west, in the usual map orientation. When $y = 100$, we have $z = 6$, so that another ordered triple on the plane is (0,100,6). Thus, we can write $z = \frac{6}{100}y = 0.06y$ as the equation of the plane. The gradient of f is $\nabla f(x, y) = \langle 0, 0.06\rangle$. Note that *x* is an independent variable but has no effect on the values of *z*. If it helps, write the plane as $z = 0x + 0.06y$.

Furthermore, at the origin, we still have $\nabla f(0,0) = \langle 0, 0.06\rangle$. Meanwhile, movement to the northeast can be modeled by the vector $\langle 1,1\rangle$, or as a unit vector, $\mathbf{u} = \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\rangle$.

The slope at the origin in the direction of northeast is given by

$$\begin{aligned} D_{\mathbf{u}} f(0,0) &= \nabla f(0,0) \cdot \mathbf{u} \\ &= \langle 0, 0.06\rangle \cdot \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\rangle \\ &= \frac{0.06}{\sqrt{2}} \approx 0.0424. \end{aligned}$$

The grade can be inferred by the fact that 1 foot of movement in the northeast direction results in a rise of 0.0424 feet vertically. Thus, the grade is about 4.24%.

Note that a movement east or west would result in no change in *z*. The directional derivative in either direction (the positive or negative *x* direction) is 0. Let $\mathbf{u} = \langle 1,0\rangle$ or $\langle -1,0\rangle$ and verify that the directional derivative would be 0.

♦ • ♦ • ♦ • ♦ • ♦

Section 3-5 : Lagrange Multipliers

In the previous section we optimized (*i.e.* found the absolute extrema) a function on a region that contained its boundary. Finding potential optimal points in the interior of the region isn't too bad in general, all that we needed to do was find the critical points and plug them into the function. However, as we saw in the examples finding potential optimal points on the boundary was often a fairly long and messy process.

In this section we are going to take a look at another way of optimizing a function subject to given constraint(s). The constraint(s) may be the equation(s) that describe the boundary of a region although in this section we won't concentrate on those types of problems since this method just requires a general constraint and doesn't really care where the constraint came from.

So, let's get things set up. We want to optimize (*i.e.* find the minimum and maximum value of) a function, $f\left(x,y,z\right)$, subject to the constraint $g\left(x,y,z\right)=k$. Again, the constraint may be the equation that describes the boundary of a region or it may not be. The process is actually fairly simple, although the work can still be a little overwhelming at times.

Method of Lagrange Multipliers

1. Solve the following system of equations.
$$\begin{aligned}\nabla f\left(x,y,z\right)&=\lambda\ \nabla g\left(x,y,z\right)\\ g\left(x,y,z\right)&=k\end{aligned}$$
2. Plug in all solutions, $\left(x,y,z\right)$, from the first step into $f\left(x,y,z\right)$ and identify the minimum and maximum values, provided they exist and $\nabla g\neq\vec{0}$ at the point.

The constant, λ, is called the **Lagrange Multiplier**.

Notice that the system of equations from the method actually has four equations, we just wrote the system in a simpler form. To see this let's take the first equation and put in the definition of the gradient vector to see what we get.

$$\left\langle f_x,f_y,f_z\right\rangle=\lambda\left\langle g_x,g_y,g_z\right\rangle=\left\langle\lambda g_x,\lambda g_y,\lambda g_z\right\rangle$$

In order for these two vectors to be equal the individual components must also be equal. So, we actually have three equations here.

$$f_x=\lambda g_x\qquad f_y=\lambda g_y\qquad f_z=\lambda g_z$$

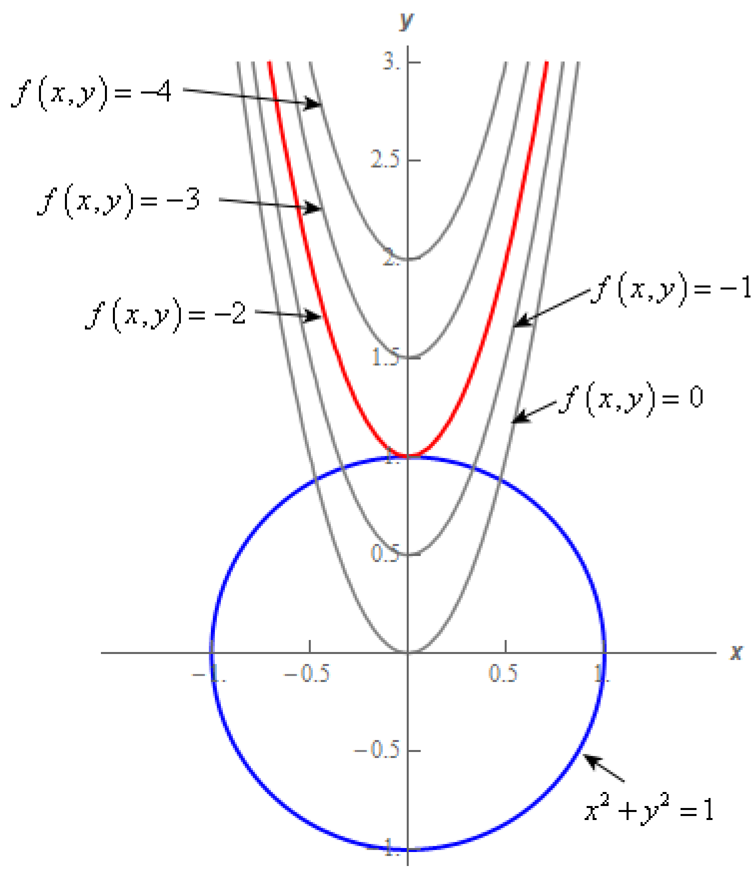
These three equations along with the constraint, $g\left(x,y,z\right)=c$, give four equations with four unknowns *x*, *y*, *z*, and λ.

Note as well that if we only have functions of two variables then we won't have the third component of the gradient and so will only have three equations in three unknowns *x*, *y*, and λ.

As a final note we also need to be careful with the fact that in some cases minimums and maximums won't exist even though the method will seem to imply that they do. In every problem we'll need to make sure that minimums and maximums will exist before we start the problem.

To see a physical justification for the formulas above. Let's consider the minimum and maximum value of $f(x,y) = 8x^2 - 2y$ subject to the constraint $x^2 + y^2 = 1$. In the practice problems for this section (problem #2 to be exact) we will show that minimum value of $f(x,y)$ is -2 which occurs at $(0,1)$ and the maximum value of $f(x,y)$ is 8.125 which occurs at $\left(-\frac{3\sqrt{7}}{8}, -\frac{1}{8}\right)$ and $\left(\frac{3\sqrt{7}}{8}, -\frac{1}{8}\right)$.

Here is a sketch of the constraint as well as $f(x.y) = k$ for various values of *k*.



First remember that solutions to the system must be somewhere on the graph of the constraint, $x^2 + y^2 = 1$ in this case. Because we are looking for the minimum/maximum value of $f(x,y)$ this, in turn, means that the location of the minimum/maximum value of $f(x,y)$, *i.e.* the point (x,y), must occur where the graph of $f(x,y) = k$ intersects the graph of the constraint when *k* is either the minimum or maximum value of $f(x,y)$.

Now, we can see that the graph of $f(x,y) = -2$, *i.e.* the graph of the minimum value of $f(x,y)$, just touches the graph of the constraint at $(0,1)$. In fact, the two graphs at that point are tangent.

If the two graphs are tangent at that point then their normal vectors must be parallel, *i.e.* the two normal vectors must be scalar multiples of each other. Mathematically, this means,

$$\nabla f(x,y,z) = \lambda \ \nabla g(x,y,z)$$

for some scalar λ and this is exactly the first equation in the system we need to solve in the method.

Note as well that if *k* is smaller than the minimum value of $f(x,y)$ the graph of $f(x,y) = k$ doesn't intersect the graph of the constraint and so it is not possible for the function to take that value of *k* at a point that will satisfy the constraint.

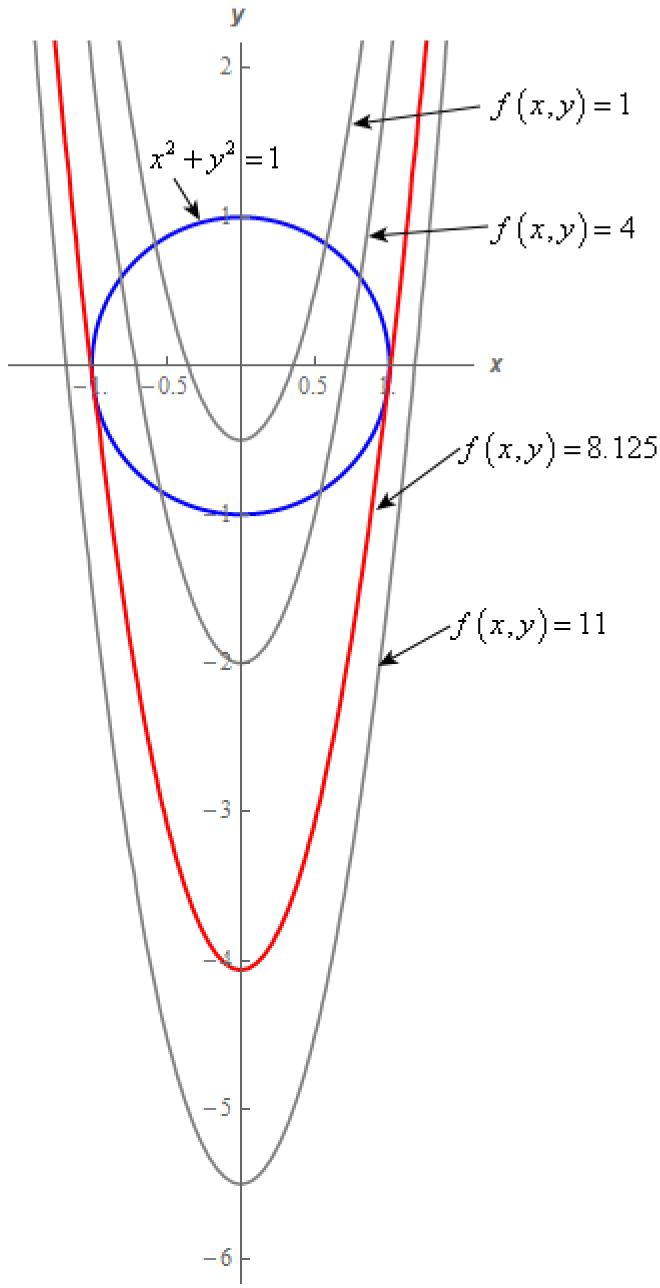
Likewise, if *k* is larger than the minimum value of $f(x,y)$ the graph of $f(x,y) = k$ will intersect the graph of the constraint but the two graphs are not tangent at the intersection point(s). This means that the method will not find those intersection points as we solve the system of equations.

Next, the graph below shows a different set of values of *k*. In this case, the values of *k* include the maximum value of $f(x,y)$ as well as a few values on either side of the maximum value.

Again, we can see that the graph of $f(x,y) = 8.125$ will just touch the graph of the constraint at two points. This is a good thing as we know the solution does say that it should occur at two points. Also note that at those points again the graph of $f(x,y) = 8.125$ and the constraint are tangent and so, just as with the minimum values, the normal vectors must be parallel at these points.

Likewise, for value of *k* greater than 8.125 the graph of $f(x,y) = k$ does not intersect the graph of the constraint and so it will not be possible for $f(x,y)$ to take on those larger values at points that are on the constraint.

Also, for values of *k* less than 8.125 the graph of $f(x,y) = k$ does intersect the graph of the constraint but will not be tangent at the intersection points and so again the method will not produce these intersection points as we solve the system of equations.



So, with these graphs we've seen that the minimum/maximum values of $f(x,y)$ will come where the graph of $f(x,y)=k$ and the graph of the constraint are tangent and so their normal vectors are parallel. Also, because the point must occur on the constraint itself. In other words, the system of equations we need to solve to determine the minimum/maximum value of $f(x,y)$ are exactly those given in the above when we introduced the method.

Note that the physical justification above was done for a two dimensional system but the same justification can be done in higher dimensions. The difference is that in higher dimensions we won't be working with curves. For example, in three dimensions we would be working with surfaces. However, the same ideas will still hold. At the points that give minimum and maximum value(s) of $f\left(x,y,z\right)$ the surfaces would be parallel and so the normal vectors would also be parallel.

Let's work a couple of examples.

Example 1 Find the dimensions of the box with largest volume if the total surface area is 64 cm^2.

Solution

Before we start the process here note that we also saw a way to solve this kind of problem in Calculus I, except in those problems we required a condition that related one of the sides of the box to the other sides so that we could get down to a volume and surface area function that only involved two variables. We no longer need this condition for these problems.

Now, let's get on to solving the problem. We first need to identify the function that we're going to optimize as well as the constraint. Let's set the length of the box to be *x*, the width of the box to be *y* and the height of the box to be *z*. Let's also note that because we're dealing with the dimensions of a box it is safe to assume that *x*, *y*, and *z* are all positive quantities.

We want to find the largest volume and so the function that we want to optimize is given by,

$$f\left(x,y,z\right)=xyz$$

Next, we know that the surface area of the box must be a constant 64. So this is the constraint. The surface area of a box is simply the sum of the areas of each of the sides so the constraint is given by,

$$2xy+2xz+2yz=64 \qquad \Rightarrow \qquad xy+xz+yz=32$$

Note that we divided the constraint by 2 to simplify the equation a little. Also, we get the function $g\left(x,y,z\right)$ from this.

$$g\left(x,y,z\right)=xy+xz+yz$$

The function itself, $f\left(x,y,z\right)=xyz$ will clearly have neither minimums or maximums unless we put some restrictions on the variables. The only real restriction that we've got is that all the variables must be positive. This, of course, instantly means that the function does have a minimum, zero, even though this is a silly value as it also means we pretty much don't have a box. It does however mean that we know the minimum of $f\left(x,y,z\right)$ does exist.

So, let's now see if $f\left(x,y,z\right)$ will have a maximum. Clearly, hopefully, $f\left(x,y,z\right)$ will not have a maximum if all the variables are allowed to increase without bound. That however, can't happen because of the constraint,

$$xy+xz+yz=32$$

Here we've got the sum of three positive numbers (remember that we *x*, *y*, and *z* are positive because we are working with a box) and the sum must equal 32. So, if one of the variables gets very large, say *x*, then because each of the products must be less than 32 both *y* and *z* must be very small to make sure the first two terms are less than 32. So, there is no way for all the variables to increase without bound and so it should make some sense that the function, $f(x,y,z) = xyz$, will have a maximum.

This is not an exact proof that $f(x,y,z)$ will have a maximum but it should help to visualize that $f(x,y,z)$ should have a maximum value as long as it is subject to the constraint.

Here are the four equations that we need to solve.

$$yz = \lambda(y+z) \qquad (f_x = \lambda g_x) \qquad (1)$$

$$xz = \lambda(x+z) \qquad (f_y = \lambda g_y) \qquad (2)$$

$$xy = \lambda(x+y) \qquad (f_z = \lambda g_z) \qquad (3)$$

$$xy + xz + yz = 32 \qquad (g(x,y,z) = 32) \qquad (4)$$

There are many ways to solve this system. We'll solve it in the following way. Let's multiply equation (1) by *x*, equation (2) by *y* and equation (3) by *z*. This gives,

$$xyz = \lambda x(y+z) \qquad (5)$$

$$xyz = \lambda y(x+z) \qquad (6)$$

$$xyz = \lambda z(x+y) \qquad (7)$$

Now notice that we can set equations (5) and (6) equal. Doing this gives,

$$\begin{aligned} \lambda x(y+z) &= \lambda y(x+z) \\ \lambda(xy+xz) - \lambda(yx+yz) &= 0 \\ \lambda(xz-yz) &= 0 \qquad \Rightarrow \qquad \lambda = 0 \quad \text{or} \quad xz = yz \end{aligned}$$

This gave two possibilities. The first, $\lambda = 0$ is not possible since if this was the case equation (1) would reduce to

$$yz = 0 \qquad \Rightarrow \qquad y = 0 \ \text{ or } \ z = 0$$

Since we are talking about the dimensions of a box neither of these are possible so we can discount $\lambda = 0$. This leaves the second possibility.

$$xz = yz$$

Since we know that $z \neq 0$ (again since we are talking about the dimensions of a box) we can cancel the *z* from both sides. This gives,

$$x = y \qquad (8)$$

Next, let's set equations (6) and (7) equal. Doing this gives,

$$\lambda y\left(x+z\right)=\lambda z\left(x+y\right)$$
$$\lambda\left(yx+yz-zx-zy\right)=0$$
$$\lambda\left(yx-zx\right)=0 \qquad \Rightarrow \qquad \lambda=0 \text{ or } yx=zx$$

As already discussed we know that $\lambda=0$ won't work and so this leaves,

$$yx=zx$$

We can also say that $x\neq 0$ since we are dealing with the dimensions of a box so we must have,

$$z=y \tag{9}$$

Plugging equations (8) and (9) into equation (4) we get,

$$y^2+y^2+y^2=3y^2=32 \qquad y=\pm\sqrt{\frac{32}{3}}=\pm 3.266$$

However, we know that *y* must be positive since we are talking about the dimensions of a box. Therefore, the only solution that makes physical sense here is

$$x=y=z=3.266$$

So, it looks like we've got a cube.

We should be a little careful here. Since we've only got one solution we might be tempted to assume that these are the dimensions that will give the largest volume. Anytime we get a single solution we really need to verify that it is a maximum (or minimum if that is what we are looking for).

This is actually pretty simple to do. First, let's note that the volume at our solution above is,

$$V=f\left(\sqrt{\frac{32}{3}},\sqrt{\frac{32}{3}},\sqrt{\frac{32}{3}}\right)=\left(\sqrt{\frac{32}{3}}\right)^3=34.8376$$

Now, we know that a maximum of $f\left(x,y,z\right)$ will exist ("proved" that earlier in the solution) and so to verify that that this really is a maximum all we need to do if find another set of dimensions that satisfy our constraint and check the volume. If the volume of this new set of dimensions is smaller that the volume above then we know that our solution does give a maximum.

If, on the other hand, the new set of dimensions give a larger volume we have a problem. We only have a single solution and we know that a maximum exists and the method should generate that maximum. So, in this case, the likely issue is that we will have made a mistake somewhere and we'll need to go back and find it.

So, let's find a new set of dimensions for the box. The only thing we need to worry about is that they will satisfy the constraint. Outside of that there aren't other constraints on the size of the dimensions. So, we can freely pick two values and then use the constraint to determine the third value.

Let's choose $x = y = 1$. No reason for these values other than they are "easy" to work with. Plugging these into the constraint gives,

$$1 + z + z = 32 \qquad \to \qquad 2z = 31 \qquad \to \qquad z = \frac{31}{2}$$

So, this is a set of dimensions that satisfy the constraint and the volume for this set of dimensions is,

$$V = f\left(1, 1, \frac{31}{2}\right) = \frac{31}{2} = 15.5 < 34.8376$$

So, the new dimensions give a smaller volume and so our solution above is, in fact, the dimensions that will give a maximum volume of the box are $x = y = z = 3.266$

Notice that we never actually found values for λ in the above example. This is fairly standard for these kinds of problems. The value of λ isn't really important to determining if the point is a maximum or a minimum so often we will not bother with finding a value for it. On occasion we will need its value to help solve the system, but even in those cases we won't use it past finding the point.

Example 2 Find the maximum and minimum of $f(x, y) = 5x - 3y$ subject to the constraint $x^2 + y^2 = 136$.

Solution

This one is going to be a little easier than the previous one since it only has two variables. Also, note that it's clear from the constraint that region of possible solutions lies on a disk of radius $\sqrt{136}$ which is a closed and bounded region, $-\sqrt{136} \le x, y \le \sqrt{136}$, and hence by the Extreme Value Theorem we know that a minimum and maximum value must exist.

Here is the system that we need to solve.

$$\begin{aligned} 5 &= 2\lambda x \\ -3 &= 2\lambda y \\ x^2 + y^2 &= 136 \end{aligned}$$

Notice that, as with the last example, we can't have $\lambda = 0$ since that would not satisfy the first two equations. So, since we know that $\lambda \ne 0$ we can solve the first two equations for *x* and *y* respectively. This gives,

$$x = \frac{5}{2\lambda} \qquad\qquad y = -\frac{3}{2\lambda}$$

Plugging these into the constraint gives,

$$\frac{25}{4\lambda^2} + \frac{9}{4\lambda^2} = \frac{17}{2\lambda^2} = 136$$

We can solve this for λ.

$$\lambda^2 = \frac{1}{16} \qquad \Rightarrow \qquad \lambda = \pm\frac{1}{4}$$

Now, that we know λ we can find the points that will be potential maximums and/or minimums.

If $\lambda = -\frac{1}{4}$ we get,

$$x = -10 \qquad\qquad y = 6$$

and if $\lambda = \frac{1}{4}$ we get,

$$x = 10 \qquad\qquad y = -6$$

To determine if we have maximums or minimums we just need to plug these into the function. Also recall from the discussion at the start of this solution that we know these will be the minimum and maximums because the Extreme Value Theorem tells us that minimums and maximums will exist for this problem.

Here are the minimum and maximum values of the function.

$$f(-10,6) = -68 \qquad \text{Minimum at } (-10,6)$$
$$f(10,-6) = 68 \qquad \text{Maximum at } (10,-6)$$

In the first two examples we've excluded $\lambda = 0$ either for physical reasons or because it wouldn't solve one or more of the equations. Do not always expect this to happen. Sometimes we will be able to automatically exclude a value of λ and sometimes we won't.

Let's take a look at another example.

Example 3 Find the maximum and minimum values of $f(x,y,z) = xyz$ subject to the constraint $x + y + z = 1$. Assume that $x, y, z \ge 0$.

Solution
First note that our constraint is a sum of three positive or zero number and it must be 1. Therefore, it is clear that our solution will fall in the range $0 \le x, y, z \le 1$ and so the solution must lie in a closed and bounded region and so by the Extreme Value Theorem we know that a minimum and maximum value must exist.

Here is the system of equation that we need to solve.

$$yz = \lambda \qquad (10)$$
$$xz = \lambda \qquad (11)$$
$$xy = \lambda \qquad (12)$$

$$x + y + z = 1 \tag{13}$$

Let's start this solution process off by noticing that since the first three equations all have λ they are all equal. So, let's start off by setting equations (10) and (11) equal.

$$yz = xz \qquad \Rightarrow \qquad z(y - x) = 0 \qquad \Rightarrow \qquad z = 0 \text{ or } y = x$$

So, we've got two possibilities here. Let's start off with by assuming that $z = 0$. In this case we can see from either equation (10) or (11) that we must then have $\lambda = 0$. From equation (12) we see that this means that $xy = 0$. This in turn means that either $x = 0$ or $y = 0$.

So, we've got two possible cases to deal with there. In each case two of the variables must be zero. Once we know this we can plug into the constraint, equation (13), to find the remaining value.

$$\begin{aligned} z = 0,\ x = 0: &\qquad \Rightarrow \qquad y = 1 \\ z = 0,\ y = 0: &\qquad \Rightarrow \qquad x = 1 \end{aligned}$$

So, we've got two possible solutions $(0,1,0)$ and $(1,0,0)$.

Now let's go back and take a look at the other possibility, $y = x$. We also have two possible cases to look at here as well.

This first case is $x = y = 0$. In this case we can see from the constraint that we must have $z = 1$ and so we now have a third solution $(0,0,1)$.

The second case is $x = y \neq 0$. Let's set equations (11) and (12) equal.

$$xz = xy \qquad \Rightarrow \qquad x(z - y) = 0 \qquad \Rightarrow \qquad x = 0 \text{ or } z = y$$

Now, we've already assumed that $x \neq 0$ and so the only possibility is that $z = y$. However, this also means that,

$$x = y = z$$

Using this in the constraint gives,

$$3x = 1 \qquad \Rightarrow \qquad x = \frac{1}{3}$$

So, the next solution is $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

We got four solutions by setting the first two equations equal.

To completely finish this problem out we should probably set equations (10) and (12) equal as well as setting equations (11) and (12) equal to see what we get. Doing this gives,

$$yz = xy \qquad \Rightarrow \qquad y(z-x) = 0 \qquad \Rightarrow \qquad y = 0 \text{ or } z = x$$
$$xz = xy \qquad \Rightarrow \qquad x(z-y) = 0 \qquad \Rightarrow \qquad x = 0 \text{ or } z = y$$

Both of these are very similar to the first situation that we looked at and we'll leave it up to you to show that in each of these cases we arrive back at the four solutions that we already found.

So, we have four solutions that we need to check in the function to see whether we have minimums or maximums.

$$f(0,0,1) = 0 \qquad f(0,1,0) = 0 \qquad f(1,0,0) = 0 \qquad \text{All Minimums}$$
$$f\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) = \frac{1}{27} \qquad \text{Maximum}$$

So, in this case the maximum occurs only once while the minimum occurs three times.

Note as well that we never really used the assumption that $x, y, z \geq 0$ in the actual solution to the problem. We used it to make sure that we had a closed and bounded region to guarantee we would have absolute extrema. To see why this is important let's take a look at what might happen without this assumption. Without this assumption it wouldn't be too difficult to find points that give both larger and smaller values of the functions. For example.

$$x = -100,\ y = 100,\ z = 1: \ -100 + 100 + 1 = 1 \qquad f(-100,100,1) = -10000$$
$$x = -50,\ y = -50,\ z = 101: \ -50 - 50 + 101 = 1 \qquad f(-50,-50,101) = 252500$$

With these examples you can clearly see that it's not too hard to find points that will give larger and smaller function values. However, all of these examples required negative values of *x*, *y* and/or *z* to make sure we satisfy the constraint. By eliminating these we will know that we've got minimum and maximum values by the Extreme Value Theorem.

Before we proceed we need to address a quick issue that the last example illustrates about the method of Lagrange Multipliers. We found the absolute minimum and maximum to the function. However, what we did not find is all the locations for the absolute minimum. For example, assuming $x, y, z \geq 0$, consider the following sets of points.

$$(0, y, z) \qquad \text{where} \quad y + z = 1$$
$$(x, 0, z) \qquad \text{where} \quad x + z = 1$$
$$(x, y, 0) \qquad \text{where} \quad x + y = 1$$

Every point in this set of points will satisfy the constraint from the problem and in every case the function will evaluate to zero and so also give the absolute minimum.

So, what is going on? Recall from the previous section that we had to check both the critical points and the boundaries to make sure we had the absolute extrema. The same was true in Calculus I. We had to check both critical points and end points of the interval to make sure we had the absolute extrema.

It turns out that we really need to do the same thing here if we want to know that we've found all the locations of the absolute extrema. The method of Lagrange multipliers will find the absolute extrema, it just might not find all the locations of them as the method does not take the end points of variables ranges into account (note that we might luck into some of these points but we can't guarantee that).

So, after going through the Lagrange Multiplier method we should then ask what happens at the end points of our variable ranges. For the example that means looking at what happens if $x = 0$, $y = 0$, $z = 0$, $x = 1$, $y = 1$, and $z = 1$. In the first three cases we get the points listed above that do happen to also give the absolute minimum. For the later three cases we can see that if one of the variables are 1 the other two must be zero (to meet the constraint) and those were actually found in the example. Sometimes that will happen and sometimes it won't.

In the case of this example the end points of each of the variable ranges gave absolute extrema but there is no reason to expect that to happen every time. In Example 2 above, for example, the end points of the ranges for the variables do not give absolute extrema (we'll let you verify this).

The moral of this is that if we want to know that we have every location of the absolute extrema for a particular problem we should also check the end points of any variable ranges that we might have. If all we are interested in is the value of the absolute extrema then there is no reason to do this.

Okay, it's time to move on to a slightly different topic. To this point we've only looked at constraints that were equations. We can also have constraints that are inequalities. The process for these types of problems is nearly identical to what we've been doing in this section to this point. The main difference between the two types of problems is that we will also need to find all the critical points that satisfy the inequality in the constraint and check these in the function when we check the values we found using Lagrange Multipliers.

Let's work an example to see how these kinds of problems work.

Example 4 Find the maximum and minimum values of $f\left(x,y\right) = 4x^2 + 10y^2$ on the disk $x^2 + y^2 \le 4$.

Solution

Note that the constraint here is the inequality for the disk. Because this is a closed and bounded region the Extreme Value Theorem tells us that a minimum and maximum value must exist.

The first step is to find all the critical points that are in the disk (*i.e.* satisfy the constraint). This is easy enough to do for this problem. Here are the two first order partial derivatives.

$$\begin{aligned} f_x &= 8x & &\Rightarrow & 8x &= 0 & &\Rightarrow & x &= 0 \\ f_y &= 20y & &\Rightarrow & 20y &= 0 & &\Rightarrow & y &= 0 \end{aligned}$$

So, the only critical point is $\left(0,0\right)$ and it does satisfy the inequality.

At this point we proceed with Lagrange Multipliers and we treat the constraint as an equality instead of the inequality. We only need to deal with the inequality when finding the critical points.

So, here is the system of equations that we need to solve.

$$8x = 2\lambda x$$
$$20y = 2\lambda y$$
$$x^2 + y^2 = 4$$

From the first equation we get,

$$2x(4-\lambda) = 0 \qquad \Rightarrow \qquad x = 0 \quad \text{or} \quad \lambda = 4$$

If we have $x = 0$ then the constraint gives us $y = \pm 2$.

If we have $\lambda = 4$ the second equation gives us,

$$20y = 8y \qquad \Rightarrow \qquad y = 0$$

The constraint then tells us that $x = \pm 2$.

If we'd performed a similar analysis on the second equation we would arrive at the same points.

So, Lagrange Multipliers gives us four points to check : $(0,2)$, $(0,-2)$, $(2,0)$, and $(-2,0)$.

To find the maximum and minimum we need to simply plug these four points along with the critical point in the function.

$$f(0,0) = 0 \qquad \text{Minimum}$$
$$f(2,0) = f(-2,0) = 16$$
$$f(0,2) = f(0,-2) = 40 \qquad \text{Maximum}$$

In this case, the minimum was interior to the disk and the maximum was on the boundary of the disk.

The final topic that we need to discuss in this section is what to do if we have more than one constraint. We will look only at two constraints, but we can naturally extend the work here to more than two constraints.

We want to optimize $f(x,y,z)$ subject to the constraints $g(x,y,z) = c$ and $h(x,y,z) = k$. The system that we need to solve in this case is,

$$\nabla f(x,y,z) = \lambda \nabla g(x,y,z) + \mu \nabla h(x,y,z)$$
$$g(x,y,z) = c$$
$$h(x,y,z) = k$$

So, in this case we get two Lagrange Multipliers. Also, note that the first equation really is three equations as we saw in the previous examples. Let's see an example of this kind of optimization problem.

Example 5 Find the maximum and minimum of $f(x, y, z) = 4y - 2z$ subject to the constraints $2x - y - z = 2$ and $x^2 + y^2 = 1$.

Solution

Verifying that we will have a minimum and maximum value here is a little trickier. Clearly, because of the second constraint we've got to have $-1 \le x, y \le 1$. With this in mind there must also be a set of limits on *z* in order to make sure that the first constraint is met. If one really wanted to determine that range you could find the minimum and maximum values of $2x - y$ subject to $x^2 + y^2 = 1$ and you could then use this to determine the minimum and maximum values of *z*. We won't do that here. The point is only to acknowledge that once again the possible solutions must lie in a closed and bounded region and so minimum and maximum values must exist by the Extreme Value Theorem.

Here is the system of equations that we need to solve.

$$0 = 2\lambda + 2\mu x \qquad (f_x = \lambda g_x + \mu h_x) \qquad (14)$$

$$4 = -\lambda + 2\mu y \qquad (f_y = \lambda g_y + \mu h_y) \qquad (15)$$

$$-2 = -\lambda \qquad (f_z = \lambda g_z + \mu h_z) \qquad (16)$$

$$2x - y - z = 2 \qquad (17)$$

$$x^2 + y^2 = 1 \qquad (18)$$

First, let's notice that from equation (16) we get $\lambda = 2$. Plugging this into equation (14) and equation (15) and solving for *x* and *y* respectively gives,

$$0 = 4 + 2\mu x \qquad \Rightarrow \qquad x = -\frac{2}{\mu}$$

$$4 = -2 + 2\mu y \qquad \Rightarrow \qquad y = \frac{3}{\mu}$$

Now, plug these into equation (18).

$$\frac{4}{\mu^2} + \frac{9}{\mu^2} = \frac{13}{\mu^2} = 1 \qquad \Rightarrow \qquad \mu = \pm\sqrt{13}$$

So, we have two cases to look at here. First, let's see what we get when $\mu = \sqrt{13}$. In this case we know that,

$$x = -\frac{2}{\sqrt{13}} \qquad y = \frac{3}{\sqrt{13}}$$

Plugging these into equation (17) gives,

$$-\frac{4}{\sqrt{13}} - \frac{3}{\sqrt{13}} - z = 2 \qquad \Rightarrow \qquad z = -2 - \frac{7}{\sqrt{13}}$$

So, we've got one solution.

Let's now see what we get if we take $\mu = -\sqrt{13}$. Here we have,

$$x = \frac{2}{\sqrt{13}} \qquad\qquad y = -\frac{3}{\sqrt{13}}$$

Plugging these into equation (17) gives,

$$\frac{4}{\sqrt{13}} + \frac{3}{\sqrt{13}} - z = 2 \qquad \Rightarrow \qquad z = -2 + \frac{7}{\sqrt{13}}$$

and there's a second solution.

Now all that we need to is check the two solutions in the function to see which is the maximum and which is the minimum.

$$f\left(-\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}, -2 - \frac{7}{\sqrt{13}}\right) = 4 + \frac{26}{\sqrt{13}} = 11.2111$$

$$f\left(\frac{2}{\sqrt{13}}, -\frac{3}{\sqrt{13}}, -2 + \frac{7}{\sqrt{13}}\right) = 4 - \frac{26}{\sqrt{13}} = -3.2111$$

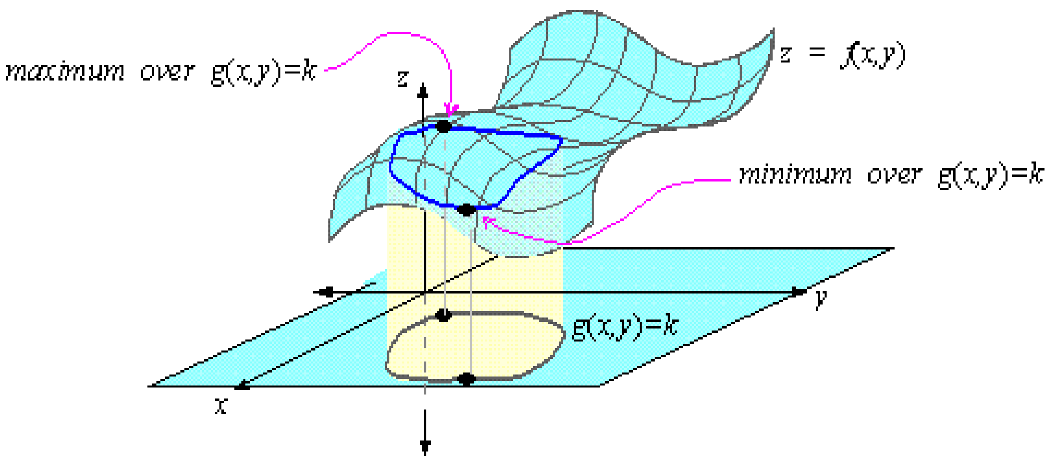
So, we have a maximum at $\left(-\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}, -2 - \frac{7}{\sqrt{13}}\right)$ and a minimum at $\left(\frac{2}{\sqrt{13}}, -\frac{3}{\sqrt{13}}, -2 + \frac{7}{\sqrt{13}}\right)$.

Lagrange Multipliers

Optimization with Constraints

In many applications, we must find the extrema of a function $f(x,y)$ subject to a *constraint* $g(x,y) = k$. Such problems are called *constrained optimization* problems.

For example, suppose that the constraint $g(x,y) = k$ is a smooth closed curve parameterized by $\mathbf{r}(t) = \langle x(t), y(t) \rangle$ on $[a,b]$, and suppose that $f(x,y)$ is differentiable at each point on the constraint. Then finding the extrema of $f(x,y)$ subject to $g(x,y) = k$ is equivalent to finding the absolute extrema of the function $z(t) = f(x(t), y(t))$ for t in $[a,b]$.



In a first calculus course, we learn that the extrema of $z(t)$ over $[a,b]$ must exist and occur either at the critical points or the endpoints of $[a,b]$. Since the curve is closed, we only need consider the critical points of $z(t)$ in $[a,b]$, which are solutions to

$$\frac{dz}{dt} = \nabla f \cdot \mathbf{v} = 0$$

where $\mathbf{v}$ is the velocity of $\mathbf{r}(t)$. That is, the critical points of $z(t)$ occur when $\nabla f \perp \mathbf{v}$. Since also $\nabla g \perp \mathbf{v}$, it follows that the extrema of $f(x,y)$ subject to $g(x,y) = k$ occur when ∇f is parallel to ∇g.

If ∇f is parallel to ∇g, then there is a number λ for which

$$\nabla f = \lambda \nabla g$$

Thus, the extrema of $f(x,y)$ subject to $g(x,y) = k$ must occur at the points which are the solution to the system of equations

$$\langle f_x, f_y \rangle = \lambda \langle g_x, g_y \rangle, \qquad g(x,y) = k \tag{1}$$

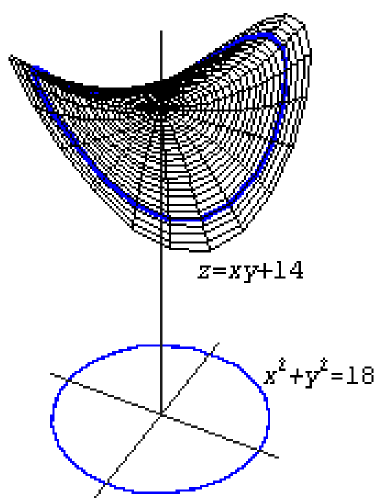
We call (1) a *Lagrange multiplier problem* and we call λ a *Lagrange Multiplier.*

A good approach to solving a Lagrange multiplier problem is to *first eliminate the Lagrange multiplier* λ using the two equations $f_x = \lambda g_x$ and $f_y = \lambda g_y$. Then solve for x and y by combining the result with the constraint $g(x, y) = k$, thus producing the critical points. Finally, since the constraint $g(x, y) = k$ is a closed curve, the extrema of $f(x, y)$ over $g(x, y) = k$ are the largest and smallest values of $f(x, y)$ evaluated at the critical points.

EXAMPLE 1 Find the extrema of $f(x, y) = xy + 14$ subject to

$$x^2 + y^2 = 18$$

Solution: That is, we want to find the highest and lowest points on the surface $z = xy + 14$ over the unit circle:



If we let $g(x, y) = x^2 + y^2$, then the constraint is $g(x, y) = 18$. The gradients of f and g are respectively

$$\nabla f = \langle y, x \rangle \qquad \text{and} \qquad \nabla g = \langle 2x, 2y \rangle$$

As a result, $\nabla f = \lambda \nabla g$ implies that

$$y = \lambda 2x \quad \textit{and} \quad x = \lambda 2y$$

. Clearly, $x = 0$ only if $y = 0$, but $(0, 0)$ is not on the unit circle. Thus, $x \neq 0$ and $y \neq 0$, so that solving for λ yields

$$\lambda = \frac{y}{2x} \quad \text{and} \quad \lambda = \frac{x}{2y} \qquad \Longrightarrow \qquad \frac{y}{2x} = \frac{x}{2y}$$

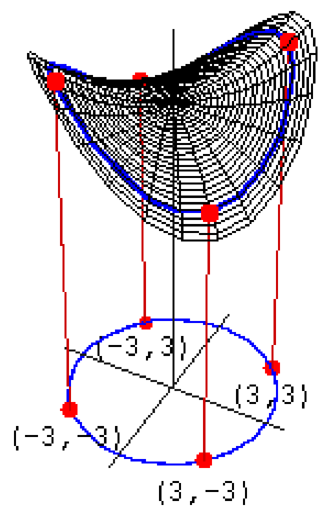
Cross-multiplying then yields $2y^2 = 2x^2$, which is the same as $y^2 = x^2$. Thus, the constraint $x^2 + y^2 = 18$ becomes

$$x^2 + x^2 = 18, \qquad x^2 = 9, \qquad x = \pm 3$$

Moreover, $y^2 = x^2$ implies that either $y = x$ or $y = -x$, so that the solutions to (2) are

$$(3,3)\,,\ (-3,3)\,,\ (3,-3)\,,\ (-3,-3)$$

However, $f(3,3) = f(-3,-3) = 23$, while $f(-3,3) = f(3,-3) = 5$. Thus, the maxima of $f(x,y) = xy + 4$ over $x^2 + y^2 = 18$ occur at $(3,3)$ and $(-3,-3)$, while the minima of $f(x,y) = xy + 4$ occur at $(-3,3)$ and $(3,-3)$.



More generally, finding the extrema of a differentiable function $f(x,y)$ subject to a constraint $g(x,y) = k$ is defined in terms of a *Lagrangian*, which is a function of *three variables* x, y, and λ of the form

$$L(x,y,\lambda) = f(x,y) - \lambda(g(x,y) - k)$$

This is because the critical points of $L(x,y,\lambda)$ occur when

$$\frac{\partial L}{\partial x} = 0, \qquad \frac{\partial L}{\partial y} = 0, \qquad \frac{\partial L}{\partial \lambda} = 0$$

and $L_x = f_x - \lambda g_x$, $L_y = f_y - \lambda g_y$, and $L_\lambda = g(x,y) - k$. That is, the critical points of $L(x,y,\lambda)$ are solutions of the system of equations

$$f_x = \lambda g_x, \qquad f_y = \lambda g_y, \quad g(x,y) = k \tag{2}$$

which is the same as the equations in (1).

blue*EXAMPLE 2* blackShow that the Lagrangian for the problem of finding the extrema of $f(x,y) = xy + 14$ subject to

$$x^2 + y^2 = 18$$

reduces to the solution in example 1.

Solution: The Lagrangian for example 1 is

$$L(x,y,\lambda) = xy + 14 - \lambda\left(x^2 + y^2 - 18\right)$$

and correspondingly, $L_x = y - \lambda(2x)$, $L_y = y - \lambda(2y)$, and

$$L_\lambda = -\left(x^2 + y^2 - 18\right)$$

The critical points of L satisfy $L_x = 0$, $L_y = 0$, and $L_\lambda = 0$, which results in

$$y = \lambda 2x \quad and \quad x = \lambda 2y$$

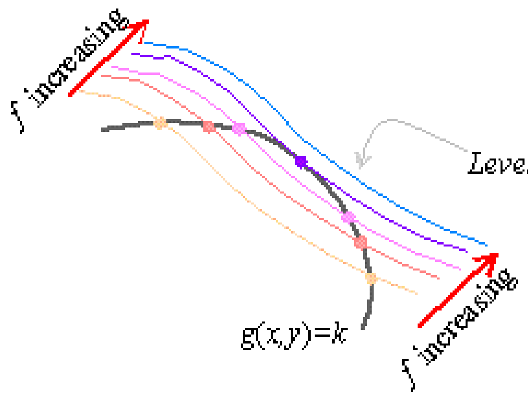
along with $x^2 + y^2 = 18$. The remainder of the solution is the same as in example 1.

Check your Reading: Can you identify the maxima and minima on the graph shown above.

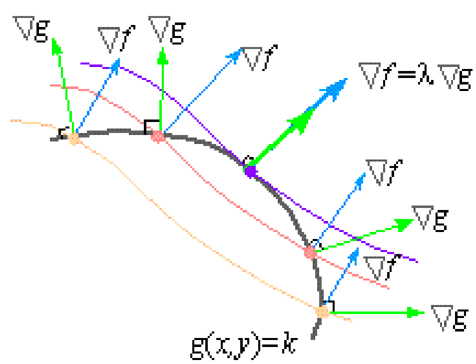
Lagrange Multipliers and Level Curves

Let's view the Lagrange Multiplier method in a different way, one which only requires that $g(x,y) = k$ have a *smooth* parameterization $\mathbf{r}(t)$ with t in a closed interval $[a,b]$. Such constraints are said to be *smooth and compact.*

If $f(x,y)$ is differentiable and to be optimized subject to a smooth compact constraint $g(x,y) = k$, then as the levels of $f(x,y)$ increase, short sections of level curves of $f(x,y)$ form secant curves to $g(x,y)=k$



It follows that the highest level curve of $f(x,y)$ intersecting a section of $g(x,y)=k$ must be tangent to the curve $g(x,y)=k$, which is possible only if their gradients ∇f and ∇g are parallel.



Consequently, if $g(x,y)=k$ is smooth and compact, then finding the extrema of $f(x,y)$ subject to $g(x,y)=k$ is equivalent to finding the critical points onf the function $L(x,y,\lambda)$, and then evaluating $f(x,y)$ at those critical points (and if applicable the boundary points of $g(x,y)=k$).

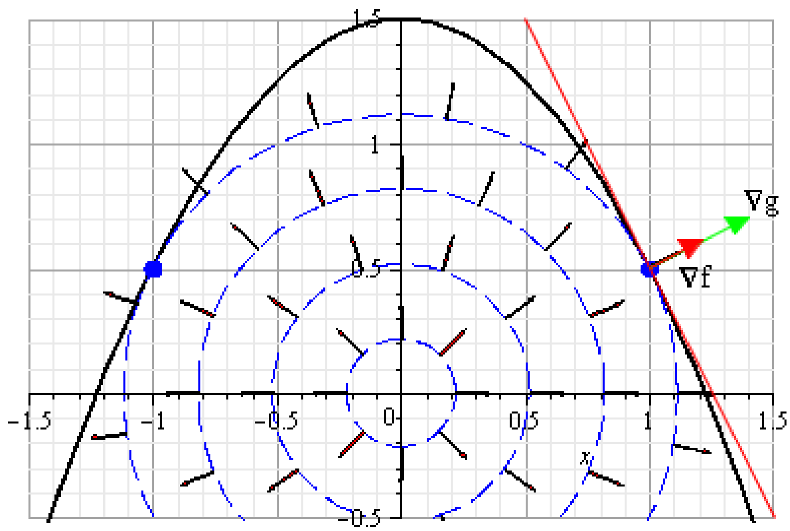
> *EXAMPLE 3* Find the point(s) on the curve $y=1.5-x^2$ closest to the origin both visually and via the Lagrange Multiplier method.

> **Solution:** If we let $f(x,y)$ be the square of the distance from a point (x,y) to the origin $(0,0)$, then our constrained optimization

problem is to

$$\begin{aligned} \text{Minimize } \ f(x,y) &= x^2+y^2 \\ \text{Subject to } x^2+y &= 1.5 \end{aligned}$$

We will thus let $g(x,y) = x^2 + y$. Graphically, we can find the point on $y = 1.5 - x^2$ closest to the origin by drawing concentric circles centered at the origin with ever greater radii until they interesect the curve. The first intersection of a circle with the curve will correspond to a circle tangent to the curve – i.e., a point where ∇f is parallel to ∇g.



Points with $|x| > 1.5$ are more distant than any of the points for $|x| \le 1.5$, so the point $(0.5, 1)$ and $(-0.5, 1)$ are the closest to the origin. Alternatively, the Lagrangian for this problem is

$$L(x,y,\lambda) = x^2 + y^2 - \lambda\left(x^2 + y - 1.5\right)$$

Since $L_x = 2x - \lambda(2x)$, $L_y = 2y - \lambda(1)$, and $L_\lambda = x^2 + y - 1.5$, the critical points of L occur when

$$2x = \lambda 2x, \qquad 2y = \lambda, \qquad x^2 + y = 1.5$$

To eliminate λ, we substitute the second equation $\lambda = 2y$ into the first equation to obtain

$$2x = (2y)\,2x, \qquad x = 2xy$$

If $x = 0$, then $y = 1.5$. Thus, $(0, 1.5)$ is the critical point corresponding to $\lambda = 3$. If $x \neq 0$, then $1 = 2y$ so that $y = 0.5$ and

$$x^2 + 0.5 = 1.5, \qquad x = \pm 1$$

Thus, the critical points are $(0, -1)$, $(1, 0.5)$ and $(-1, 0.5)$. However,

$$f(1, 0.5) = f(-1, 0.5) = 1.25, \qquad f(0, 1.5) = 2.25$$

Since we need only consider points with $|x| \leq 1.5$, the points on $y = 1.5 - x^2$ which are closest to the origin are

$$(1, 0.5) \quad \text{and} \quad (-1, 0.5)$$

Similarly, an equation of the form $g(x, y, z) = k$ defines a *level surface* in 3-dimensions, and finding the extrema of a function $f(x, y, z)$ subject to a constraint $g(x, y, z) = k$ is equivalent to finding a level surface of $f(x, y, z)$ that is tangent to a constraint surface $g(x, y, z) = k$. It follows that ∇f is parallel to ∇g at this point, so that if $g(x, y, z) = k$ is a *closed, bounded* surface (i.e,. a surface with finite extent that contains its boundary points), then solving

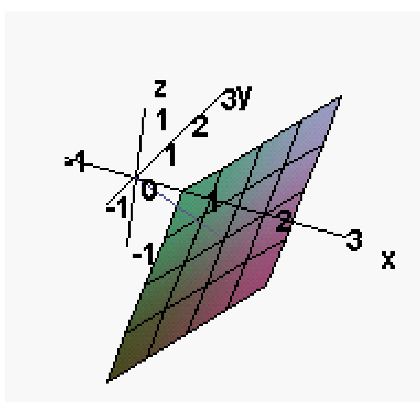
$$\begin{aligned} \textit{Optimize } w &= f(x, y, z) \\ \textit{Subject to } g(x, y, z) &= k \end{aligned}$$

is equivalent to finding the critical points of the Lagrangian in 4 variables given by

$$L(x, y, z, \lambda) = f(x, y, z) - \lambda (g(x, y, z) - k)$$

and then evaluating $f(x, y, z)$ at those critical points (and if applicable the boundary points of $g(x, y, z) = k$). Let's revisit a problem from the previous section to see this idea at work.

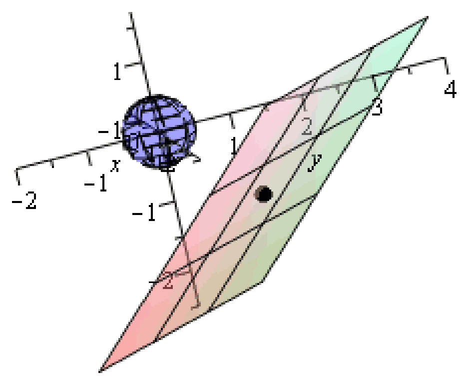
blue*EXAMPLE 4* blackFind the point(s) on the plane $x + y - z = 3$ that are closest to the origin.



Solution: To begin with, we let f denote the *square* of the distance from a point (x, y, z) to the origin. Consequently, $f = x^2 + y^2 + z^2$. Thus, we want to

$$\begin{aligned} \text{minimize } f(x,y,z) &= x^2+y^2+z^2 \\ \text{subject to:} \quad & x+y-z=3 \end{aligned}$$

Graphically, we can locate the closest point by drawing concentric spheres expanding until they intersect the plane, thus resulting in the "closest point" on a sphere that is tangent to the surface.



Alternatively, we can define the Lagrangian to be

$$L(x,y,z,\lambda) = x^2+y^2+z^2-\lambda(x+y-z-3)$$

Since $L_x = 2x - \lambda$, $L_y = 2y - \lambda$, $L_z = 2z + \lambda$, and $L_\lambda = -(x+y-z-3)$, the critical points are solutions to

$$2x = \lambda, \; 2y = \lambda, \; -2z = \lambda, \; \textit{and } x+y-z=3 \tag{3}$$

To eliminate λ, we notice that the first two equations imply that $y = z$, while the first and third imply that $z = -x$. Substituting into the constraint (the last equation in (3)) leads to

$$x + x - (-x) = 3, \quad 3x = 3, \; x = 1$$

Thus, $y = 1$ and $z = -1$, so the critical point is $(1, 1, -1)$. Points of the plane not shown in the figure above are further from the origin than the points that are shown in the figure, and the section of the plane shown in the figure is a closed subset of the plane whose boundary points (i.e., the edges) are more than $\sqrt{3}$ away from the origin. Since the set is closed, a minimum of $f(x, y, z)$ must occur, yet it cannot occur on the boundary. Instead, it must occur at the critical point $(1, 1, -1)$ which is a distance of $\sqrt{3}$ from the origin.

Check your Reading: Can $\lambda = 0$ in example 4?

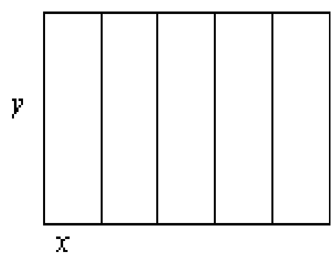
Applications

Many of the optimization word problems in a first calculus course are, in fact, constrained optimization problems of the form

$$\begin{array}{lcl} \textit{Optimize} & : & f(x,y) \\ \textit{Subject to} & : & g(x,y) = k \end{array}$$

In such problems, the Lagrange multiplier method produces a family of "extrema" of $f(x,y)$ parameterized by λ, so that eliminating λ and combining the result with the constraint is essentially equivalent to finding which member of the family satisfies the constraint.

EXAMPLE 5 John happens to acquire 420 feet of fencing and decides to use it to start a kennel by building 5 identical adjacent rectangular runs (see diagram below). Find the dimensions of each run that maximizes its area.



Solution: We let A denote the area of a run, and we let x, y be the dimensions of each run. Clearly, there are to be 10 sections of fence corresponding to widths x and 6 sections of fence corresponding to lengths y. Thus, we desire to maximize $A = xy$ subject to the constraint

$$10x + 6y = 420$$

Since x and y cannot be negative, we need only find absolute extrema for x in $[0, 42]$.
The Lagrangian for the problem is

$$L(x,y,\lambda) = xy - \lambda(10x + 6y - 420)$$

and $L_x = y - 10\lambda$, $L_y = x - 6\lambda$, and $L_\lambda = 10x + 6y - 420$. Thus, the critical points of $L(x,y,\lambda)$ satisfy

$$y = 10\lambda, \quad x = 6\lambda, \qquad 10x + 6y = 420$$

The first two equations parameterize the extrema in the parameter λ, which is why we eliminate λ to obtain $\lambda = y/10$ and $\lambda = x/6$. Thus,

$$\frac{y}{10} = \frac{x}{6}, \qquad y = \frac{10x}{6} = \frac{5x}{3}$$

Substituting into the constraint thus yields

$$10x + 6\left(\frac{5x}{3}\right) = 420, \qquad x = 21\ feet$$

Moreover, we also have $y = 5 \cdot 21/3 = 5 \cdot 7 = 35$ feet. At $x = 0$ and $x = 42$, the area is 0, while at the critical point $(20, 35)$, the area is 700 square feet. Thus, the maximum occurs when $x = 21$ feet and $y = 35$ feet.

If possible, a good approach to eliminating λ in a system of equations of the form

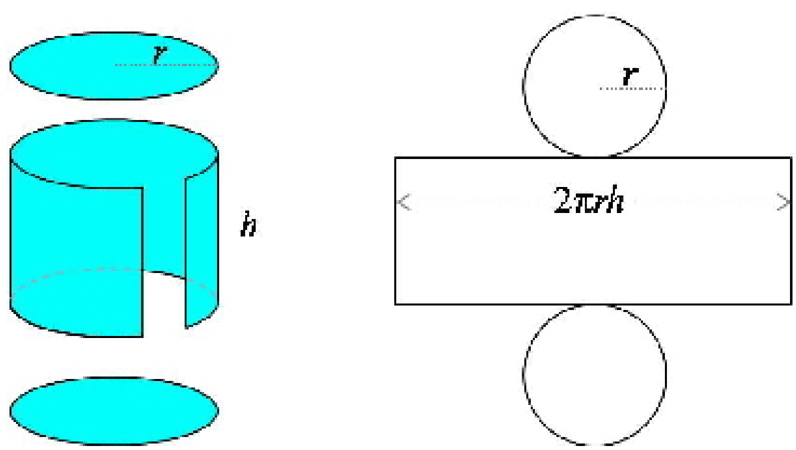
$$f_x = \lambda g_x, \quad f_y = \lambda g_y$$

is that of dividing the former by the latter to obtain

$$\frac{f_x}{f_y} = \frac{\lambda g_x}{\lambda g_y} \quad \Longrightarrow \quad \frac{f_x}{f_y} = \frac{g_x}{g_y}$$

and then cross-multiplying to obtain $f_x g_y = f_y g_x$. However, this method is not possible if one or more of the factors is zero.

EXAMPLE 6 A right cylindrical can is to have a volume of 0.25 cubic feet (approximately 2 gallons). Find the height h and radius r that will minimize surface area of the can. What is the relationship between the resulting r and h?



Solution: The surface area S is the sum of the areas of 2 circles of radius r and a rectangle with height h and width $2\pi r$. Thus,

$$S = 2\pi r^2 + 2\pi r h$$

This is constrained by a volume of $\pi r^2 h = 0.25\ ft^3$, so that the Lagrangian is

$$L(r, h, \lambda) = 2\pi r^2 + 2\pi r h - \lambda\left(\pi r^2 h - 0.25\right)$$

Setting $L_r = 0$ and $L_h = 0$ leads to

$$4\pi r + 2\pi h = \lambda(2\pi r h), \quad 2\pi r = \lambda\left(\pi r^2\right)$$

The ratio of the two equations is

$$\frac{4\pi r + 2\pi h}{2\pi r} = \frac{\lambda(2\pi r h)}{\lambda(\pi r^2)} \qquad \Longrightarrow \qquad \frac{2r + h}{r} = \frac{2h}{r}$$

Cross-multiplying yields $2r^2 + rh = 2rh$, which in turn yields

$$rh = 2r^2 \quad or \quad h = 2r$$

since r cannot be 0. That is, all cans with minimal surface area will have $h = 2r$, which means a height equal to the diameter. To determine which such can satisfies the constraint, we substitute to obtain

$$\pi r^2(2r) = 0.25, \quad r^3 = \sqrt[3]{\frac{0.25}{2\pi}}$$

which leads to $r = 0.3414$ feet, with $h = 0.6818$ feet. To see that a minimum must occur, we notice that the constraint implies that $h = 0.25/\left(\pi r^2\right)$, which leads to S as a function of r in the form

$$S = 2\pi r^2 + \frac{0.5}{r^2}$$

Straightforward differentiation shows that $S''(r) > 0$ for all $r > 0$, so that any extremum must be a minimum.

Check Your Reading: What is the value of the Lagrange multiplier in example 5?

Multiple Constraints

Typically, if given a constraint of the form $g(x, y) = k$, we instead let $g_1(x, y) = g(x, y) - k$ and use the constraint $g_1(x, y) = 0$. Thus, Lagrangians are usually of the form

$$L(x, y, z, \lambda) = f(x, y, z) - \lambda g_1(x, y, z)$$

Correspondingly, to find the extrema of a function $f(x,y,z)$ subject to *two* constraints,

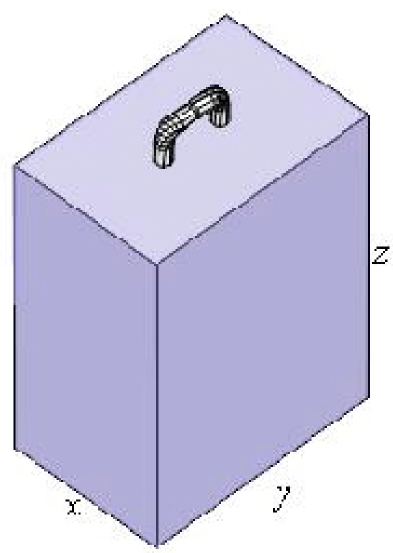
$$g(x,y,z)=k, \qquad h(x,y,z)=l$$

we define a function of the 3 variables x, y,and z, and the Lagrange multipliers λ and μ by

$$L(x,y,z,\lambda,\mu)=f(x,y,z)-\lambda g_1(x,y,z)-\mu h_1(x,y,z)$$

where $g_1(x,y,z)=g(x,y,z)-k$ and $h_1(x,y,z)=h(x,y,z)-l$. As before, the goal is to determine the critical points of the Lagrangian.

blue*EXAMPLE 7* blackMany airlines require that carry-on luggage have a linear distance (sum of length, width, height) of no more than 45 inches with an additional requirement of being able to slide under the seat in front of you.



If we assume that the carry-on is to have (at least roughly) the shape of a rectangular box and one dimension is no more than half of one of the other dimensions (to insure "slide under seat" is possible), then what dimensions of the carryon lead to maximum storage (i.e., maximum volume)?

Solution: If we let x, y, and z denote length, width, and height, respectively, then our goal is to maximize the volume $V(x,y,z)$ subject to the constraints

$$x+y+z=45 \quad and \quad y=2x$$

(i.e., x is $1/2$ of y). Thus, $g_1(x,y,z) = x+y+z-45$ and $h_1(x,y,z) = y-2x$ leads to a Lagrangian of the form

$$\begin{aligned} L(x,y,z,\lambda,\mu) &= f(x,y,z,) - \lambda g_1(x,y,z,) - \mu h_1(x,y,z,) \\ &= xyz - \lambda(x+y+z-45) - \mu(y-2x) \end{aligned}$$

The partial derivatives of L are

$$\begin{aligned} L_x &= yz - \lambda - \mu(-2), \quad L_y = xz - \lambda - \mu \\ L_z &= xy - \lambda \end{aligned}$$

and $L_\lambda = x+y+z-45$, $L_\mu = y-2x$. The critical points thus must satisfy

$$yz = \lambda - 2\mu, \quad xz = \lambda + \mu, \quad xy = \lambda$$

along with the constraints. Combining the last two equations yields $xz = xy + \mu$, so that the first equation becomes

$$yz = xy - 2(xz - xy) \quad or \quad yz = 3xy - 2xz$$

Since $y = 2x$, this becomes

$$2xz = 6x^2 - 2xz \quad or \quad 4xz = 6x^2$$

Since $x = 0$ leads to a zero volume, we must have $2z = 3x$, or $z = 1.5x$. Substituting into the first constraint yields

$$x + 2x + 1.5x = 45$$

which is $4.5x = 45$ or $x = 10$. If $x = 10$, then $y = 2x = 20$ and $z = 1.5x = 15$, so that the critical point is $(10,20,15)$. Since x, y, and z must all be in $[0,45]$, we are seeking the extrema of the volume over a closed set (in particular, the closed box $[0,45]\times[0,45]\times[0,45]$) and the volume is zero on the boundary. Thus, the maximum volume must occur, and the only place left for it to occur is at the critical point $(10,20,15)$.

Of course, we could substituted $y = 2x$ directly to reduce example 7 to an ordinary Lagrange multiplier problem:

$$\begin{aligned} Maximize\ Volume &: \quad V = 2x^2 z \\ Subject\ to &: \quad x + 2x + z = 45 \end{aligned}$$

However, not all multiple constraint problems share this feature. Moreover, example 7 illustrates how the Lagrange multiplier method can be applied to optimizing a function f of any number of variables subject to any given collection of constraints. Specifically, as is shown in the accompanying worksheet, the associated Lagrangian has a multiplier term corresponding to each constraint.

Exercises

Use the method of Lagrange Multipliers to find the extrema of the following functions subject to the given constraints. (Notice that in each problem below the constraint is a closed curve).

1.	$f(x,y) = 3x + 2y$ subject to: $x^2 + y^2 = 1$	2.	$f(x,y) = 2x - y$ subject to: $x^2 + y^2 = 1$
3.	$f(x,y) = x - 2y$ subject to: $x^2 + y^2 = 25$	4.	$f(x,y) = x + y$ subject to: $x^2 + 2y^2 = 1$
5.	$f(x,y) = x^2 + 2y^2$ subject to: $x^2 + y^2 = 1$	6.	$f(x,y) = x^2 y$ subject to: $x^2 + y^2 = 1$
7.	$f(x,y) = x^4 + y^2$ subject to: $x^2 + y^2 = 1$	8.	$f(x,y) = x^4 + y^4$ subject to: $x^2 + y^2 = 1$
9.	$f(x,y) = x \sin(y)$ subject to: $x^2 + y^2 = 1$	10.	$f(x,y) = \sin(x)\cos(y)$ subject to: $x^2 + y^2 = 1$

Find the point(s) on the given curve closest to the origin.

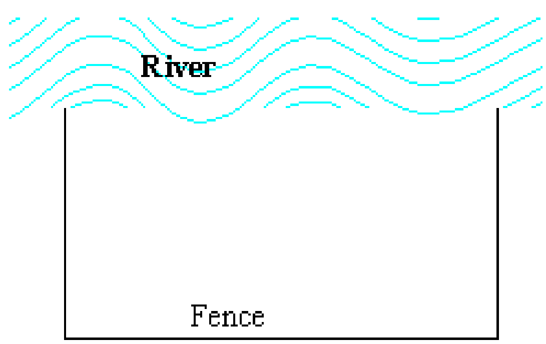
11.	$x + y = 1$	12.	$x + 2y = 5$
13.	$x = y^2 - 1$	14.	$x^2 + 2y^2 = 1$
15.	$y = e^{-x^2/2+2}$	16.	$y = 2e^{-x^2/2}$

Use the method of Lagrange multipliers in problems 17-30

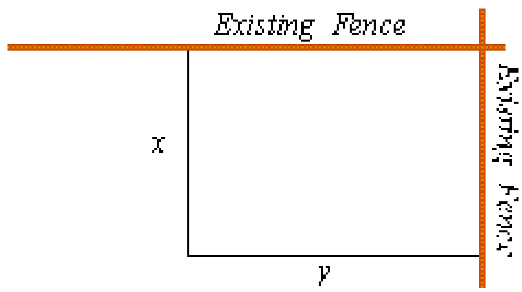
17. Maximize the product of two positive numbers whose sum is 36.

18. Minimize the sum of two positive numbers whose product is 36.

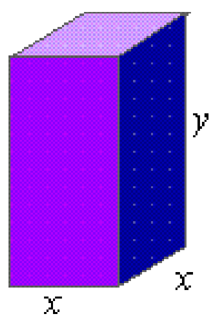
19. A farmer has 400 feet of fence with which to enclose a rectangular field bordering a river. What dimensions of the field maximize the area if the field is to be fenced on only three sides (see picture below)?



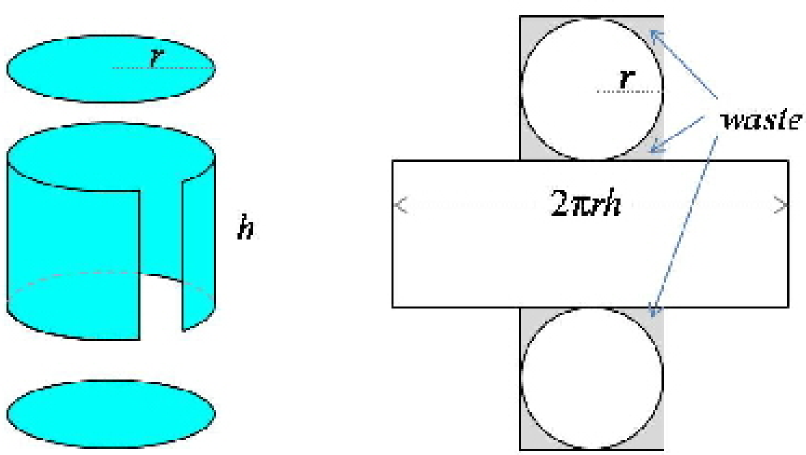
20. A farmer has 400 feet of fence with which to fence in a rectangular field adjoining two existing fences which meet at a right angle. What dimensions maximize the area of the field?



21. A rectangular box with a square bottom is to have a volume of 1000 ft^3. What dimensions for the box yield the smallest surface area?



22. A soupcan can is to have a volume of 25 cubic inches, and it is to be made with as little metal as possible. The manufacturing process makes the cans by rolling up rectangles of metal and capping each end with circles *punched from a square whose sides are the lengths of the diameters of the top and bottom.*



This means minimizing the surface area of the can *as well as the "wasted area" between the ends and the squares they are punched from.* Use Lagrange multipliers to find r and h that minimize the metal used to make the can. What is the relationship between the resulting r and h?

23. What dimensions of the carry on in example 7 yield maximum volume if we drop the requirement of one dimension being no more than half of another.

24. What dimensions of the carry on in example 7 yield maximum volume if we drop the requirement of one dimension being no more than half the other and *set* $z = 9$ (i.e., find x and y).

25. Redo example 7 for those airlines that allow 51 linear inches for a carryon.

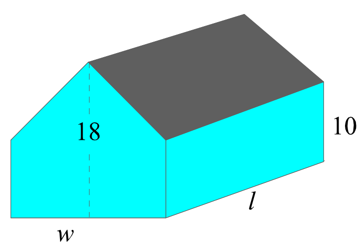
26. What dimensions of the carry on in example 7 yield maximum volume if we replace the constraint $y = 2x$ by the constraint $xy = 300$? What does this new constraint represent?

27.

27. The intersection of the plane $z = 4.5 + 0.5x$ and the cone $x^2 + y^2 = z^2$ is an ellipse. What point(s) on the ellipse are furthest from or closest to the origin? What is significant about these points?

28. The intersection of the plane $z = 1 + x$ and the cone $x^2 + y^2 = z^2$ is a parabola. What point on the parabola is closes to the origin? What is significant about these point?

29. A house with width w and length l is be 18 feet tall at its tallest and 10 feet tall at each corner.



What dimensions for a 2000 square foot house (i.e., $wl = 2000$) minimize the area of the roof and sides of the house?

30. The moon's orbit about the earth is well-approximated by the curve

$$x^2 + y^2 = (238,957 - 0.0549y)^2$$

where distance is in miles. How close is the moon to the earth at its closest point? What is the greatest distance between the moon and the earth?

Optimization with constraints occurs frequently in business settings. For example, if L denotes the number of manhours and K denotes the number of units of capital required to produce q units of a commodity, then q is often related to L and K by a Cobb-Douglas function

$$q = AL^{\alpha}K^{\beta} \tag{4}$$

where A is a constant, α is the product elasticity of labor, and β is the product elasticity of capital. If one unit of labor costs w dollars and one unit of capital costs r dollars, then the cost to manufacture q units of a commodity is

$$C = wL + rK \tag{5}$$

Often the number of items to be produced, q, is a constant, so that (4) is used as a constraint to (5). Use the Cobb-Douglas production model in exercises 31-34 to find the relationship between L and K that minimizes total cost.

31. Cobb and Douglas introduced the **idealized production function**

$$q = AL^{3/4}K^{1/4}$$

as a model of the interplay of Labor and Capital in the U.S. economy from 1889 to 1929. Find the values of L and K that minimize total cost and determine the **ratio** of labor to capital when total cost is minimized, given that $A = 0.8372$ during that time and that one unit of capital has the same cost as one unit of capital.

32. Erwin is contracted to produce 50 corner cabinets. He knows he will need help (labor), and he knows that he must expand his basement workshop (capital). If Erwin estimates his production will satisfy

$$50 = 0.4L^{0.8}K^{0.4}$$

then what amount of labor and capital will minimize the cost to produce the cabinets when labor is \$20 per hour and capital is \$10 per unit (A unit of capital in Erwin's case may be considered the amount to increase his workshop by the equivalent of one handtool.)?

33. A firm produces a commodity with a Cobb-Douglas production model of $q = 4L^{2/3}K^{1/3}$, where labor costs $w = \$15$ per manhour and capital costs $r = \$10$ per unit. What values of L and K minimize the cost of producing $q = 1000$ units of the commodity?

34. Suppose in exercise 33 that the firm has only \$10,000 to spend on producing the commodity. What values of L and K will maximize the output, q?

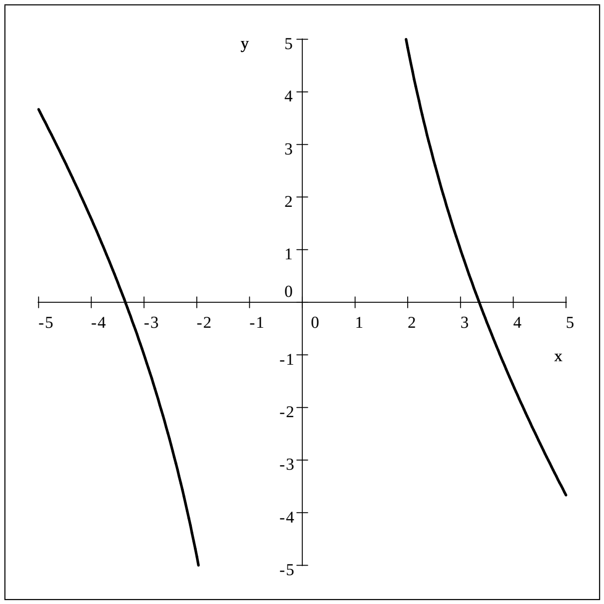
35. Find the extrema of $f(x, y, z) = x + yz$ subject to the constraints

$$x^2 + y^2 + z^2 = 1,\ z^2 = x^2 + y^2$$

36. Find the extrema of $f(x, y, z) = xyz$ subject to the constraints

$$x^2 + y^2 + z^2 = 1 \quad \textit{and} \quad xy + yz + zx = 1$$

37. Find the point P on the curve $4x^2 + 3xy = 45$ that is closest to the origin, and then show that the line from the origin through that point P is perpendicular to the tangent line to the curve at P. In light of the method of Lagrange multipliers, why would we expect this result? (Hint: you may prefer to eliminate λ using the method described between examples 5 and 6).



38. Find the point P on the curve $y = x^2 + x - 1.5$ that is closest to the origin, and then show that the line from the origin through that point P is perpendicular to the tangent line to the curve at p. In light of the method of Lagrange multipliers, why would we expect this result?

39. Find the maximum of

$$f(x, y) = e^{-x^2 - y^2}$$

subject to the constraint $x + y = 1$. Explain why we know that we have a maximum even though the constraint is not a closed curve.

40. Find the maximum of

$$f(x, y) = e^{-x^2 - y^2}$$

subject to the constraint $x^2 - y^2 = 1$. Explain why we know that we have a maximum even though the constraint is not a closed curve.

41. Show that ∇f is parallel to ∇g only if

$$\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial y} = 0 \tag{6}$$

Then use (6) to find the extrema of $f(x, y) = 3x + 4y$ subject to $x^2 + y^2 = 25$.

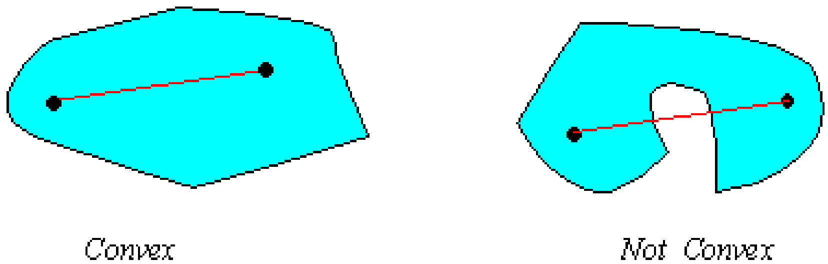
42. Use the method in problem 39 to find the extrema of $f(x, y) = xy$ subject to $x^2 + 2y^2 = 5$.

43. Write to Learn: Write a short essay in which you explain why if $g(x, y) = k$ is a smooth closed curve, then a continuous function $f(x, y)$ must attain its maximum and minimum values at solutions to the Lagrange multiplier problem. (Hint: if $\mathbf{r}(t) = \langle x(t), y(t) \rangle$, t in $[a, b]$, is a parametrization of $g(x, y) = k$, then finding the extrema of $f(x,y)$ subject to $g(x,y)=k$ is the same as finding the extrema of

$$z(t)=f(x(t),y(t))$$

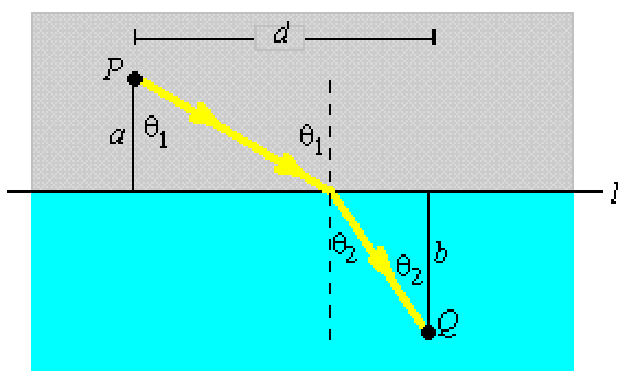
when $z(a)=z(b)$.)

44. Write to Learn: A region in the xy-plane is said to be *convex* if the line segment joining any two points in the region is also in the region.



Show that if R is a closed convex region and if $f(x,y)=mx+ny$ is a linear function (i.e., m,n constant), then the largest possible value of $f(x,y)$ over the region R must occur on the boundary of R. (Hint: let $g(x,y)=k$ be any line connecting two boundary points of R)

45. Write to Learn: Suppose that light travels from a point P with a constant speed v_1 in the medium above a horizontal line l and suppose that it travels to Q with a constant speed v_2 in the medium below l.



In a short essay, use Lagrange multipliers to explain why that the angles θ_1 and θ_2 that minimize the time required to travel from P to Q must satisfy

$$\frac{\sin(\theta_1)}{v_1}=\frac{\sin(\theta_2)}{v_2} \tag{7}$$

(Hint: Constraint is that horizontal distance traveled is equal to a fixed distance d).

Maple Extra: Create a worksheet which demonstrates Snell's law (7) using animation and through calculation of time of travel for different values of θ_1 and θ_2.