Computer Laboratory Manual: Calculus for the Life Sciences

Excel and Maple Edition

Joseph M. Mahaffy

March 2014

Contents

CHAPTER 1:

INTRODUCTION

This document includes a LaTeX version of many WeBWorK Laboratory problems created in the past. The problems are divided according to the mathematics that they use. The five categories are functions, discrete dynamical models, differentiation and optimization, differential equations, and integration. This manual begins with a brief introduction to basic graphing in Excel and Maple of simple functions and commands to use functions. Most computer applications in this lab manual are introduced with an example, then problems are presented to practice the computer techniques on interesting applications.

Excel is a spreadsheet software package developed by Microsoft in 1987, becoming quite dominant in the area. It is packaged in the Microsoft Office software, which is very common. Its primary strength is as a spreadsheet software for managing business accounts. However, it has many features that are useful to scientists, especially when managing large data sets. This manual will concentrate on using the data management and graphing abilities of Excel. Some of the numerical packages for fitting curves and finding parameters will be applied to modeling examples.

Maple is a Symbolic Algebra package. Maple was developed in 1980 at the University of Waterloo to become a more efficient program to its predecessor Macsyma. Most science, mathematics, and engineering students need to know some symbolic algebra program to handle complex modeling problems in the science of today. Maple's chief competitor is Mathematica, but at the level of this manual either language could be used with only minor variations in the syntax of the commands. Maple is continually upgraded and can manage a wide range of mathematics. This manual only provides some basics to introduce students to the program.

The Problems in this Lab Manual were originally developed for a computer laboratory component of a Calculus for the Life Sciences course. Students can see a more complete development of many of the Calculus and modeling ideas in this Lab Manual in the texts *Calculus: A Modeling Approach for the Life Sciences*[10, 9]. The original problems were designed to introduce students to modeling applications of Calculus and teach basic programming techniques. The primary computer applications taught were Excel, Maple, and Word. To provide additional variation in the problems, they were adapted to WeBWorK, using its ability to individualize the problems with random changes to the parameters. (The problems in this text include the WeBWorK file name for easy reference, though some modifications are made.) This text has one version of the WeBWorK problems including most of the blanks that the students input to WeBWorK. The problems also include parts that are meant to be written into a lab report. The lab report is designed to have students write complete paragraphs, explaining what was done, and to demonstrate their ability to create good graphs. The lab report portion of the problem will say, "In the Lab Report..."

CHAPTER 2:

INTRODUCTION TO EXCEL AND MAPLE

Keywords: Graphing, Excel, Maple

This chapter introduces the reader to Excel and Maple. Some basic commands are shown to use these programs for studying functions and mathematical models. The primary purpose of this chapter is to introduce the reader to making good graphs in Excel and demonstrate the algebraic power of Maple. The chapter centers around an example for graphing a line and a quadratic using both programs. Problems are presented to allow the reader to practice using Excel and Maple with simple functions. As the manual progresses, the reader will learn to build on these basic commands to produce more intricate graphs. There are tutorials for both Excel and Maple available through the program *Help* pages, which the novice should consider viewing. For a good overview of the **Excel** program the reader may want to take advantage of the *Getting Started* Help page, which has both text and video overviews of the program. This can be accessed by going to the *Main Menu* on a spreadsheet, then clicking on the tab *File*. The *File* tab has the option *Help*, which has the links to the *Getting Started* tutorials. Because of the popularity of Excel, there are many excellent references and videos available on the web for studying properties of this program. **Maple** also has an introduction called *Maple Tour*, which can be found under the *Help* tab of the *Main Menu* with the title *Take a Tour of Maple*. Maple is also used extensively, so many websites exist to provide additional learning tools. This manual will present a number of good references for additional reading.

1 BASICS IN EXCEL

This chapter introduces basic operations in Excel with the primary goal being the creation of good graphs for a Lab Report. Before discussing the graphics features in Excel, some basic spreadsheet operations are presented. The techniques are illustrated with an example of a line and a parabola.

Example 2.1 (Line and Parabola)**.** *This example considers linear and quadratic functions. Features of these functions are analyzed, and a graph of the functions with key points is created.*

The author has designed a WeBWorK problem to examine the linear and quadratic functions:

$$f(x) = 2\,x - 2.5 \qquad \text{and} \qquad g(x) = 5 - 3\,x - x^2. \tag{2.1}$$

In order to create a good graph in Excel, the important features of the graph need to be found. Of particular interest are x and y-intercepts, points of intersection, and the vertex of the quadratic. Later in this chapter (9) instructions are given to use the Excel package *Solver* to numerically find the x-intercepts and the points of intersection. Since this problem considers only linear and quadratic functions, basic techniques from algebra are applied to find the points of interest.

The linear function, $f(x)$, has a y-intercept, $y = -2.5$, and an x-intercept, $x = 1.25$. The quadratic formula gives the x-intercepts for $g(x)$, as

$$x = \frac{-3 \pm \sqrt{29}}{2},$$

and the y-intercept is clearly, $y = 5$ ($g(0) = 5$). By setting $f(x) = g(x)$, we obtain

$$x^2 + 5\,x - 7.5 = 0,$$

so the quadratic formula gives the x-values of intersection as

$$x = \frac{-5 \pm \sqrt{55}}{2}.$$

It follows that the two points of intersection are

$$(x_1, y_1) = \left(\frac{-5 - \sqrt{55}}{2}, -\frac{15}{2} - \sqrt{55} \right) \quad \text{and} \quad (x_2, y_2) = \left(\frac{-5 + \sqrt{55}}{2}, -\frac{15}{2} + \sqrt{55} \right).$$

Excel can be used as a simple calculator to help transform the algebraic expressions above into decimals.

Excel as a Calculator

Excel excels as a spreadsheet calculator. It rapidly computes formulas and performs basic computations very well. It tabulates results, so provides a basis for basic data analyses. The software has many built-in functions, which makes it a very good choice for simulating models or evaluating functions.

The numbers found by the quadratic formula above need to be converted to decimals to determine the range of interest for graphing our functions. Excel evaluates a formula in a cell whenever the formula begins with '=.' As an example, we show how to enter the x-intercept formulas. In any cell, enter:

```
= (-3 + SQRT(29))/2
```

After hitting *ENTER*, Excel returns 1.192582. Note that Excel understands *SQRT* as the square root function. (Excel is not case sensitive, so this can be typed in any combination of upper and lower case letters.) By enlarging or shrinking the width of the cell, more or fewer digits can be obtained. (The size of Excel cells are varied by clicking at the top between columns and moving the mouse to obtain the desired spacing.) Clearly, a similar formula finds the other x-intercept for the quadratic to be $x = -4.19258$.

Next we want to find the decimal values for the points of intersection. Suppose we select the cells $A3$ and $A4$ to enter the x-values of the points of intersection. In $A3$, we enter:

```
= (-5 - SQRT(55))/2
```

which yields, $x_1 = -6.2081$. A similar calculation gives $x_2 = 1.208099$ in $A4$. Since these are the points of intersection of $f(x)$ and $g(x)$, the y-values are obtained by either of these functions, using the x-values computed above. In Excel we find these y-values by evaluating either $f(x)$ or $g(x)$ at x_1 and x_2. This is done very easily by entering one function in $B3$ and having it call on the value in $A3$. Thus, using $f(x)$, we type in $B3$:

```
= 2*A3 - 2.5
```

and it gives the result $y_1 = -14.9162$. To obtain the value y_2, we use Excel's *pull down* or *updating* feature. This feature is central to Excel, so it can be accessed in multiple ways. The easiest way is to click on Cell $B3$. When it is highlighted, the user moves the cursor with the mouse to the lower right corner of $B3$ (marked with a small black square). The cursor changes from a broad white plus sign to a small '+' sign. At this point the user holds the left mouse button down and *pulls down* to $B4$. Upon releasing the mouse, the answer is *updated* in $B4$ to $y_2 = -0.0838$. Clicking on $B4$, the user sees this *updated* formula:

```
= 2*A4 - 2.5
```

This *pull down* function of Excel is critical to many applications. Below are listed some ways to execute the *pull down* function in Excel, and the reader needs to decide which method he or she prefers. In all of the cases below, we assume that the formula that needs updating is in a particular *source cell* and the *target cells*, which will be updated, are below this.

1. In the *source cell*, the cursor is moved to the lower right corner until a '+' sign appears, then with the left button of the mouse held down the user pulls down until the desired number of cells are filled with the updated formula.

2. Highlight the *source cell* and all the cells below that are meant to be filled with the updated formula. Go to the the *Main Menu* under the *Home* tab and select under the *Editing* options, the *cell icon* with the down pointing arrow and select *Down* or type *D*.

3. Highlight the *source cell* and *Copy* (keystrokes *ctrl-C*) this cell. Next highlight all the cells below this cell, which are meant to contain the updated formula, and *Paste* (keystrokes *ctrl-V*) to fill all the cells.

4. Related to the previous statement, the *source cell* is highlighted and the user *Right clicks* on the cell and selects *Copy*. Next the user highlights all cell to be filled with the updated formula, then *Right clicks* in this region and selects *Paste*, choosing the *icon* with the clipboard and a page with a folded corner.

Note that these options will often allow the user to fill and update cells in other directions, including the ability to simultaneously perform the operation on multiple columns or rows. This feature is critical to efficient use of Excel, so learn the technique that is most comfortable to you.

To further use Excel to help estimate the x-intercepts of $g(x)$ without the quadratic formula, we take advantage of the *pull down* feature above. Suppose that it is given or known that the x-intercepts are between $x = -10$ and $x = 10$. For this example, insert the label x in Cell $A1$ and the label $g(x)$ in Cell $B1$. It is a good practice to use the top row for labeling columns for future reference. The steps below will describe how to obtain the first x-intercept, and the reader can extend the method to find the other x-intercept.

In Cell $A2$, type '-10,' and in Cell $B2$, type '`= 5 - 3*A2 - A2^2`.' (You should see -65 appear in $B2$.) In Cell $A3$, type '`= A2 + 1`.' Now use Excel's *pull down* feature to fill cells $A4$ to $A22$. Then use the *pull down* feature to fill cells $B3$ to $B22$. At this point, we see that the the function $g(x)$ changes sign between $B7$ and $B8$ and again between $B13$ and $B14$, so the x-intercepts lie with $x \in [-5, -4]$ and $x \in [1, 2]$, giving very crude estimates of the x-intercepts.

The next step in the procedure is to insert '-5' in Cell $A2$ and change Cell $A3$ to '`= A2 + 0.1`.' Now we *pull down* $A3$ to $A12$, and we observe that $g(x)$ changes signs between $B10$ and $B11$, giving the x-intercept with $x \in [-4.2, -4.1]$. Repeat this process inserting '-4.2' in Cell $A2$ and changing Cell $A3$ to '`= A2 + 0.01`.' The procedure gives $x \in [-4.20, -4.19]$. Taking it another step, we find $x \in [-4.193, -4.192]$. This procedure can be repeated as often as necessary to get the desired accuracy, gaining one more significant figure each time. It also has to be repeated to find the second solution, $x \in [1, 2]$.

The reader should realize that this is not a very efficient way to obtain the x-intercept. However, it does illustrate how the *pull down* feature can be used to efficiently generate function evaluations very rapidly.

2 BASIC GRAPHING IN EXCEL

The main purpose of this chapter is creating a graph in Excel, which can be used in a Lab Report. Key features for creating good graphs in Excel 2010 are shown. This document will illustrate the essential steps for all graphs and several regular optional features.

A Graphing Template

Excel provides a very good package for producing quality graphics. This chapter of the Lab Manual concentrates on the function graphing abilities of Excel. This section shows how to produce a general purpose *graphing template*, which can be used for many functions used in the modeling problems. It is illustrated using Example 2.1.

Excel is a spreadsheet based software, so the functions need to be entered as tables of data. To create good smooth curves, always use 40-60 data points (exceptions being a straight line, requiring very few points, or complex curves, which require more points). Below are instructions to produce a general *Graphing Template*, which automatically has 50 data points and uses the variable x in a function. The reader should follow the directions below to create this *Graphing Template*, then *Save* the Excel file under a generic name to be used for graphing other functions. A copy of this *Graphing Template* can be renamed for a new problem, which will require minimal changes to obtain the graph of the new function.

To create the *Graphing Template*, open Excel. Use the first row to provide labels for the variables and functions. We insert the following labels, which will be used for Example 2.1 and more generally other functions. In Cell $A1$ insert x, then follow that with the labels $f(x)$ and $g(x)$ in Cells $B1$ and $C1$. We will want to be able to easily vary the domain for different functions, so enter the labels $firstx$ and $lastx$ in the Cells $E1$ and $G1$.

For Example 2.1, the calculations above show that all of the interesting points will be contained if the functions are graphed for $x \in [-7, 3]$. We enter -7 and 3 in the Cells $F1$ and $H1$, respectively, to provide the limits of the domain. In Cell $A2$, '`= F1`' is entered as the first point in the domain of interest. To obtain 50 evenly spaced points, in Cell $A3$ '`= A2 + ($H$1 - $F$1)/50`' is entered, then this formula is *pulled down* to Cell $A52$, which will be the same value as Cell $H1$. Note that the formula in Cell $A3$ has the expressions '`$H$1`' and '`$F$1`.' The '`$`' preceding the letter or number tells Excel to *freeze* this character when doing calculations, instead of updating the Cell as it is *pulled down*.

Excel has a valuable feature that allows the user to *name variables*. For this *Graphing Template*, we create the named variable x for the first column. To accomplish this, we highlight the first column using the mouse to click on A above the first column. Next we select the tab *Formulas* from the *Main Menu*. Under this tab, we find the entry *Create from Selection* and click on this. A new window will pop up, where it will already presume that you want the name of the variable from the *Top row*. By clicking on *OK*, the entries in Column A are now named x.

The next step is to enter the functions in Columns B and C. From Example 2.1, we enter the formulas '`= 2*x - 2.5`' and '`= 5 - 3*x - x^2`' in Cells $B2$ and $C2$, respectively. These formulas are *pulled down* to Cells $B52$ and $C52$ to complete the function evaluations across the domain of interest. The functions are readily graphed by highlighting all cells in the first three columns of the spreadsheet. Return to the *Main Menu* and select the tab *Insert*. Search for the option *Scatter*, then select the bottom left entry (connecting data points with line segments). At this point, the Excel spreadsheet should appear similar to the image in Fig. 2.1.

At this point the user should save this file (say *graph_tmpl.xlsx*) for future use with other functions. When the graph of a new function is needed, then the user can make a copy of this file (renaming it for the new problem). One opens the copy, then by changing the limits of the domain in Cells $F1$ and $H1$ to the new desired domain and changing Columns B and C to the new function or functions, which are to be graphed, the new graph is easily produced using the *Scatter* option under the *Insert* tab.

The instructions for the *Graphing Template* provide the basis for evaluating enough points to obtain a reasonable graph for many functions. The basic graph is shown in Fig. 2.1; however, this basic graph **must** be improved before it is inserted into a Lab Report. Excel has many features that allow the user to improve the appearance of the graphs.

Producing a Good Graph

The graph in Fig. 2.1 is not very attractive and lacks information. A graph must always have a title and labels for the axes. The curves need to be labeled through either direct labels or a legend, and finally the domain must be changed so that the functions span the domain. Below we give some details on how to create the graph shown in Fig. 2.2, which could be used in a Lab Report.

The process begins by clicking on the graph in the Excel spreadsheet as seen in Fig. 2.1. The *Main Menu* now has a category named *Chart Tools*. The middle tab of *Chart Tools* is *Layout*, which gives many

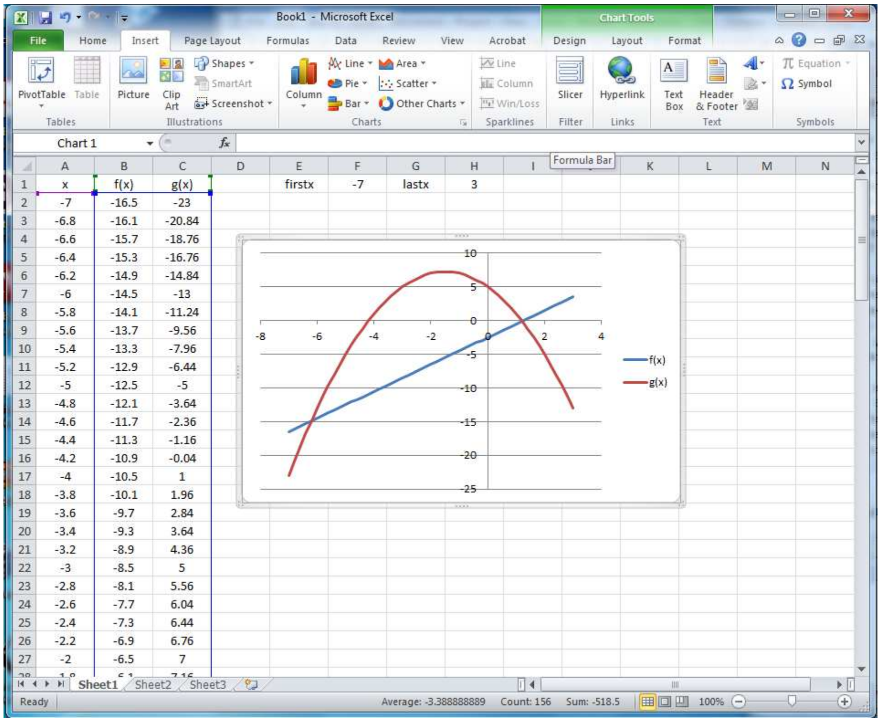

Figure 2.1: Screenshot of the Excel *Graphing Template* after graphing the two functions from Example 2.1.

important tools for formatting the graph. On the left, we select *Chart Title*, then choose the option *Above Chart*. By right clicking on the title, we can vary the *font*. In Fig. 2.2, the font chosen was *Times New Roman* in *Bold* with font size of 14 points.

The next step is to add *Axis Titles*. Select this option, then use both *Primary Horizontal Axis Title* and *Primary Vertical Axis Title* to create appropriate axis titles. Choose the *Rotated Title* for vertical alignment. Again by right clicking on the axis title, we can vary the *font*. In Fig. 2.2, the font chosen was *Times New Roman* in *Bold Italic* with font size of 12 points.

Excel has an odd feature that it always adds additional space on the horizontal axis. Since we want the graphed domain of the functions to match with x-axis labeling, the domain virtually always must be adjusted. Under *Layout* there is an *Axes* tab, where the user selects the box to adjust the *Primary Horizontal Axis*. At the bottom of the box are *More Primary Horizontal Axis Options ...*, which allow changing the domain of your function. Check the *fixed* option for both the *Minimum* and *Maximum* and insert the appropriate values for the domain of interest. In this example these are -7 and 3, respectively. Usually the default choice of Excel works fine for the *Primary Vertical Axis*. Excel automatically inserts *horizontal gridlines*; however, we want to also include *vertical gridlines*. These are easily inserted by going to the tab *Gridlines*, selecting *Primary Vertical Gridlines*, and choosing *Major Gridlines*.

In Fig. 2.2, the two points of intersection are shown with black circled markers and labeled with the appropriate values. We accomplish this by using several features in Excel's *Chart Tools*. We begin by inserting the values of the points of intersection in the spreadsheet. For example, put the x-values in Cells $E3$ and $E4$ and the corresponding y-values in the Cells $F3$ and $F4$. Next *right click* on the graph and a menu pops up, where the user selects the option *Select Data ...*. This pops up another window, *Select Data Source*, where the user clicks the button *Add*. A label can be given in the box *Series name:*, but we don't want one in Fig. 2.2. In the box labeled *Series X values:*, the user clicks, then returns to the spreadsheet and highlights the appropriate cells, which are Cells $E3$ and $E4$ for this example. Next the user goes to the box

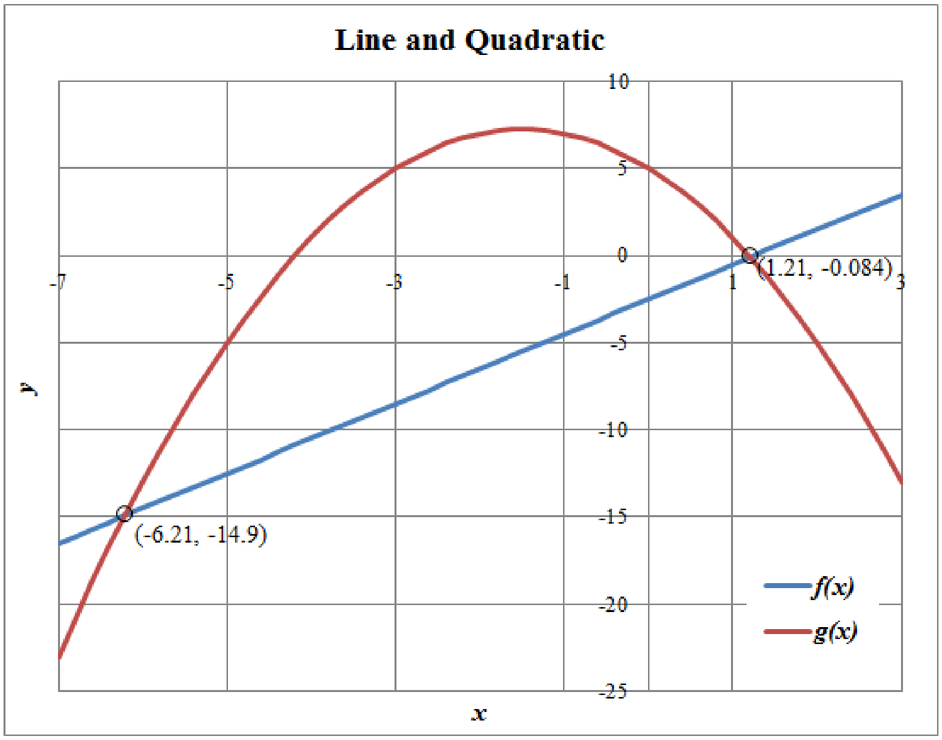

Figure 2.2: Final graph from the Excel *Graphing Template* showing the two functions from Example 2.1.

labeled *Series Y values:*, where the {1} is replaced by highlighting the appropriate cells, which are Cells $F3$ and $F4$ for this example. By clicking on *OK*, a new line appears between the points of intersection.

To change this green line segment into the points of intersection with the appropriate markers, the user moves the mouse to one of the points of intersection where Excel states that the point is *Series 3 Point ...* At this point, the user *right clicks* on the point and selects the option *Format Data Series ...* To obtain the *black open circle* seen in Fig. 2.2 there are several steps. First under *Marker Options*, we select *Built-in* and choose the *Type* matching a circle with a size of 7. Under *Marker Fill*, we select *No fill*. Under *Line Color*, we choose the *No line* option. Finally, under *Marker Line Color*, we select *Solid line* with *Color:* black.

To label the points of intersection, right click again on the black circle, and choose *Add Data Labels*. Another right click will bring up the option to *Format Data Labels*. The x and y values should be selected, and the *number* should be given to the appropriate decimal places - in this case, two. A left click is made on the *data label*, then the text is highlighted. Parentheses can be typed appropriately. The text is highlighted, and a right click allows one to select *Font ...*, which can be varied to match the style for the rest of the graph and for Fig. 2.2 is *Times New Roman* with *Regular* and a font size of 11 points.

Alternately, the points of intersection can be labeled with the *Text Box* option under the *Chart Tools* category *Layout*. The user selects this option, then clicks near one of the points of intersection. A text box appears, where one can type the appropriate label for this point of intersection. The text box is manipulated with the mouse to be resized and repositioned, as seen in Fig. 2.2. Finally, a left click is made on the *Text Box*, then the text is highlighted. A right click now allows one to select *Font ...*, which can be varied to match the style for the rest of the graph and for Fig. 2.2 is *Times New Roman* with *Regular* and a font size of 11 points.

We note that the legend has the new entry *Series 3*. We remove this legend entry by *left clicking* on the entry and *deleting* it with the *delete key*. To further improve the appearance of the legend, we *right click* on the legend and select *Font ...*, where we enter *Times New Roman* with *Bold Italic* and a font size of 12 points. To have the legend stand out against the background, we right click on the legend and select *Format Legend*. In this window we select the *Fill* option with the *Solid fill*, choosing the white *Fill Color*. This graph is now ready to be copied and pasted into a Lab Report document.

Optional Excel Techniques

Excel contains a number of powerful tools for performing calculations. Excel's *Solver* uses advanced numerical methods for solving equations. In this section, we introduce *Solver* to numerically find the x-intercepts and points of intersection, which were done algebraically above. Excel's *Solver* routine is not normally loaded with the standard installation of Excel, so to access this package the user goes to the *Main Menu* under the tab *File*. This tab has the choice *Options*, which has the further choice of *Add-Ins*. The *Solver Add-in* is highlighted, then the user selects *Go...* at the bottom and a new window pops up. Finally, the choice *Solver Add-in* is checked and *OK* is selected. This installs *Solver*, and it stays available for future use on the computer in which it is installed.

We return to Example 2.1 to let *Solver* find the x-intercepts of the parabola and the points of intersection for the line and the parabola. With a new spreadsheet, once again in the top row, insert labels x, $f(x)$, $g(x)$, and $f(x) - g(x)$ for the x values, the two functions, and the difference of the functions. In Cell $B2$, insert the equation of the line, '`= 2*A2 - 2.5`,' and in Cell $C2$, insert the equation of the parabola, '`= 5 - 3*A2 - A2^2`.' *Pull down* these equations to $B5$ and $C5$, since we are seeking 4 numerical results.

The graph shows x-intercepts near -4 and 1, so enter guesses of -4 in Cell $A2$ and 1 in Cell $A3$. *Solver* is found under *Data* in the *Main menu*. Click on *Solver* from the *Analysis* section (right side of menu) and a window pops up. In the window labeled *Set Objective:* either enter $C2$ or click on $C2$, the equation for the parabola. On the next line, select *Value Of:* and leave the value as 0. In the box labeled *By Changing Variable Cells:* either enter $A2$ or click on $A2$, the x value. It is **important** to uncheck the box *Make Unconstrained Variables Non-Negative*. It only remains for the user to click on the *Solve* button at the bottom, and the Cell $A2$ changes to a very good approximation of the x-intercept, yielding $x = -4.19258$. We repeat this process in *Row 3* to obtain a very good approximation of the x-intercept with $x = 1.192582$.

The points of intersection are found by entering in Cell $D4$, '`= B4 - C4`.' When the difference of the line and the parabola are 0, then the curves intersect. This formula is *pulled down* to Cell $D5$, then estimates of the x-values of the points of intersection are entered into Cells $A4$ and $A5$. The procedure with *Solver* in the previous paragraph is repeated with the primary difference being that the *Set Objective:* is now $D4$ and $D5$ with the *By Changing Variable Cells:* being $A4$ and $A5$. If we guess in $A4$ that $x = -6$, then *Solver* produces the point of intersection as $(-6.2081, -14.9162)$ seen in Cells $A4$ and either $B4$ or $C4$. Similarly, a guess of $x = 1$ in $A5$ has *Solver* yield the point of intersection as $(1.208099, -0.0838)$ seen in Cells $A5$ and either $B5$ or $C5$.

3 BASIC FUNCTIONS AND GRAPHING IN MAPLE

This section introduces a few basic commands in Maple, which allows entering functions and creating graphs. Maple is a very powerful Symbolic Algebra package with extensive abilities to handle complex mathematical problems symbolically. This lab manual is meant only as a brief introduction to Maple with the idea that users will become comfortable with basic commands and be able to solve increasingly complex problems in the future. This section shows how to enter functions and use these functions to solve equations and graph simple graphs.

Maple is an interactive line command language. Students may want to think of Maple as a very smart graphing calculator in early use. The syntax of some basic commands are shown below to make operation easy; however, note that Maple does have excellent *Help* available. The reader may want to delve deeper into what Maple can do by following the *Take a Tour of Maple* under the *Help* tab.

In this section, the problem given by Eqns. (2.1) is examined with Maple. The functions are input into Maple by typing the following commands:

```
f := x -> 2*x - 2.5
g := x -> 5 - 3*x - x^2
```

Maple will automatically modify these to appear as
$f := x \to 2 \cdot x - 2.5$
$g := x \to 5 - 3 \cdot x - x^2$
The functions $f(x)$ and $g(x)$ behave like functions, so by typing $g(2)$, Maple gives the answer -5.

The roots of $g(x)$ are found by the command:

```
solve(g(x) = 0,x)
```

which gives the answer

$$-\frac{3}{2} - \frac{1}{2}\sqrt{29}, -\frac{3}{2} + \frac{1}{2}\sqrt{29}.$$

To get a decimal answer one replaces *solve* with *fsolve*. The x-values of intersection are found and stored in a vector variable $xsoln$ by typing

```
xsoln := fsolve(f(x) = g(x),x)
```

Maple responds with the answers 1.208099244 and -6.208099244. The y-values are obtained by typing

```
f(xsoln[1]); g(xsoln[2])
```

The points of intersection found are

$$(x_1, y_1) = (-6.2081, -14.916) \qquad \text{and} \qquad (x_2, y_2) = (1.2081, -0.0838015).$$

The points above indicate a good range to graph. Maple has a very easy plot command to plot functions, and it is given by:

```
plot({f(x), g(x)}, x = -7 .. 2)
```

This command says to plot the set of functions $f(x)$ and $g(x)$ on the interval $x \in [-7, 2]$. Below is the plot that Maple produces. By right clicking on the graph, a menu appears that allows the user to edit the graph. For example, a title or axis label can be added, and the graph can be exported in a number of formats for use in a paper. Maple graphs are much harder to get the quality of output that is available with Excel, but the command above shows how quickly a sketch of a graph can be made in Maple. In this mode Maple acts as a very smart graphing calculator.

4 PROBLEMS

Below is the first problem of the Lab Manual. The problems have been created from a collection of WeBWorK problems, and the WeBWorK file name is provided, though some modifications are necessary. This Lab Manual includes blanks corresponding to the blanks in the WeBWorK problems, and every problem has a graphing portion that is meant to be written into a Lab Report. Most problems have an interesting biological example, and a complete list of the problems is provided in an appendix at the end of the Lab Manual with a brief summary of the problem.

The problems below allow the students to either use their algebra skills or employ the tools of Excel or Maple. These problems should provide a good introduction to Excel and Maple. The graphing uses easy modifications of the Excel instructions provided above.

Problem 2.1 (Two Lines)**.** *(Lab121_A1_lin_quad_graph.pg) This problem introduces graphing methods with a review of the algebra for two intersecting lines.*

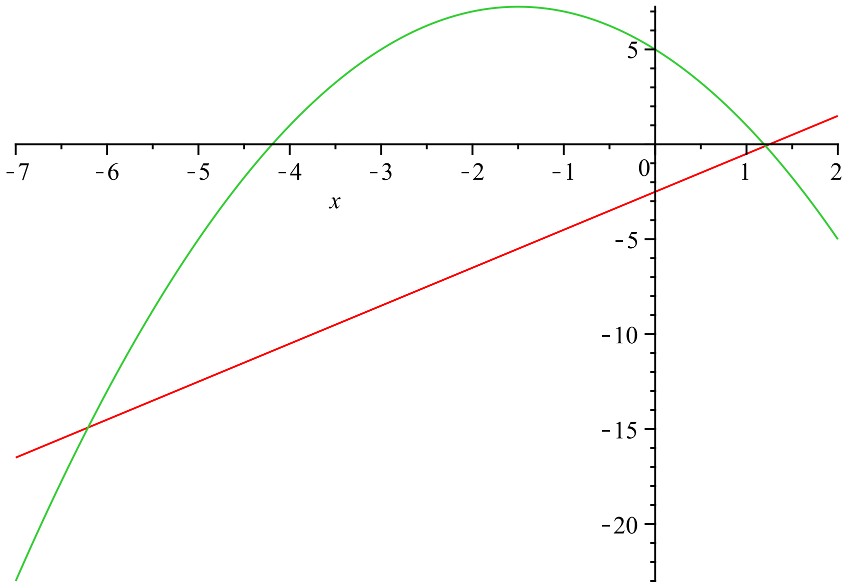

Figure 2.3: Graph of a line and quadratic from Maple.

a. You are given the following two lines:

$$y = 5x - 2 \qquad \text{and} \qquad y = -\frac{x}{2} + 3.$$

Find the slopes and the x and y-intercepts of these lines.

First Line:

Slope = ________ x-intercept = ________ y-intercept = ________

Second Line:

Slope = ________ x-intercept = ________ y-intercept = ________

Are these lines perpendicular? (Yes or No)

Find the point of intersection. $x =$ ________ $y =$ ________

b. In the Lab Report, create a graph with the two lines given above. Your graph needs to include a title and have the x and y axes labeled correctly. Adjust the scale of the graph so that the domain x and range y both go from -10 to 10, then expand the graph into a square. Label each line with its equation. Also, label the coordinates of the point of intersection on the graph. Write a sentence explaining how you determine if the lines are perpendicular or not. Write a short paragraph detailing how the point of intersection is found. Explain the process, not using the specific numbers for this problem.

Problem 2.2 (Line and Quadratic)**.** *(Lab121_A2_quad_line.pg) This problem studies the intersection of a line and a quadratic function.*

a. Consider the following functions, $f(x)$ and $g(x)$:

$$f(x) = 2 - 2.5x \quad \text{and} \quad g(x) = x^2 + 3x - 3.$$

Beginning with the line, find the x and y-intercepts. Also, determine the slope of the line.

x-intercept = ____________ y-intercept = ____________ Slope = ____________

For the parabola, find the x and y-intercepts. Also, determine the coordinates of the vertex.

x-intercepts = ____________ and ____________ y-intercept = ____________

Vertex = (____________, ____________).

Find the points of intersection of the two curves.

First point of intersection = (____________, ____________).

Second point of intersection = (____________, ____________).

b. In the Lab Report, create a graph with the two functions given above. Be sure that the domain for the graph includes all intercepts, points of intersection, and the vertex of the parabola. Do NOT label these points, but do be sure to label the respective curves with an appropriate identifying label. Label the axes and give a title to the graph. Write a brief paragraph about how to find the points of intersection. Explain the process, not using the specific numbers for the problem.

Problem 2.3 (Line, Quadratic, and Cubic)**.** *(Lab121_C1_lin_quad_cub.pg) This problem reviews a quadratic function and studies the intersection of a line and a cubic function.*

Maple is a powerful symbolic algebra computer package and is capable of solving cubic equations, which are extremely difficult to do by hand (so are rarely taught).

a. Consider the quadratic function:

$$g(x) = x^2 + 2x - 8.$$

Find the x and y-intercepts. Also, determine the coordinates of the vertex.

x-intercepts = ____________ and ____________ y-intercept = ____________

Vertex = (____________, ____________).

b. In the Lab Report, create a graph of the parabola given above. Be sure that the domain for the graph shows all intercepts and the vertex.

c. Now consider these functions, $f(x)$ and $g(x)$.

$$f(x) = -2x - 3.1 \qquad \text{and} \qquad g(x) = 1.2 + 2.8x - 2.5x^2 - x^3.$$

Find the x and y-intercepts and the slope for the line.

x-intercept = ____________ y-intercept = ____________ Slope = ____________

For the cubic equation, find the x and y-intercepts.

x-intercepts = ____________, ____________, and ____________ y-intercept = ____________

Find the points of intersection between $f(x)$ and $g(x)$.

Points of intersection = (____________, ____________), (____________, ____________), and (____________, ____________).

d. In the Lab Report, graph this line and cubic function. Choose a domain such that the graph clearly shows all the points identified above.

CHAPTER 3:

FUNCTION MODELS

Keywords: Linear model, linear least squares, exponential, logarithm, power law, allometric

This section contains the laboratory problems from WeBWorK that use basic functions for modeling. The first problems use the simplest of models, the linear models of the form:

$$y = ax + b,$$

with slope, a, and y-intercept, b. The model is found using a *linear least squares* best fit to the data, which can be found readily from web sources. This manual shows how computer programs can be applied to find this best fitting model.

1 ERROR ANALYSIS

When discussing models, it is very important to know how well a particular model represents the data that is being analyzed. There are a number of ways to measure how well data fit a model. This section presents important measures of error to compare models.

Least Squares Error

The most common unbiased error used to fit a model to data is the *least squares error*. Consider a set of n data points: $(x_1, y_1), (x_2, y_2), ...(x_i, y_i), ..., (x_n, y_n)$, where (x_i, y_i) is the i^{th} data point. A linear least squares best fit to these data considers a linear model

$$y(x) = ax + b,$$

where the slope, a, and the intercept, b, are chosen to best fit the data in some sense.

The least squares best fit minimizes the square of the error in the distance between the y_i values of the data points and the $y(x_i)$ value of the line, which depends on the selection of the slope, a, and the intercept, b. Fig. 3.1 provides an illustration of this linear model fitting data. Note the the least squares best fit method can be generalized to other functions with $y(x_i) = f(x_i)$, but fitting the parameters in $f(x)$ is more difficult for nonlinear functions.

Define the *signed error* or *actual error* between each of the data points and the line as

$$e_i = y_i - y(x_i) = y_i - (ax_i + b), i = 1, ...n,$$

and the *absolute error* between each of the data points and the line as

$$|e_i| = |y_i - y(x_i)| = |y_i - (ax_i + b)|, i = 1, ...n.$$

We can see that e_i varies as a and b vary.

Fig. 3.1 displays the graph showing these error measurements.

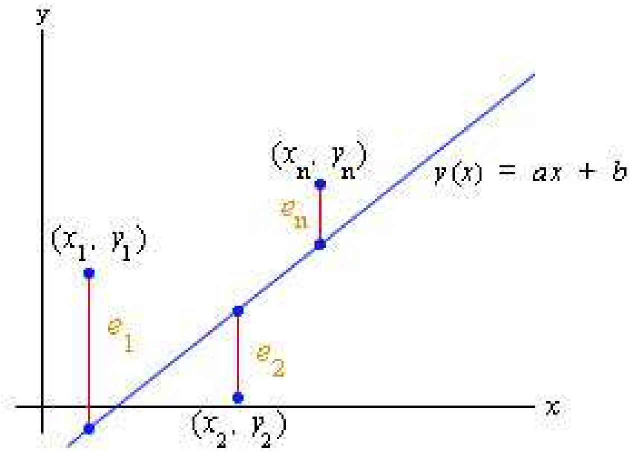

Figure 3.1: Graphic representation of the error measurements for the Least Squares Best Fit method.

The *least squares best fit* finds the minimum value of the function

$$J(a,b) = e_1^2 + e_2^2 + ... + e_n^2 = \sum_{i=1}^{n} e_i^2,$$

as the slope, a, and intercept, b, vary.

The function $J(a,b)$ sums all the squared errors between the linear model and the data points. The slope, a, and the intercept, b, of the linear model are varied until the function, $J(a,b)$, achieves its smallest value. The unique line that produces this minimum $J(a,b)$ is the least squares best fit or best linear model for the given data set. There are formulas for the best values of a and b, which are given by:

General Linear Least Squares Best Fit Method

Define the mean of the x values of the data points as

$$\bar{x} = \frac{x_1 + x_2 + ... + x_i + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

The value for the slope of the line that best fits the data is given by

$$a = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

With the slope computed, the intercept is found from the formula

$$b = \frac{1}{n}\sum_{i=1}^{n} y_i - a\bar{x} = \bar{y} - a\bar{x}.$$

These formulas are available in many computer programs and calculators, so the user only needs the basic program commands to obtain the best fitting slope and intercept for a linear model through a set of data.

Error Computations

There are several standard computations of error, *actual*, *absolute*, *relative*, and *percent* errors. The actual and absolute error between a data point and a model were defined above. The most common computations of error are *relative error* and *percent error*. Let X_e be an *experimental measurement* or the measurement, which is considered the *worst value*, and X_t be the *theoretical value* or the measurement, which is considered the *best value*. The problems in this lab manual will usually let the reader know which value is considered *best* or *worst*. Often the models are being tested, so the model will be considered the *worst value*.

The *relative error* is the difference between the experimental value and the theoretical value **divided** by the theoretical value, so

$$\text{Relative Error}= \frac{X_e - X_t}{X_t}.$$

The *percent error* is closely related to the *relative error*, except that the value is multiplied by 100% to change the fractional value to a percent, so

$$\text{Percent Error}= \frac{X_e - X_t}{X_t} \times 100\%.$$

Another way to think of the relative error is that it is the difference between the least accurate and most accurate values divided by the most accurate value:

$$\text{Percent Error} = \frac{Worst\ Value - Best\ Value}{Best\ Value} \times 100\%.$$

We generally divide by the most accurate value, which is usually the one provided by the data tables when comparing models.

2 LINEAR MODELS AND DATA

This laboratory manual is designed to introduce students to basic modeling from a variety of examples and apply the computer programs Excel and Maple to graph and solve computational aspects of the problems. The simplest model is the *linear model*. The examples and problems in this section introduce linear models and show how computer programs can produce graphs for these models, showing the data and the model. The programs use the least squares best fit analysis described above. The techniques are demonstrated with an example of growing children and the Lambert-Beer law. Three problems are presented on crickets chirping, urea excretion of hummingbirds, and growth of yeast.

Example 3.1 (Juvenile Height)**.** *The rate of growth in height of children is almost linear for a range of ages.*

Age	1	3	5	7	9	11	13
Height (cm)	75	92	108	121	130	142	155

Table 3.1: The average juvenile height of American children as a function of age [7].

Table 3.1 gives data for the height of the average American child with respect to age [7]. This example shows how to use Excel and Maple to find the best fitting linear model and produce a graph. The model is given by

$$h(a) = ma + b,$$

where the best slope, m, and intercept, b are found by our programs. The least square error analysis showed the formulas for finding these values, which the reader may want to verify by hand. This least square error is computed.

Excel for Data and Linear Least Squares Fit

This example introduces techniques to use Excel to find the best linear model through the set of data for juvenile height. We demonstrate how to accurately copy data into an Excel spreadsheet and graph the data. Excel has a built in program for finding the linear least squares best fit to the data called *Trendline*. This routine gives the best linear model for juvenile height as a function of age.

The first step to finding the best linear model is to accurately copy the data set into an Excel spreadsheet. Table 3.1 is presented above in this document, which has a PDF format. We highlight the elements of the Table 3.1 from this document and *Copy* (*crtl C*) them. The next step is to *Paste* (*ctrl V*) the table into a new Excel spreadsheet. This process enters all the data into Column *A* in the spreadsheet. From the Excel *Main menu*, select *Data*, then choose the item labeled *Text to Columns*. This will pop up a box where the user selects the choice of *Delimited*, then clicks on the button *Next*. In Step 2 of this process, the user checks the boxes for *Tab* and *Space* and clicks on the box *Finish*. The result should be the data separated into distinct cells of accurately copied data from which we can proceed. The data is not aligned because of the two cells with `height` and `(cm)`. We choose to *right click* on the cell with `(cm)`, then select *Delete...* and allow Excel to *Shift cells left*.

This example has a small data set, where we only seek the best *linear model*, so the user can highlight the data, proceed to the *Main menu* under the tab *Insert*, and choose the graph with the *Scatter* plot. Since the graph is showing data, the user selects the default setting of *data points* (top left). Often the data set is larger, so it is preferred that the data set is in columns. This is also the preferred direction for computing the sum of square errors. To achieve this with the horizontal data set, the user highlights the data and copies it (*ctrl C*). Proceed to an open cell below the data and *right click* on that cell. A menu pops up, and we left click on the category *Paste Special...* A new box pops up, and we check the box for *Transpose* and click on *OK*. This transforms the data set from one in rows into one in columns. This data set can be highlighted and made into a *scatter plot* as described above.

The graph, as always, **must** be improved in appearance, so follow the steps outlined in the previous chapter to create a *Chart Title*, *Axes Titles*, and *Vertical Gridlines*. Since there is only one model and data set, then the legend is removed by clicking on the legend label and pressing the *delete* key. For this graph, we modified the *font* to *Times New Roman* with sizes 14 and 12 points for the title and axes, respectively. Also, the data points were changed from the default setting.

After finishing this initial phase, there are two good methods of getting the best fitting linear model. After clicking on the graph, the most straightforward technique is to go to the *Main Menu* under *Chart Tools* and click on *Layout*. From there we find the tab for *Trendline*. Under this tab we select the *More Trendline Options ...* and a box appears with *Format Trendline*. Alternately, we can simply *right click* on any of the data points and select the option to *Add Trendline ...*, then the box appears with *Format Trendline*.

In the box *Format Trendline*, we allow the default option *Linear* for our choice of *Trendline/Regression Type*. We note that Excel's *Trendline* linear fit is using the formulas given earlier in this chapter. Since we are interested in knowing the mathematical model produced by Excel's *Trendline*, we **always** check the box to *Display Equation on Chart*. For this data set the data goes from $a = 1$ to 13, and we want to view the model over the range $a \in [0, 16]$. Thus, the *Trendline* needs extending, which is accomplished with the option *Forecast*. In this case we *forecast Forward:* 3 units and *Backward:* 1 unit. As is always the case with Excel, the *Axes Properties* have to be adjusted to fix the range with $a \in [0, 16]$.

When the *Trendline* box is closed, the graph will have the formula for the best model on the graph with the *linear model* shown through the data. In this example, the formula shown has all the accuracy that is required for proceeding with answering other questions posed by the problem. Often this is **not** the case, so we need to obtain more *significant figures*. The easiest way to obtain more significant figures from the *Trendline* model is right click on the formula box in the graph. Select *Format Trendline Label ...* and the

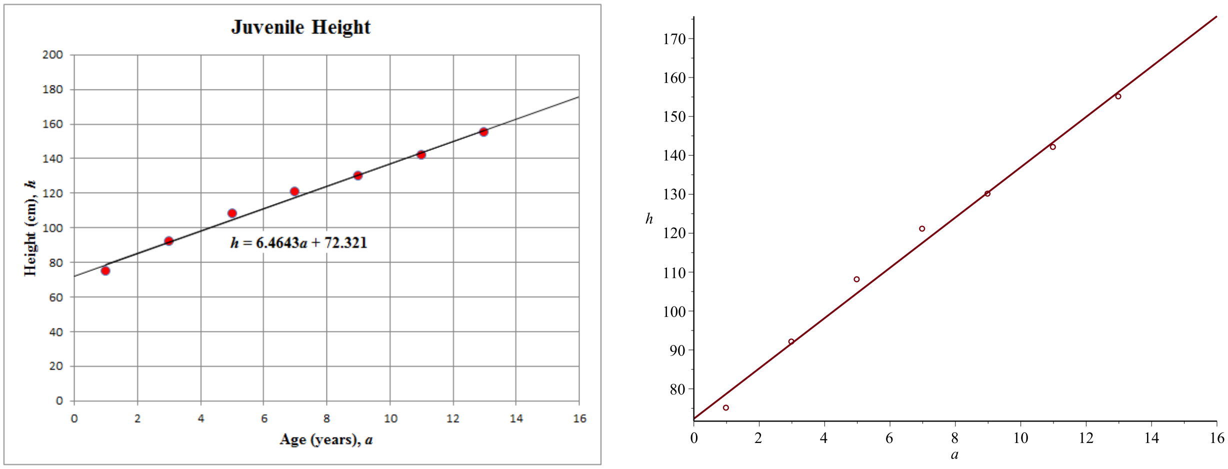

Figure 3.2: Graphs of child height data and the least squares best fit linear model from Excel (left) and Maple (right).

default box says *General*. In this box we select *Scientific*, using at least 4 *Decimal places*. (This will give the equation formula with 5 significant figures.) This **must** be done before anything is modified in the formula box or Excel will not allow the formula to change.

The final step for creating a good looking graph is to move the *Trendline* equation to a more visible part of the graph. We *right click* on the equation box, then make a number of changes to the font, making it larger and using *Times New Roman*. We also use the appropriate variables, a, and h, which are in *Bold italic*. To make the equation stand out, we *Format Trendline Label*, selecting the *Fill/Solid fill* option with a *white background*. The resulting graph is shown on the left in Fig. 3.2 and can be used in a Lab Report.

This procedure gives the best slope, $m = 6.4643$, and h-intercept, $b = 72.321$ for the juvenile height model. Thus,

$$h(a) = 6.4643\,a + 72.321,$$

is the best fitting model to the data in Table 3.1. To find the sum of square errors, this model is compared against data in the table. Assume that the data are arranged in *Columns* A and B with $A1$ and $B1$ holding the labels *Age* and *Height*, respectively. In Cells $E1$ and $E2$, we insert the labels m and b, then give their values 6.4643 and 72.321, respectively, in Cells $F1$ and $F2$. These variables are *named* and assigned their values by highlighting the four cells in *Columns* E and F, then going to the *Main menu* under *Formulas* and selecting *Create from Selection*. After accepting the box that pops up, which has *Create names from values in the: Left column*, then these variables can be used in any formula.

In $C1$, we insert the label '`model`.' In $C2$, the model formula is written, '`= m*A2 + b`,' and this formula is *pulled down* to Cell $C8$ (the end of the data). This produces the heights, h, that the model predicts for all the ages in the data set. In $D1$, we insert the label '`SSE`' for the *sum of square errors*. In $D2$, the square error between the data height and the model is computed by typing '`= (B2 - C2)^2`.' This formula is *pulled down* to Cell $D8$. In Cell $D9$, we want to compute the *sum of square errors*, so we click on this cell. Under the *Main menu* tab *Formulas*, there is the option $\sum$ *AutoSum*, which when clicked automatically enters '`=SUM(D2:D8)`,' which computes the sum of all the entries from $D2$ to $D8$. The result is the sum of square errors:

$$J(m, b) = 41.5714,$$

which is the least sum of square errors between these data and the best fitting linear model. If the user tries to input any other numbers (to 5 significant figures) into the values for m or b, then the resulting value of $J(m, b)$ will be larger.

This example gives interesting information about child growth based on the linear model. The slope gives a good estimate of how much a child grows each year. The slope indicates that between 1 and 13 years of age, the average child grows about 6.5 cm/yr. The h-intercept is 72.3 cm, which gives a rough estimate of the length of a newborn child. However, the h-intercept is outside the range of the data, so it provides a poor estimate for birth length.

Maple for Linear Least Squares Fit

Maple can be used to find the linear least squares best fitting model to the juvenile height example above. Maple is organized into distinct packages to allow more efficient computing. The *polynomial fit* program is in the *Statistics* package, while techniques to graph both points and lines use the *plots* package.

To find the best linear least squares model to the juvenile height data, the following Maple commands are issued:

```
with(Statistics):                  # Invoke Statistics package
ad := Vector([1, 3, 5, 7, 9, 11, 13]):              # Age data
hd := Vector([75, 92, 108, 121, 130, 142, 155]):   # Height data
p := PolynomialFit(1, ad, hd)      # Store Coefficients for Best Fit in p
```

The coefficients found with Maple are clearly going to have the same values as the ones found with Excel. The graph should have data plotted as points and the linear model as a continuous curve. To have two different types of curves, the following Maple commands are issued:

```
with(plots):                        # Invoke plots package
P1 := plot(ad, hd, style = point, symbol = circle):   # Plots the data
P2 := plot(p[2]*a+p[1], a = 0 .. 16, h):              # Plots the best model
display({P1, P2})                   # Displays both the data and model
```

Note that the hashtag, #, is used for Maple comments. For plotting graphs in Maple, one usually invokes the plot command to observe the graph, then stores the plot in a variable, which in this case is P1 and P2. Note that whenever a colon ends a line in Maple, then Maple suppresses the output. The resulting graph from the commands above is displayed on the right in Fig. 3.2.

The problems below provide applications of the *linear least squares* best fit to some biological data. These problems allow the student to practice the techniques given above in the juvenile height example. Fitting linear models is very common in scientific studies, so these techniques are important to learn.

3 PROBLEMS USING LINEAR MODELS

This lab manual includes many data sets, which will be modeled in various ways. Below is an example of some data on crickets chirping from Bessey and Bessey [1], which are modeled with a *linear model*.

Example 3.2 (Cricket Model)**.** *Experiments have shown that the rate at which a cricket chirps is approximately a linear relationship with the temperature. This example takes a set of cricket data and graphs it with the best linear model fitting the data.*

The Bessey brothers collected a very large set of data of snowy tree crickets (*Oecanthulus fultoni*) from Lincoln, NE in 1897. They carefully measured the temperature where the crickets were chirping and the rate of chirps/min. Below is a table of a small set of their measurements.

As noted in the Example 3.1, the first step to creating an accurate model is entering the data into Excel correctly. Details of this process are given in Example 3.1. This example already has its data in column form, which is better for managing the data. Briefly, the user *highlights* all the table entries from a PDF file and copies (*ctrl C*) the data. On a new Excel spreadsheet with the *Cell A*1 highlighted, the user pastes

chirps/min	°F	chirps/min	°F	chirps/min	°F
77	54.2	123	67.5	173	76.3
85	58.5	136	69.2	168	77.2
95	59.5	146	71	162	78.2
101	61.5	144	73.5	183	79
112	64.8	158	75	208	80.9
122	66.5	172	75	195	83

Table 3.2: A collection of data from Bessey and Bessey [1] on cricket chirping rate as a function of temperature (°F).

(*ctrl V*) the data. Under the *Data* tab of the *Main menu*, the *Text to Columns* feature is used to transform the data into a useable form by choosing *Delimited* with *Space*. With Table 3.2, the data is rearranged on the Excel spreadsheet using *Cut* (*ctrl Z*) and *Paste* (*crtl V*) until all the chirping rates are in *Column A* and the temperatures are in *Column B*.

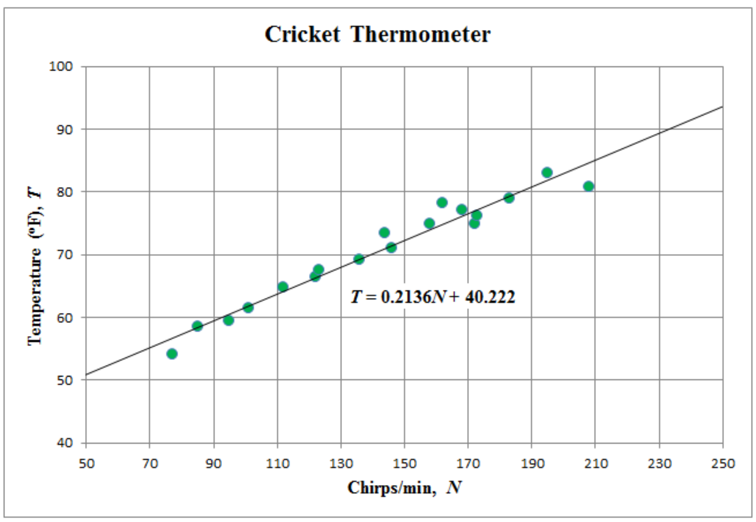

Figure 3.3: Graph showing the data and linear model for rate of crickets chirping for different temperatures.

We follow the same directions as in Example 3.1 for graphing the data and adding a *Linear Trendline* model to produce the graph in Fig. 3.3. The linear least squares best fitting model satisfies the equation:

$$T = 0.2136\,N + 40.222.$$

We note that Fig. 3.3 uses several different features to create the final graph. The graph required the restriction of the range of the graph, as Excel defaults to 0 for the lower bound, which would not make much sense with the crickets only living above freezing. For the choice of font, we selected the *Times New Roman* and made the variables *Bold italic*. The way that we produced the °F was to type '`oF`,' then highlight the '`o`' and after *right clicking*, select *Font...* and check the *Superscript* option. Once again, we see that Fig. 3.3 presents a graph that is very good for including in a Lab Report.

Problem 3.1 (Cricket Thermometer)**.** *(Lab121_A3_cricket.pg) Data indicate that the rate at which crickets are chirping is a linear relationship with temperature.*

For many years people have recognized a relationship between the temperature and the rate at which crickets are chirping. The folk method of determining the temperature in degrees Fahrenheit is to count the number of chirps in a minute and divide by 4, then add 40. In 1897, A. E. Dolbear [5] noted that "crickets in a field [chirp] synchronously, keeping time as if led by the wand of a conductor." In his paper, he appears to be the first person to write down a formula in a scientific publication giving a linear relationship for the temperature based on the chirp rate of crickets.

a. The classic folk "cricket thermometer," formalized by Dolbear, satisfies the linear relationship:

$$T = \frac{N}{4} + 40,$$

where T and N were the temperature and the number of chirps/minute, respectively. Determine the slope and T-intercept for this line.

Slope = __________ T-intercept = __________

The Bessey brothers [1] later made careful measurements and did a linear least squares best fit to their data and obtained the linear relationship

$$T = 0.213\,N + 40.4.$$

Determine the slope and T-intercept for this line.

Slope = __________ T-intercept = __________

b. Below are recordings of four crickets chirping at different temperatures. In this question, you time the number of chirps/minute of four crickets. By clicking on the different sound tracks for the crickets, the hyperlink goes to a recording for each of the four crickets.

Sound for Cricket A **Sound for Cricket B**

Sound for Cricket C **Sound for Cricket D**

(These sounds can be obtained through the hyperlinks:
http://www-rohan.sdsu.edu/∼jmahaffy/courses/s00a/math121/labs/laba/X
where
X = **w585_21.html** X = **w585_5.html** X = **w585_25.html** X = **w585_30.html**)

Below are a list of chirping rates for crickets, and you select the answer that is closest to the chirping rate for the sounds that you measured above.

______ Cricket A	A. 221 chirps/min
	B. 69 chirps/min
______ Cricket B	C. 99 chirps/min
	D. 39 chirps/min
______ Cricket C	E. 131 chirps/min
	F. 276 chirps/min
______ Cricket D	G. 161 chirps/min
	H. 199 chirps/min

c. In the Lab Report, create a graph of each of the models (one graph with both models). Show clearly the data points that are gathered in Part b. Write a short paragraph on how the data was collected to find the correct answers.

d. The cricket thermometer model is a linear model. Each of the variables has physical units. For example, T has units °F. Use the units of °F, chirps, and min.

What are the units for the coefficient representing the slope? ____________________

What are the units for the coefficient representing the T-intercept? ____________________

e. Errors are introduced when counting the number of chirps/min. Suppose that the error for counting the chirps/min for Cricket B is ±5 chirps/min.

Find the range of temperatures predicted by the Dolbear model

_____________ $\leq T \leq$ _____________

Also, find the range of temperatures predicted by the Bessey model

_____________ $\leq T \leq$ _____________

Suppose that the error for counting the chirps/min for Cricket D is ±10 chirps/min.

Find the range of temperatures predicted by the Dolbear model

_____________ $\leq T \leq$ _____________

Also, find the range of temperatures predicted by the Bessey model

_____________ $\leq T \leq$ _____________

f. Below is a table of a subset of the Bessey brother data [1].

chirps/min	°F	chirps/min	°F
77	54.2	145	73.5
85	58.5	150	74.8
95	59.5	158	75
105	61	174	76.3
101	61.5	168	77.2
111	65.5	189	78
123	67.5	162	78.2
128	68	183	79
132	69	191	80.9
141	69.5	198	82.5

Find the equation of the least squares best fit line through these data.

$T =$ ________________________________

Find the percent error between slope and T-intercept of this least squares best fit line and the Bessey model given in Part a. (Assume the model in Part a is the better value.)

Percent Error in slope = __________% Percent Error in T-intercept = __________%

g. In the Lab Report, create a graph showing the data above with its least squares best fit line. On the graph include both the Dolbear and Bessey brother models from Part a. Write a brief paragraph discussing the accuracies of the models from the lab experience, what are the major sources of error (list at least two). How much agreement is there between the different models have in predicting the temperature.

In the problem below, the WeBWorK version notes that all answers should be kept to 5 or 6 significant figures. This is required because the WeBWorK default numerical answer accuracy is 0.1% relative error. Excel's *Linear Trendline* will only fit the data with 3 or 4 significant figures, so solving this problem requires extending the significant figures of the *linear model* produced by Excel's *Trendline*, which is done by *right clicking* on the formula, selecting *Format Trendline Label ...*, and choosing *Scientific* with at least 4 *Decimal* places.

Problem 3.2 (Hummingbirds and Urea). *(Lab121_B2_absorb.pg) Spectrophotometry is used to quantitatively measure urea concentration, and this quantitative analysis gives information about animal excretion patterns.*

This problem examines some physiological data from the laboratory of Professor Carol Beuchat at San Diego State University [2]. Animals have evolved different mechanisms for excreting waste nitrogen. The principle means of excreting nitrogen are uric acid, urea, and ammonia. Unfortunately, the latter two are toxic so require larger volumes of water for excretion. Uric acid uses less water, but it requires more energy (ATP) to produce. Thus, animals must weigh their needs of water versus energy when selecting a means of excretion.

a. Here we only examine the amount of urea excreted. First, a standard is run to determine the absorbance at 570 nm as a function of the concentration of urea. (This is a standard technique using spectrophotometry.) The data are listed below:

Urea (mg/dl)	Absorbance	Urea (mg/dl)	Absorbance
2	0.139	70	2.17
10	0.3	100	3.117
20	0.624	130	3.826
40	1.236	150	4.545
50	1.51	200	6.149

Find the best straight line through the data, where

$$A = mu + b,$$

is the straight line describing absorbance, A, as a function of the concentration of urea, u, determining the slope, m, and intercept, b.

$m =$ __________ $b =$ __________

Find the expected absorbance for a sample containing 59 mg/dl of urea

$A =$ __________

and a sample containing 184 mg/dl of urea.

$A =$ __________

b. In the Lab Report, create a graph showing both the data and this linear model. Describe how well the linear model fits the data.

c. Use the formula for the best fitting line in Part a to find the predicted value of absorbance, A, for the concentrations of urea, u, at 50, 100, and 200.

$A(50) =$ __________ $A(100) =$ __________ $A(200) =$ __________

Determine the absolute and percent errors for the concentrations of urea, u, at 50, 100, and 200. In this case, assume that the absorbance given by the best fitting line is the best value.

When $u = 50$, the absolute error = ____________ and percent error = ____________%

When $u = 100$, the absolute error = ____________ and percent error = ____________%

When $u = 200$, the absolute error = ____________ and percent error = ____________%

Compute the sum of square errors, $J(m, b)$, that least squares function minimized to obtain the best fitting line to the data.

$J(m, b) =$ ____________

d. In practice, one uses the spectrograph to measure the absorbance, and use the relationship between the two to calculate the urea levels. In order to do this you must now solve for u as a function of A, the inverse function. That is find a function $f(A)$ such that

$$u = f(A)$$

$u =$ ____________ $A+$ ____________

e. In Professor Beuchat's laboratory they found that the urine from a hummingbird kept at 10°C had an absorbance of 0.138.

Concentration of urea = ____________mg/dl.

When the hummingbird was kept at 20°C, the absorbance for a urine sample was 0.206.

Concentration of urea = ____________mg/dl.

When the hummingbird was kept at 40°C the absorbance was 0.29.

Concentration of urea = ____________mg/dl.

f. In the Lab Report, give an explanation of these results with regard to either energy or water conservation by these hummingbirds?

g. We would like to see if the analysis of urine samples tells us about other species. The table below lists different animals and the corresponding absorbances measured.

Animal	Absorbance
Chicken	3.145
Duck (Fresh Water)	0.442
Duck (Salt Water)	0.768
Frog	0.266
Turtle	1.082
Tortoise	6.908

Concentration of urea for the chicken = ____________mg/dl.

Concentration of urea for the duck (fresh water) = ____________mg/dl.

Concentration of urea for the duck (salt water) = ____________mg/dl.

Concentration of urea for the frog = ____________mg/dl.

Concentration of urea for the turtle = ____________ mg/dl.

Concentration of urea for the tortoise = ____________ mg/dl.

A hummingbird gets its energy from nectar, which is high in water content.
Which animal has an excretion concentration most similar to a hummingbird? __________________

h. In the Lab Report, write a paragraph discussing why the one animal's excretion concentration is similar to a hummingbird's excretion. Also discuss why you would see the different concentrations of urea for the different animals in the table above. Is there a pattern between the different animals here, and can you offer some explanations?

Additional Commands for Function Fitting

There are examples where a general polynomial fit is not appropriate. The most common case occurs when the model must necessarily pass through the *origin*. This removes the constant term from a polynomial. Finding a model of this type can be readily done in both Excel and Maple.

In Excel, a mathematical model passing through the origin is very easy. We follow the same procedure as given before to invoke *Trendline*, selecting the desired model, which at this point has only been the *Linear* model. To obtain the model passing through the origin, we simply add the option *Set Intercept* = 0.0. For a *linear model*, this results in a model of the form $y = mx$.

Maple has the ability to fit more general forms of models, including one that passes through the *origin*. Other cases include situations when there is special symmetry, so the polynomial has only even or odd terms. Maple has the function *LinearFit* in its Statistics package, which can easily find these special forms of a polynomial. The *LinearFit* function takes the form:

```
LinearFit(flst,X,Y,v)
```

where *flst* is the list of functions being used, X and Y are the x and y data entries, and v is the variable in the functions.

Example 3.3 (Lambert-Beer Law)**.** *The Lambert-Beer law for absorbance of light by a spectrophotometer is a linear relationship, which can have the form*

$$c = mA,$$

where c is the concentration of the sample, A is absorbance, and m is the slope that must be determined from experiments.

The ion dichromate forms an orange/yellow solution, which has a maximum absorbance at 350 nm and is often used in oxidation/reduction reactions. Below is a table of data from known samples.

A	0.12	0.32	0.50	0.66
c (mM)	0.05	0.14	0.21	0.30

Table 3.3: Absorbance, A, readings from samples of known concentrations, c of dichromate.

Use the data in Table 3.3 to find the best fit for parameter m in the Lambert-Beer law, then determine the concentration of dichromate in an unknown sample with $A = 0.45$.

Solution: The solution using Excel is readily found by entering the data into a new spreadsheet with the data for A in *Column A* and the data for c in *Column B*. These data are plotted with *Scatter* plot, and

Trendline is applied to the data. The *Trendline* options that are checked are *Linear*, *Set Intercept* = 0.0, and *Display Equation on chart.* The result is seen in Fig. 3.4 with the best model given by

$$c = 0.44093\,A.$$

Note that additional accuracy is obtained in this solution by using the *Format Trendline label ...* option.

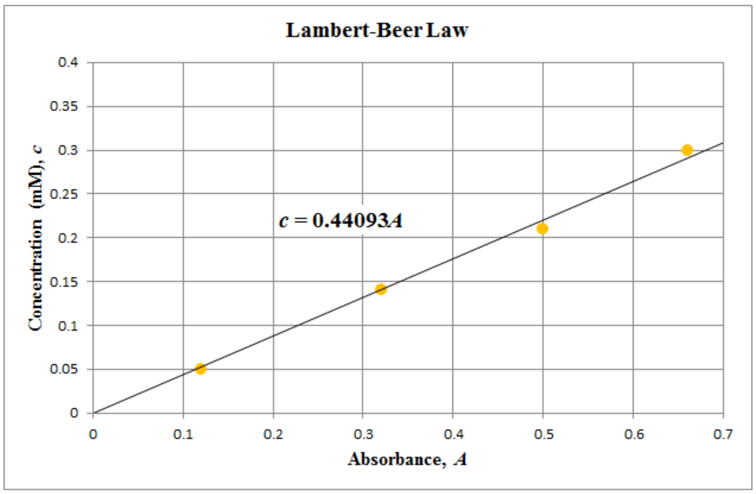

Figure 3.4: Graph showing the data and Lambert-Beer law model for dichromate.

Alternately, a solution to this problem is found with the following set of Maple commands:

```
with(Statistics):
X := Vector([0.12,0.32,0.50,0.66]):
Y := Vector([0.05,0.14,0.21,0.30]):
LinearFit([A], X, Y, A)
```

These commands open the Maple Statistics package, store the data in the vectors X and Y, and apply a model with only the linear term A. The resulting best model given by Maple is

$$c = 0.440927218344965\,A.$$

It follows that an unknown sample with $A = 0.45$ gives $c = 0.1984172482$.

Since the Lambert-Beer law has a simple form for the linear function, the sum of square errors is easily written from the data.

$$J(m) = (0.05 - 0.12m)^2 + (0.14 - 0.32m)^2 + (0.21 - 0.5m)^2 + (0.30 - 0.66m)^2,$$

which can be easily input into Maple and simplified with the following commands:

```
J := m -> (0.05-0.12*m)^2+(0.14-0.32*m)^2+(0.21-0.5*m)^2+(0.30-0.66*m)^2
simplify(J(m))
```

producing

$$J(m) = 0.1562 - 0.7076\,m + 0.8024\,m^2.$$

Maple's symbolic algebra capability reduces the quadratic function to standard form for easy evaluation.

Problem 3.3 (Yeast Growth). *(Lab121_C3_yeast_growth.pg) The growth rate of yeast at low densities is linear, which satisfies the conditions for Malthusian growth.*

In 1913, Carlson[3] studied the growth of a culture of yeast, *Saccharomyces cerevisiae*. Over time this culture levels off, but its initial growth is exponential or Malthusian. A Malthusian growth model is given by

$$P_{n+1} = P_n + rP_n.$$

Simply put, the population at the next time $(n+1)$ is equal to the population at the current time (P_n) plus some growth term, which is simply proportional (r) to the current population. Thus, we have a growth function

$$g(P) = rP.$$

Below is a table from Carlson's data showing the population, P, and the rate of growth, $g(P)$, at that particular population

Population	Growth/hr
18,600	10,800
29,500	18,200
48,200	24,600
71,800	48,100

a. Use the data above to find the best straight line passing through the origin. (This is a special case of the linear least squares best fit, with the line having a slope and a zero intercept.) What is the slope of the line that best fits through the data?

Slope $r =$ ____________

Find the sum of square errors with this model and the data.

Sum of Square Errors = ____________

b. In the Lab Report, create a graph of the data and the best fitting linear model. Briefly describe how well the line fits the data.

c. For the linear model (passing through the origin) given above, it is easy to find the sum of squares function. Consider a data point $(P_i, g(P_i))$. The absolute error between this data point and our model is given by

$$e_i = |g(P_i) - rP_i|.$$

Thus, $e_1 = |10800 - 18600r|$. Similarly, you can find e_2, e_3, and e_4. The sum of squares function is given by

$$J(r) = e_1^2 + e_2^2 + e_3^2 + e_4^2.$$

Find the expression for the quadratic function of the slope of the model, r (in simplest form).

$$J(r) = \_\_\_\_\_\_ r^2 + \_\_\_\_\_\_ r + \_\_\_\_\_\_ .$$

Find the coordinates of the vertex.

$$(r_v, J(r_v)) = (\_\_\_\_\_\_, \_\_\_\_\_\_)$$

d. In the Lab Report, create a graph of $J(r)$ for $r \in [0.3, 0.8]$. Compare the value of the vertex of the parabola, r_v, and the slope of the best fitting line through the origin found by Excel's *Trendline* or Maple's *LinearFit*. Also, compare the sum of square errors and the value of $J(r_v)$.

e. From the best model, find the growth for a population of 100,000 yeast.

Growth for 100,000 yeast $g(100,000) =$ ____________

Determine the population of another culture of yeast, given that their growth rate is measured to be 75,000 yeast/hour.

For $g(P) = 75,000$, $P =$ ____________

4 EXPONENTIAL AND LOGARITHMIC FUNCTIONS

The exponential and logarithmic functions are very important in a number of modeling applications. In Calculus it is natural to use the exponential function to the base e. Similarly, this is the natural logarithmic base. Most computer languages also rely on e as the natural base. For computer languages we generally write e^x in the following manner:

```
exp(x)
```

Thus, in either Maple or Excel, we have the following:

```
exp(1)
        2.718281828
```

For Maple, we need to type `evalf(exp(1))` or `exp(1.)` to obtain the decimal value. The Maple command `evalf(·)` is used to force Maple to produce decimal answers. Note that neither language recognizes e as a character alone. Maple is case sensitive, so the expressions must be typed exactly as presented here. However, Excel is **not** case sensitive, so any combination of upper and lower case letters give the same result.

For the natural logarithm, one generally writes $\log(x)$ in a computer language. However, many computer languages recognize the form $\ln(x)$, also. Maple recognizes both forms, while Excel only recognizes `ln(x)`. Excel recognizes $\log(x)$ as the logarithm base 10. Thus, to obtain $\ln(2)$, we type

```
log(2.)
        0.6931471806
```

in Maple or

```
= ln(2)
        0.6931471806
```

in Excel.

Below is a problem to explore the asymptotic size of exponential and logarithmic functions compared to power functions.

Problem 3.4 (Exponential, Logarithmic, and Power Functions)**.** *(Lab121_E1_exp_ln_pwr.pg) This problem explores the exponential, logarithmic, and power functions.*

This problem is intended to compare the exponential function to high order polynomials and compare the logarithmic function to fractional powers of x.

a. Consider the functions:

$$f(x) = 0.9e^{1.1x} \qquad \text{and} \qquad g(x) = 3.1x^8.$$

The domain for $f(x)$ is ________________ and for $g(x)$ is ________________.

Find all points of intersection (x and y values).

Points of intersection = (__________, __________), (__________, __________),

and (__________, __________).

Which of the following are statements are appropriate for $f(x)$?

A. This function has a Vertical Asymptote at $x = 0$.
B. This function has a Horizontal Asymptote as $x \to -\infty$.
C. This function has a Vertical Asymptote at some $x \neq 0$.
D. This function has a Horizontal Asymptote as $x \to \infty$.
E. None of the Above

If $f(x)$ has a Horizontal Asymptote or a Vertical Asymptote with $x \neq 0$, then give the x or y value of the asymptote or type DOES NOT APPLY.

____

Which of the following are statements are appropriate for $g(x)$?

A. This function has a Horizontal Asymptote as $x \to -\infty$.
B. This function has a Horizontal Asymptote as $x \to \infty$.
C. This function has a Vertical Asymptote at some $x \neq 0$.
D. This function has a Vertical Asymptote at $x = 0$.
E. None of the Above.

If $g(x)$ has a Horizontal Asymptote or a Vertical Asymptote with $x \neq 0$, then give the x or y value of the asymptote or type DOES NOT APPLY .

____

b. In the Lab Report, create two graphs for $f(x)$ and $g(x)$. For the first graph, choose a symmetric domain ($x \in [-A, A]$) that clearly shows the first two points of intersection, *i.e.*, take a value of A that is slightly larger than either of the x-values (in absolute value) for the first two points of intersection. For the second graph choose a larger symmetric domain that includes the third point of intersection. Note that this point of intersection has a very large y value, so the other two points of intersection are not very visible. Briefly describe the appearance of the graphs of each of these functions. Briefly discuss how any vertical or horizontal asymptotes are found. Also, describe how the points of intersection are found. Which function dominates (has larger y-values) for large values of x?

c. Now consider the functions:

$$h(x) = 1.2\ln(x) \qquad \text{and} \qquad k(x) = 1.2x^{1/7}.$$

The domain for $h(x)$ is ______________ and for $k(x)$ is ______________.

Find all points of intersection (x and y values).

Points of intersection = (__________, __________) and (__________, __________)

Which of the following statements are appropriate for $h(x)$?

A. This function has a Horizontal Asymptote as $x \to \infty$.
B. This function has a Vertical Asymptote at $x = 0$.
C. This function has a Horizontal Asymptote as $x \to -\infty$.
D. This function has a Vertical Asymptote at some $x \neq 0$.
E. None of the Above

If $h(x)$ has a Horizontal Asymptote or a Vertical Asymptote with $x \neq 0$, then give the x or y value of the asymptote or type DOES NOT APPLY.

____

Which of the following statements are appropriate for $k(x)$?

A. This function has a Vertical Asymptote at $x = 0$.
B. This function has a Vertical Asymptote at some $x \neq 0$.
C. This function has a Horizontal Asymptote as $x \to \infty$.
D. This function has a Horizontal Asymptote as $x \to -\infty$.
E. None of the Above.

If $k(x)$ has a Horizontal Asymptote or a Vertical Asymptote with $x \neq 0$, then give the x or y value of the asymptote or type DOES NOT APPLY.

____

d. In the Lab Report, create two graphs for $h(x)$ and $k(x)$. For the first graph, choose the domain $x \in [0, 10]$. For the second graph choose a much larger domain, $x \in [0, A]$, that includes the second point of intersection. Note that this point of intersection has a very large x value, which again masks the first point of intersection. Briefly describe the appearance of the graphs of each of these functions. Briefly discuss how any vertical or horizontal asymptotes are found. Also, describe how the points of intersection are found. Which function dominates (has larger y-values) for large values of x?

5 ALLOMETRIC MODELING

Many modeling problems are nonlinear. One very important class fits a model in the form of an *allometric* or *power law* model. These models have the form

$$y = kx^a. \tag{3.1}$$

By taking logarithms of this equation and using the properties of logarithms, Eqn. (3.1) can be written in the following form

$$\ln(y) = a\ln(x) + \ln(k).$$

By defining the variables $Y = \ln(y)$, $X = \ln(x)$, and $K = \ln(k)$, then this equation has the form

$$Y = aX + K,$$

which is a straight line with slope a and intercept K. The techniques from before can be used to find the best linear least squares best fit through the logarithm of the data to find the best *allometric model*. This gives an easy method for finding a and K or equivalently k.

Excel's *Trendline* has the option *Power*, which gives the best fitting power law model using the algorithm above. The user should be aware that the Excel *Trendline Power* fit is actually finding the linear least squares fit to the logarithms of the data. In Maple, the user computes the logarithms of the data and employs the *linear least squares* best fit to the logarithms of the data in a manner very similar to the one shown earlier in Example 3.1.

The technique of using the *linear least squares fit* to the logarithms of the data gives the easiest way to fit an allometric model, since virtually all computer applications have linear least squares fitting algorithms. However, this method is finding the least squares best fit to the *logarithms of the data* and **not** the least squares best fit directly from the data. The latter method requires a *nonlinear least squares* best fit of the data. Though the *nonlinear least squares best fit* is technically a better unbiased fit to the data, it is much more challenging to find. In particular, it usually requires a reasonable initial guess of the parameter values. Below we show how to use Excel to find this *nonlinear least squares fit* to data.

Example 3.4 (Cumulative AIDS Cases)**.** *In the early years the numbers of AIDS cases grew rapidly, and the cumulative cases of AIDS could be modeled by an allometric model of the form*

$$c = kt^r,$$

where t is the number of years after 1980, c is cumulative number of AIDS cases, and k and r are parameters that fit the data.

Table 3.4 shows the data for the cumulative number of AIDS cases in the U. S. over the time period of 1981-1992. During these early years the disease was not understood, and its spread was very rapid.

Year	Cumulative AIDS Cases	Year	Cumulative AIDS Cases	Year	Cumulative AIDS Cases
1981	97	1985	15,242	1989	120,612
1982	709	1986	29,944	1990	161,711
1983	2,698	1987	52,902	1991	206,247
1984	6,928	1988	83,903	1992	257,085

Table 3.4: Data of the Cumulative AIDS (in thousands of cases) throughout the U. S. from 1981 until 1992.[4]

The data clearly cannot be fit well with a linear model, so one option is to use an allometric model. If we define the variable t to be the number of years after 1980 and c to be the cumulative number of AIDS cases, then an appropriate allometric model is given by:

$$c = kt^r,$$

where the parameters k and r are found from the data above. With the use of logarithms this model can be written:

$$\ln(c) = \ln(k) + r\ln(t).$$

If we let $C = \ln(c)$, $K = \ln(k)$, and $T = \ln(t)$, then the above model is written:

$$C = K + rT,$$

which is just a linear equation. Thus, the logarithms of the data can be fit to a straight line, as was done before, to give the allometric model.

Excel for Power Law Fit

The data from Table 3.4 is entered into an Excel spreadsheet and rearranged to form two columns (A and B) with the first being the *Year* and the second being the *Cumulative AIDS Cases*. A column (B) is inserted between these by *right clicking* on the B in the second column and selecting *Insert*. This column is labeled t, and its entries are generated by typing '`= A2 - 1980`' and *pulling down* to the end of the data. Columns B and C are highlighted, and the user *Inserts* a *Scatter* plot. Excel's *Trendline* is invoked from either the *Main menu* or by *right clicking* on a data point. This time in the *Format Trendline* window, the user checks the *Power* and *Display Equation on chart* options. More digits for the formula are obtained by *right clicking* on the equation and selecting *Format Trendline Label...* The window that opens allows *Number* and *Scientific* to be selected, and for accuracy required of WeBWorK *Decimal places:* 4 is chosen. Appropriate graph titles and axis labels are added and modified, and the equation label can be modified to show the correct variables and look better. Fig. 3.5 on the left shows this graph.

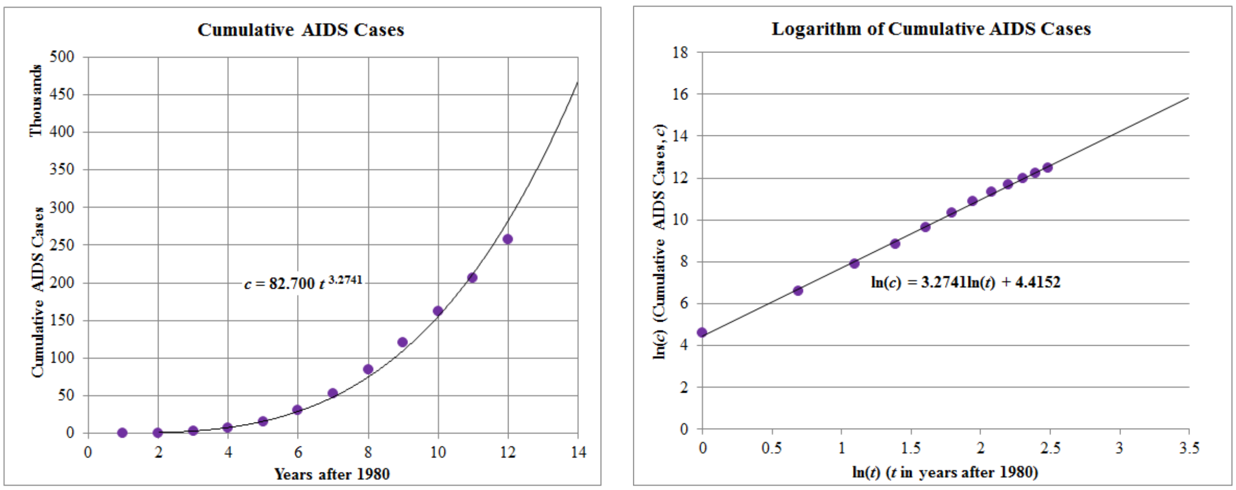

Figure 3.5: The left graph shows cumulative AIDS cases with the Excel *Trendline Power* law fit. The right graph shows the logarithms of the data with the Excel *Trendline linear* fit.

The model is entered in Column D and *pulled down* to the end of the data. Column E is used to compute the sum of square errors by entering '`= (C2 - D2)^2`' and *pulling down* this formula. Below the last cell, the user clicks on the $\sum$ symbol from either the *Home* or *Formulas* tabs to yield '`= SUM(E2:E13)`' with this sum of square errors being 932,060,425. This best *allometric model* is given by:

$$c = 82.700\, t^{3.2741}.$$

Fig. 3.5 on the right shows the linear least squares best fit to the logarithms of the data. In Columns G and H enter '`= ln(B2)`' and '`= ln(C2)`,' respectively, and *pull down* to the end of the data. These columns are highlighted, and the user once again *Inserts* a *Scatter* plot. This time Excel's *Trendline* with a *Linear* option is applied to give the linear least squares best fit. The graph is improved with the usual techniques. We make a couple of observations about this graph that are relevant to an *allometric model*. First, it is easy to see that the data appear to lie in a line, so an *allometric model* is appropriate whenever the graph of the logarithms of the data look almost linear. Second, we observe the formula of the *Linear* fit and see that the slope of the line (3.2741) agrees with the exponent of the *Power law* fit. Also, if we exponentiate the value of the intercept from the *Linear* fit ('`= EXP(4.4152)`'), then the coefficient of the *Power law* (82.700) is obtained.

The *power law* fit usually fails to give the *least sum of square errors* between the model and the data. To find this *Nonlinear Least Squares* best fitting model we use Excel's *Solver*, which was introduced in Chapter 2. For the sake of clarity, assume that we are working with the Excel spreadsheet used above for the Cumulative AIDS cases, where Column A has *Years* 1981-1992, Column B is the years after 1980, Column C has data for the Cumulative AIDS cases, Column D has the *allometric model* from *Trendline's Power law* fit, and Column E has the sum of square errors between the data and the Power law model. Insert labels k

and r in Cells $I2$ and $I3$ and give them the values from the *allometric model* in Cells $J2$ and $J3$. Highlight the four Cells $I2:J3$ and from the *Main menu* tab *Formulas* select *Create from Selection: Left column* to *name* the variables k and r. (Note that the variable r becomes $r\_$, because r is reserved for rows.) Label Columns F and G with *Nonlin Model* and *SSE*, respectively. In $F2$, enter '`= k*B2^r_`' and *pull down* to the end of the data ($F13$). (Note that one can accurately enter the formula by typing the symbols, '`=`,' '`*`,' '`^`,' and clicking on the locations of the variables, $J2$, $B2$, $J3$.) In $G2$, enter '`= (C2 - F2)^2`' and *pull down* to the end of the data. Below the last cell, click on the $\sum$ symbol from either the *Home* or *Formulas* tabs, which gives '`= SUM(G2:G13)`.'

The next step is to invoke Excel's *Solver* to minimize the *sum of square errors* in $G14$. From the *Main menu* under the tab *Data*, we find *Solver* on the right side with the category *Analysis*. By clicking on this, a window appears and we *Set Objective:* to $G14$ and check the box to *Min*. Finally, click in the box *By Changing Variable Cells:* and highlight the Cells $J2$ and $J3$. After clicking on the *Solve* button, the *sum of square errors* is reduced to 210,542,992, which is less than a quarter the *sum of square errors* computed by the *Power law model*, which used the best *linear fit* to the logarithms of the data. The result gives $k = 213.025$ and $r = 2.8630$, which gives the best *allometric model* as

$$c = 213.025\, t^{\,2.8630}.$$

It is clear that the two techniques give fairly different models.

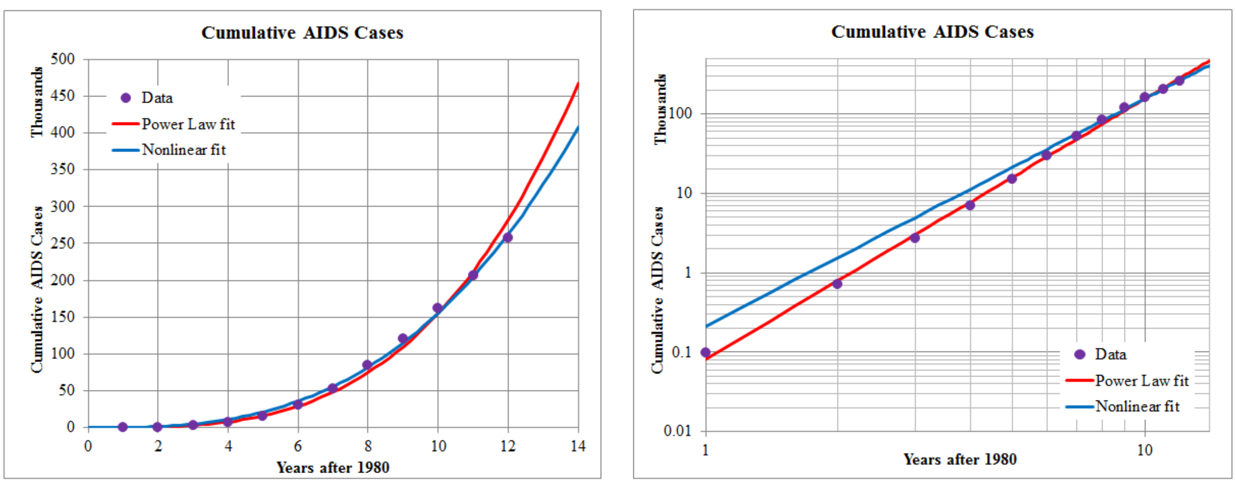

Figure 3.6: The left graph shows cumulative AIDS cases with the Excel *Trendline Power* law fit and the best nonlinear least squares fit. The right graph shows the same graph, but with logarithmic axis scales.

There are only twelve points in the data set, and graphing the models should include approximately 50 points. Choose three successive blank columns and label them t, *Power model*, and *Nonlinear model*. Suppose that the label t is in Cell $L1$. In Cell $L2$, enter 0, then in Cell $L3$, enter '`= L2 + 0.2`.' *Pull down* $L3$ to $L72$, where $t = 14$. In Columns M and N, enter the two versions of the *allometric model*. Columns L, M, and N are highlighted, and a *Scatter plot* is created. After the usual improvements are made to the graph, the left graph in Fig. 3.6 is produced. To produce the graph on the right in Fig. 3.6, we copy the graph we produced on the left and *Paste Special*, using the selection *Microsoft Office Drawing Object*. Next we either *right click* on one of the axes or use the *Main* menu Chart Tools under *Layout*, selecting either *Primary Vertical* or *Horizontal Axis* and choosing the option *Show Axis with Log Scale*. A warning will appear that *Negative or zero values cannot be plotted correctly on log charts...*, which can be ignored. Since the gridlines are fairly sparse, we have Excel also include the *Major & Minor Gridlines*. With minor adjustments the *log-log graph* of Fig. 3.6 is produced. On this graph the data appear to lie on a straight line, which again indicates that an *allometric model* is a good choice for fitting the data.

In Fig. 3.6, both models fit the data quite well. However, the *Trendline Power law* fit to the data does a poorer job fitting the larger t values. Thus, the *Nonlinear Least Squares* fit model is better at projecting future trends for cumulative AIDS cases. It is often a subjective decision as to which model is better. In *allometric modeling*, the most important parameter is the exponent, r, and for the two fits in Fig. 3.6 these differ by about 15%, which is not too bad for biological models.

Problem 3.5 (Dog Measurement Study). *(Lab121_D3_dog_study.pg) Allometric models are used to study the relationship between length, weight, and surface area of several dogs.*

Allometric modeling provides a method of understanding physical relationships between different methods of measurement. This problem examines the relationship of the length of a dog to its weight and surface area. Physiologically, it is important to note that all dogs have a basic shape that one recognizes as a dog. This is a type of *self-similarity.* A collection of dogs was measured and weighed, producing the Table 3.5 of data[12].

Length (cm)	Body Weight (gm)	Surface Area (cm^2)
48	3460	2245
63	5150	3250
72	5350	3865
76	9980	5000
97	17220	7900
101	26070	8995
105	33260	10535

Table 3.5: Data on several dogs measuring their length from nose to anus, their body weight, and surface area.

a. The first model examines Weight, w, as a function of the Length, u. Use the allometric modeling technique of fitting the logarithms of the data (Excel's *Trendline*) to find the best model fitting the data for a model of the form:

$$w = ku^a.$$

The best fit *Trendline* coefficients, k and a, are given by

$k =$ __________ and $a =$ __________

Determine the sum of square errors between the weight model and the data.

Sum of Square Errors = __________

If we assume that this model gives the ideal weight for a dog, then a dog with a length of 72 cm should weigh

Weight of 72 cm dog = __________gm.

Assuming this model gives the best value, compute the percent error that the 72 cm dog in the table varies from the ideal dog in weight.

Percent Error (weight) for 72 cm dog = __________%.

Suppose a dog has a length of 101 cm, then according to the model it should weigh

Weight of 101 cm dog = ____________gm.

Again, assuming this model gives the best value, compute the percent error that the 101 cm dog in the table varies from the ideal dog in weight.

Percent Error (weight) for 101 cm dog = ____________%.

Suppose you have a dog that weighs 14460 gm. According to the model, what would you predict would be its length?

Length of 14460 gm dog = ____________cm.

b. In the Lab Report, create a plot of this best fitting curve with the data. What are the appropriate units for the coefficient k? Write a brief paragraph to explain why the coefficient a has the value it does when fitting the weight data for the dogs.

c. Repeat the process in Part a for the surface area data, s, as a function of length, u. In this case, the model satisfies the allometric model:

$$s = ku^a,$$

where the coefficients k and a differ from Part a. The best fit *Trendline* coefficients, k and a, are given by

$k =$ ____________ and $a =$ ____________

Determine the sum of square errors between the surface area model and the data.

Sum of Square Errors = ____________.

If we assume that this model gives the ideal surface area for a dog, then a dog with a length of 72 cm should have a surface area

Surface area of 72 cm dog = ____________cm^2

Assuming this model gives the best value, compute the percent error that the 72 cm dog in the table varies from the ideal dog in surface area.

Percent Error (surface area) for 72 cm dog = ____________%.

Suppose a dog has a length of 101 cm, then according to the model it should have a surface area

Surface area of 101 cm dog = ____________cm^2.

Again, assuming this model gives the best value, compute the percent error that the 101 cm dog in the table varies from the ideal dog in surface area.

Percent Error (surface area) for 101 cm dog = ____________%.

Suppose you have a dog that has a surface area 6945 cm^2. According to the model, what would you predict would be its length?

Length of 6945 cm^2 dog = ____________cm.

d. In the Lab Report, create a plot of this best fitting curve with the data. What are the appropriate units for the coefficient k? Write a brief paragraph to explain why the coefficient a has the value it does when fitting the surface area data for the dogs.

e. Repeat the process in Part a for finding Weight, w, as a function of the Length, u. However, use the allometric modeling technique of *nonlinear least squares* to fit the data to the model:

$$w = ku^a.$$

The best *nonlinear* fit coefficients, k and a, and sum of square errors are given by

$k =$ ____________, $a =$ ____________, and Sum of Square Errors = ____________

Assume that this model gives the ideal weight (best value) for a dog. Find the weight of a dog with a length of 72 cm and compute the percent error that the 72 cm dog in the table varies from the ideal dog in weight.

Weight of 72 cm dog = ____________gm Percent Error (weight) = ____________%.

Find the weight of a dog with a length of 101 cm and compute the percent error that the 101 cm dog in the table varies from the ideal dog in weight.

Weight of 101 cm dog = ____________gm Percent Error (weight) = ____________%.

Suppose you have a dog that weighs 14460 gm. According to the model, what would you predict would be its length?

Length of 14460 gm dog = ____________cm.

f. Repeat the process in Part c for finding the surface area, s, as a function of the Length, u. However, use the allometric modeling technique of *nonlinear least squares* to fit the data to the model:

$$s = ku^a,$$

where the coefficients k and a differ from Part e. The best *nonlinear* fit coefficients, k and a, and sum of square errors are given by

$k =$ ____________, $a =$ ____________, and Sum of Square Errors = ____________

Assume that this model gives the ideal surface area (best value) for a dog. Find the surface area of a dog with a length of 72 cm and compute the percent error that the 72 cm dog in the table varies from the ideal dog in surface area.

Surface area of 72 cm dog = ____________gm Percent Error (weight) = ____________%.

Find the surface area of a dog with a length of 101 cm and compute the percent error that the 101 cm dog in the table varies from the ideal dog in surface area.

Surface area of 101 cm dog = ____________gm Percent Error (weight) = ____________%.

Suppose you have a dog that surface area 6945 cm^2. According to the model, what would you predict would be its length?

Length of 6945 cm^2 dog = ____________cm.

g. In the Lab Report add the *nonlinear least squares* fitting allometric models to the graphs created in Parts b and d. Write a brief paragraph comparing the graphs from the *logarithmic fit* to the data and the *nonlinear least squares* fit to these allometric models.

Problem 3.6 (Island Biodiversity)**.** *(Lab121_E2_biodiversity.pg) An allometric model is fitted to data on herpetofauna to study biodiversity on islands of the Caribbean.*

Currently there is a debate on the importance of preserving large tracts of land to maintain *biodiversity.* Many of the arguments for setting aside large tracts are based on studies of biodiversity on islands. In this problem an *allometric model* is used to determine the number of species of *herpetofauna* (amphibians and reptiles) as a function of island area for the given Caribbean islands. Table 3.6 provides data on biodiversity of herpetofauna on Caribbean islands[12].

Island	Area (km^2)	Species
Redunda	2.4	3
Saba	13	7
Montserrat	86.8	10
Jamaica	11,434	40
Hispaniola	76,490	91

Table 3.6: Data on the biodiversity of herpetofauna on Caribbean islands based on area.

a. Let N be the number of species and A be the area of the island. Use the *allometric modeling* technique of fitting the logarithms of the data (Excels *Trendline*) to find the best model fitting the data for a model. The *power law* expression relating the number of species to the area of the island is given by

$$N = kA^a.$$

$k =$ __________ and $a =$ __________

Determine the sum of square errors between the species model and the data.

Sum of Square Errors = __________.

b. In the Lab Report, plot the data and the best power law fit. Show the formula for the best fitting model on the graph. How well does the graph match the data?

c. Next, we want to fit a *straight line* to the logarithms of the data. From the *allometric model* above, we obtain the formula

$$\ln(N) = \ln(k) + a\ln(A).$$

In the table above, take the logarithm of the Number of Species ($\ln(N)$) and the logarithm of the Island Area ($\ln(A)$). Graph the logarithms of the data ($\ln(N)$ vs $\ln(A)$), then apply a *linear least squares* fit of the data.

What is the value of the slope of this best fitting line?

Slope = __________.

What is the value of the $\ln(N)$-intercept?

Intercept = __________.

d. In the Lab Report, plot the logarithm of the data and the best straight line fit to these data. Show the formula for the best fitting linear model on the graph. How well does the graph match the data? Write a brief discussion of how the coefficients obtained in this manner compare to the ones found in Part a.

e. The Caribbean island of Saint Croix has an area of 214 km^2. Use the allometric model found above to determine the predicted number of species of herpetofauna on this island

Number of species on Saint Croix = ____________.

If it is found that Saint Croix has 13 species of herpetofauna on the island, then determine the percent error between the model and the actual number of species found. (Use the actual number of species as the best value in the percent error formula.)

Percent Error (species) for Saint Croix = ____________.

Suppose that the island of Martinique has an area of 1133 km^2. Use the allometric model found above to determine the predicted number of species of herpetofauna on this island

Number of species on Martinique = ____________.

If it is found that Martinique has 30 species of herpetofauna on the island, then determine the percent error between the model and the actual number of species found.

Percent Error (species) for Martinique = ____________.

Now suppose that the Caribbean island of Puerto Rico is found to have 42 species of herpetofauna. Use the allometric model found above to determine the predicted an area km^2 of Puerto Rico.

Area of Puerto Rico = ____________km^2

The actual area of Puerto Rico is 8899 km^2. Use this information to determine the percent error between the model and the actual area. (Use the actual area as the best value in the percent error formula.)

Percent Error (area) for Puerto Rico = ____________

Suppose that the Caribbean island of Cuba is found to have 96 species of herpetofauna. Use the allometric model found above to determine the predicted an area km^2 of Cuba.

Area of Cuba = ____________km^2

The actual area of Cuba is 120961 km^2. Use this information to determine the percent error between the model and the actual area.

Percent Error (area) for Cuba = ____________

f. This problem helps address the importance of maintaining a large tract of land for the maintenance of biodiversity. Based on the allometric model computed above, determine a factor of how much a land area must increase to double the number of species supported by the environment.

Area is multiplied by ____________to double the number of species of herpetofauna in the Caribbean region.

g. Write a brief paragraph about this mathematical model. What does this model say about the size of land tracts needed to preserve biodiversity? How can this model be extended to different regions on Earth and different types of animals? Would you expect a similar power law relationship, and if so, which of the coefficients in a new model for a different set of species in a different geographic location is most likely to be similar to the one computed for the herpetofauna data?

h. (Extended Study) In the discussion of the allometric modeling technique, there was the alternate method of fitting the data using *nonlinear least squares*. This problem can be approached using Excel's *Solver* to solve this highly nonlinear problem. The coefficients in Part a and least squares values can be found, then applied to the specific examples in Part e and the land area factor in Part f. The percent errors will suggest which model is fitting the data better. Finally, one can add the graph of this *nonlinear least squares best fit allometric model* to the graph from Part b to visually compare the differences.

Problem 3.7 (Allegheny Forest)**.** *(Lab121_E3_allegheny.pg) Data from the Allegheny National Forest is used to model the volume of trees as a function of diameter or height. Both linear and allometric models are examined.*

The data in Table 3.7 came from the Allegheny National Forest in Pennsylvania[14]. The data were collected from Black Cherry trees with the diameter measured 4.5 ft above ground. One issue was whether either the diameter or the height of a tree accurately predicts the useful volume of wood in the tree. The volume is measured in cubic feet. This problem studies both *linear* and *allometric models*.

Diameter (in)	Height (ft)	Volume (ft^3)	Diameter (in)	Height (ft)	Volume (ft^3)	Diameter (in)	Height (ft)	Volume (ft^3)
8.3	70	10.3	11.4	76	21	14.2	80	31.7
8.6	65	10.3	11.4	76	21.4	14.5	74	36.3
8.8	63	10.2	11.7	69	21.3	16	72	38.3
10.5	72	16.4	12	75	19.1	16.3	77	42.6
10.7	81	18.8	12.9	74	22.2	17.3	81	55.4
10.8	83	19.7	12.9	85	33.8	17.5	82	55.7
11	66	15.6	13.3	86	27.4	17.9	80	58.3
11	75	18.2	13.7	71	25.7	18	80	51.5
11.1	80	22.6	13.8	64	24.9	18	80	51
11.2	75	19.9	14	78	34.5	20.6	87	77
11.3	79	24.2						

Table 3.7: Data for the diameter, height, and volume of Black Cherry trees from the Allegheny National Forest.

a. The simplest model is a *linear model*, so find the best line through the data. Begin with a linear model for the volume, V_d, in ft^3 as a function of the diameter, d, in inches

$$V_d = md + b.$$

The best linear model has a slope, m, and intercept, b with

$m =$ __________ and $b =$ __________

Determine the sum of square errors between this model and the data.

Sum of Square Errors = __________.

Repeat this process for a linear model for the volume, V_h, in ft^3 as a function of the height, h, in ft

$$V_h = kh + c.$$

This best linear model has a slope, k, and intercept, c with

$k =$ __________ and $c =$ __________

Determine the sum of square errors between this model and the data.

Sum of Square Errors = ____________.

Consider the tree that has a volume of 19.9 ft^3. The data give this tree a diameter of 11.2 in and a height of 75 ft. Use the models above to predict the expected volume and the percent error for each of the models (assuming the measured value to be the best value).

For $d = 11.2$, $V_d =$ ____________ ft^3 with percent error = ____________%.

For $h = 75$, $V_h =$ ____________ ft^3 with percent error = ____________%.

Based on the two linear models and the graphs with the data, which data set (DIAMETER or HEIGHT) better predicts the volume of a tree? __________________

b. In the Lab Report, create a graph of the data and a *linear model* for volume as a function of diameter. Show the formula for the model on the graph. Also, create a second graph of the data and a *linear model* for volume as a function of height with its formula on the graph. Write a brief explanation for why you would expect one of these measurements to better predict the volume based on the biology of trees. What happens with both models as the diameter or height gets close to zero? Is this biologically feasible?

c. For this part of the problem, we examine an *allometric model* for the volume as a function of the diameter. Consider the model given by

$$V = Kd^A.$$

The best fitting coefficients using a *linear fit* to the logarithms of the data, K and A, are found with Excel's *Trendline* and given by

$K =$ ____________ and $A =$ ____________.

Determine the sum of square errors between the *allometric* model and the data.

Sum of Square Errors = ____________.

Consider the tree that has a volume of 19.9 ft^3. The data give this tree a diameter of 11.2 in. Use this model to predict the expected volume and the percent error for the model (assuming the measured value to be the best value).

For $d = 11.2$, $V =$ ____________ ft^3 with percent error = ____________%.

d. In the Lab Report, plot the data and the best *allometric model*. Be sure to include the formula for the model on the graph. Find a reasonable explanation for the *power* that is obtained with the *allometric model*. (Round the power to the nearest integer and explain why that integer value is relevant to relating volume of a tree to its diameter to this power.) Also, have Excel plot a *log-log plot* of the data and the model that is found in Part c. Do the data roughly fall on a straight line in this log-log plot?

e. The method above of fitting an *allometric* or *power law model* to a set of data uses the best straight line through the logarithms of the data. In the case above, we have an allometric model of the form $V = Kd^A$, which can be written in the form

$$\ln(V) = \ln(K) + A\ln(d).$$

In the table above, find the logarithm of the volume ($\ln(V)$) and the logarithm of the diameter($\ln(d)$). Graph the ($\ln(V)$) vs ($\ln(d)$) data, then apply a *linear least squares best fit* to the logarithms of the data.

What is the value of the slope of this best fitting line? Slope = ____________.

What is the value of the $\ln(V)$-intercept?

Intercept = ____________.

f. In the Lab Report, plot the logarithms of the data and the best straight line fit to these logarithms. Show the formula for the best fitting linear model on the graph. How well does the graph match the data? Write a brief discussion of how the coefficients obtained in this manner compare to the ones found in Part c. Write a brief discussion comparing these different forms of the allometric model graphs shown in Parts d and f.

h. (Extended Study) In the discussion of the *allometric* modeling technique, there was the alternate method of fitting the data using *nonlinear least squares*. This problem can be approached using Excel's *Solver* to solve this highly nonlinear problem. The coefficients in Part c and least squares values can be found with this method, then applied to the specific example in Part c. Finally, one can add the graph of this *nonlinear least squares best fit allometric model* to the graph from Part d to visually compare the differences.

6 OTHER NONLINEAR FUNCTIONS

The techniques above have examined models with *linear* and *power law* functions. In the process, information has been given to enter *quadratic*, *exponential*, and *logarithmic* functions into Excel and Maple. In this section, several other examples are explored that use different special functions. The first section explores weak acid chemistry and introduces the *square root* and $\log_{10}$ functions. Later in this section *rational* functions are applied to an example of enzyme kinetics.

6.1 WEAK ACID CHEMISTRY

Many of the organic acids found in biological applications are weak acids. Formic acid (HCOOH) is a relatively strong weak acid that ants use as a defense[1]. The strength of this acid makes the ants very unpalatable to predators. The chemistry of dissociation is given by the following equation:

$$\mathrm{HCOOH} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{H}^+ + \mathrm{HCOO}^-.$$

Each acid has a distinct equilibrium constant K_a that depends on the properties of the acid and the temperature of the solution. For formic acid, $K_a = 1.77 \times 10^{-4}$. Let [X] denote the concentration of a particular chemical species X, then assuming that the formic acid is in equilibrium, it satisfies the following equation:

$$K_a = \frac{[\mathrm{H}^+][\mathrm{HCOO}^-]}{[\mathrm{HCOOH}]}.$$

The *normality* of a weak acid solution is a concentration measurement for reactions. In fact, the normality x = the molarity of the weak acid (undissociated + the number of protons from dissociated weak acid). If formic acid is added to water, then dissociation causes the acid ion (H^+) to be approximately the same as the formic ion (HCOO^-), this means $[\mathrm{H}^+] \approx [\mathrm{HCOO}^-]$. Also, if x is the normality of the solution, then $x = [\mathrm{HCOOH}] + [\mathrm{HCOO}^-]$. (The formic anion HCOO^- must be either in the bound form HCOOH or ionized

[1] en.wikipedia.org/wiki/Formic_acid, last visited 10/19/13

form $HCOO^-$ with the total representing the normality of the formic acid added to solution.) It follows that $[HCOOH] \approx x - [H^+]$. From the approximations, the equation for K_a is written:

$$K_a = \frac{[H^+][H^+]}{x - [H^+]},$$

which is easily solved using the quadratic formula to give,

$$[H^+] = \frac{1}{2}\left(-K_a + \sqrt{K_a^2 + 4K_a x}\right).$$

Notice that we only take the positive solution from the quadratic equation to make physical sense.

Excel and Maple for Square Root and $\log_{10}$

The formula above is used to find the concentration of $[H^+]$ for a solution of formic acid with a normality of 0.1 N and determine its pH. Since formic acid has $K_a = 1.77 \times 10^{-4}$ and $x = 0.1$, the formula given above yields:

$$[H^+] = \frac{1}{2}\left(-0.000177 + \sqrt{(0.000177)^2 + 4(0.000177)(0.1)}\right) \approx 0.00412.$$

In either Excel or Maple, the square root function is written `sqrt` though Maple must be lower case, while case doesn't matter in Excel, so

```
0.5*(-0.000177+sqrt(0.000177^2+4*0.000177*0.1)).
```

This expression gives the answer that the concentration of acid in the 0.1 N solution is $[H^+] = 4.12 \times 10^{-3}$. Clearly, we need '=' in front for Excel to evaluate the expression. In either program, we can substitute raising the expression following `sqrt` to the $\frac{1}{2}$ power instead.

Since the pH is defined to be $-\log_{10}[H^+]$, then for Maple, where $\log_{10}$ is written `log10`, the expression

```
-log10(0.00412)
```

gives a pH of 2.385. Excel is different from most computer languages. It assumes the user wants $\log_{10}$, so the expression

```
= -LOG(0.00412)
```

yields the pH of 2.385.

Problem 3.8 (Weak Acid Study). *(Lab121_C2_weak_acid.pg) A quadratic approximation of weak acid chemistry is used to find the* $[H^+]$ *and the pH of a solution, as the normality of the solution varies.*

Acetic acid (CH_3COOH) is a weak acid with an equilibrium constant $K_a = 1.75 \times 10^{-5}$. Use the techniques developed above to find an expression for $[H^+]$ as a function of the normality, x, of the weak acid solution.

a. In the Lab Report, write the expression for $[H^+]$ as a function of x using the value of $K_a = 1.75 \times 10^{-5}$. One should write this formula using Microsoft Equation or some other technical writing tool. (Note that this Lab Manual and many technical writings use the language LaTeX, and Maple has the ability to produce LaTeXexpressions very easily. However, the reader can study this topic on his or her own.)

b. Fill in the following table:

$x = 0.05$ N	$x = 0.45$ N	$x = 1.45$ N
$[H^+] =$ ____________	$[H^+] =$ ____________	$[H^+] =$ ____________

Also, when $[\mathrm{H}^+] = 0.004$, find

$x =$ ____________N.

c. In the Lab Report, create a graph of the $[\mathrm{H}^+]$ as a function of the normality x for $x \in [0.001, 2]$. Be sure to properly label the axes.

d. The pH of a solution is given by

$$\mathrm{pH} = -\log_{10}([\mathrm{H}^+]).$$

Fill in the following table:

$x = 0.05$ N	$x = 0.45$ N	$x = 1.45$ N
pH = ____________	pH = ____________	pH = ____________

Also, when the pH = 2.8, find

$x =$ ____________N.

e. In the Lab Report, create a graph of the pH as a function of the normality x for $x \in [0.001, 2]$. Be sure to properly label the axes.

6.2 RATIONAL FUNCTIONS

A *rational function*, $r(t)$, is defined by the quotient of two polynomials, $p(t)$ and $q(t)$, with

$$r(t) = \frac{p(t)}{q(t)},$$

and defined for all t provided $q(t) \neq 0$. This form of function is very important in *enzyme kinetic* models and some types of *population* models. Rational functions are characterized by having *vertical asymptotes*, where $q(t) = 0$ (unless $p(t) = 0$ at the same t), and *horizontal asymptotes*, if the degree of $p(t)$ is less than or equal to the degree of $q(t)$. Asymptotes complicate the graphing of this type of function.

Example 3.5 (Rational Function)**.** *This example considers the intersection of a line and a rational function and shows how to create a detailed graph of the functions.*

Consider the two functions:

$$f(x) = x - 1 \qquad \text{and} \qquad g(x) = \frac{x}{x^2 - 4}.$$

The function $f(x)$ is a *linear function* with a slope of 1 and intercept of -1. The *rational function* has a *domain* of all $x \neq \pm 2$, which indicates there are *vertical asymptotes* at $x = \pm 2$. Since the degree of the numerator is less than the degree of the denominator, there is a *horizontal asymptote* at $y = 0$. Since numerator of $g(x)$ has only *odd* powers and the denominator has only *even* powers, $g(x)$ is an *odd function* and is *symmetric* about the origin. The only x and y-intercept is the origin.

From this information, it should be easy to sketch the graph of these functions. The graph would show these functions intersecting at three points. To find these points of intersection would require solving a *cubic equation*, so it is time to turn to Maple to find the points of intersection. Graphing commands for both Maple and Excel are shown.

Maple and Excel Commands for Rational Function

As seen earlier, Maple provides an easy means of graphing the functions, and this graph is useful for estimating the points of intersection. The functions are entered in the following manner:

$$f := x \to x - 1; \qquad\qquad g := x \to \frac{x}{x^2 - 4}$$

Since $g(x)$ has vertical asymptotes, the following command makes a good graph, which is shown on the left in Fig. 3.7:

```
plot({f(x), g(x)}, x = -8 .. 8, y = -15 .. 15, discont = true)
```

Notice that we have added the additional options that limit the range of the function and a special command to tell Maple that it should allow for vertical asymptotes (*discont* = *true*).

Either from our intuition about the functions or the graph, we recognize there are three points of intersection, which are separated by the vertical asymptotes. Once again, Maple provides an easy tool to find these points of intersection. However, unlike its capability of finding points of intersection for polynomials, where it can find all solutions in one line, Maple generally only looks for a single solution when there is a nonpolynomial function. It does this to prevent falling into an infinite loop. For this problem, there are three points of intersection to find, so the following Maple commands obtain these points of intersection:

```
x1 := fsolve(f(x) = g(x), x = -5 .. -2); f(x1)
x2 := fsolve(f(x) = g(x), x = -2 .. 2);  f(x2)
x3 := fsolve(f(x) = g(x), x = 2 .. 5);   f(x3)
```

The second expression in the *fsolve* command shows a range where the user has Maple search to find the x value of the intersection point, and this value is stored in one of the variables, `x1`, `x2`, or `x3`. The command `f(x1)` provides the corresponding y-value for `x1`. The resulting points of intersection are:

$$(x_1, y_1) = (-2.1642, -3.1642), \quad (x_2, y_2) = (0.7729, -0.2271), \quad \text{and} \quad (x_3, y_3) = (2.3914, 1.3914).$$

The Excel graph shown on the right in Fig. 3.7 is significantly more complex. Many more curves were added to this graph to illustrate the asymptotes and the distinct continuous curves between the *vertical asymptotes*. Because the *rational function* is split into three distinct domains, the basic graphing template provided in Chapter 2 is not appropriate. However, the idea of having at least 40 points for any curve needs to be implemented. Including the three asymptotes, this graph is composed of seven curves, which need to be entered into Excel. This example demonstrates the amount of control and detail, which can be entered into the Excel graph to produce an annotated graph with asymptotes.

Any graph in Excel, which includes significant curvature or vertical asymptotes, is going to require more points where the function is changing rapidly. With the vertical asymptotes near $x = \pm 2$, this particular graph requires special consideration near these vertical asymptotes. The domain of interest for this graph is $x \in [-8, 8]$. The function $g(x)$ requires splitting this domain into the intervals $x \in [-8, -2)$, $x \in (-2, 2)$, and $x \in (2, 8]$. The stepsize needed near the asymptotes is small for good resolution of the curve, while away from the asymptotes the stepsize can be larger.

This particular graph is started with the *linear* function $f(x)$. In Cells $A2\!:\!A4$, enter -8, 0, and 8 (with $A1$ containing the label x). Use $B1$ for the label $f(x)$, then in $B2$, enter the formula '= A2 - 1' and *pull down* to $B4$. *Highlight* the six cells, $A2\!:\!B4$ and *Insert* a *Scatter* plot. Delete the legend, as the graph will be labeled with the *Text Box* option from the *Main menu*. As always, adjust the x-axis to go from $x = -8$ to $x = 8$. (*Chart Tools*, *Layout*, *Axes*, *Primary Horizontal Axis*, and *More Primary Horizontal Axis Options*....)

As noted above, we need the stepsize to be small (around 0.02-0.025) near the asymptote, but can be larger further away. Begin in Cell $C2$ entering -8 (with Cell $C1$ having the label x). In $C3$, enter '`= C2 + 0.1`' and *pull down* this formula to $C42$ ($x = -4$). In $C43$, enter '`= C42 + 0.05`' and *pull down* this formula to $C62$ ($x = -3$). Finally, in $C63$, enter '`= C62 + 0.025`' and *pull down* this formula to $C101$ ($x = -2.025$). In $D2$, enter '`= C2/(C2^2 - 4)`' and *pull down* this formula to $D101$. *Right click* on the graph and choose

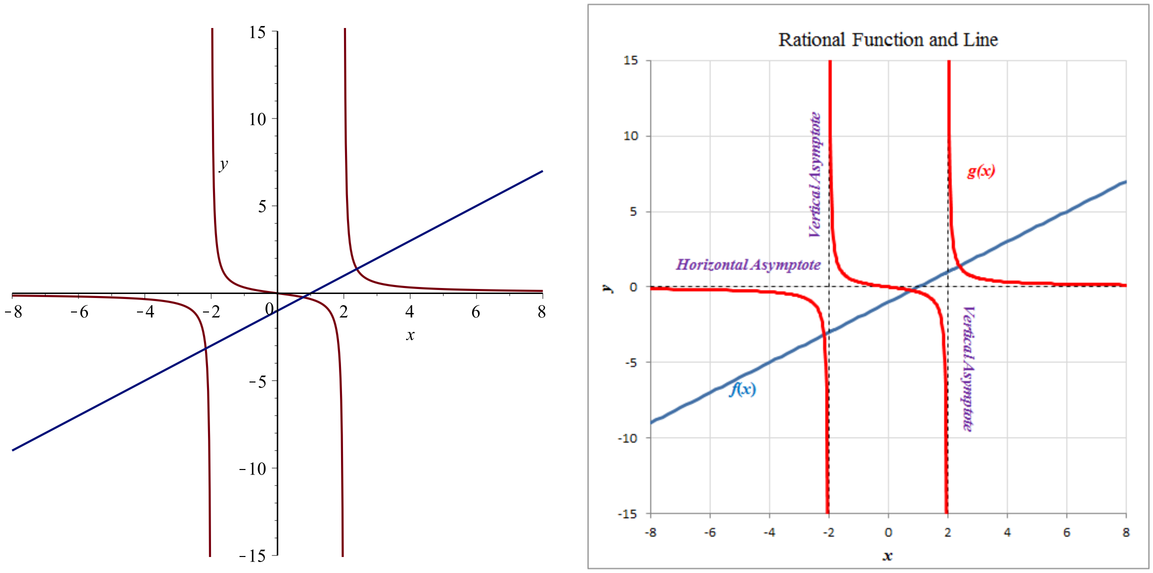

Figure 3.7: Graphs of $f(x) = x - 1$ and $g(x) = \frac{x}{x^2-4}$. On the left is the graph produced by Maple, while the graph on the right is constructed in Excel.

Select Data ... From the new window that opens, choose *Add*, then in the *Edit Series* menu, give the *Series* a name and click on the space below *Series X values:* and highlight Cells $C2:C101$. Finally, click on the space below *Series Y values:*, delete the $\{1\}$ that appears, highlight Cells $D2:D101$, and select *OK*. This should result in the left branch of $g(x)$ appearing on the graph. This would be a good time to adjust the *range* of this graph by forcing the y-axis to only go from $y = -15$ to $y = 15$.

Similarly, partition the x values for $x = -1.975$ to $x = 1.975$ in Column E (noting that the small stepsizes are needed at both the beginning and end of this interval) with $g(x)$ evaluated in Column F. This allows the user to *Add* the center section of $g(x)$ to the graph, using entries in Columns E and F. A partition of the x values from $x = 2.025$ to $x = 8$ in Column G and $g(x)$ evaluated in Column H are used to produce the right section of $g(x)$ for the graph. With all the elements of $g(x)$ available, *right click* on each segment selecting *Format Data Series...*, then choose *Line Color* and *Solid line* with the choice of *red* to give uniformity to $g(x)$.

The next step is adding the asymptotes, which are simply lines. Excel handles lines best when three points are given (though it can be done with just two points). The first *vertical asymptote* occurs at $x = -2$, so enter -2 in say Cells $I2 : I4$ with the y-values -16, 0, and 16 in Cells $J2 : J4$. Once again, *Add* the data from these six cells to the graph to create this *vertical asymptote*. *Right click* on this line segment and *Format Data Series...* to become a *dashed black line*. Enter the points $(2, -16)$, $(2, 0)$, and $(2, 16)$ elsewhere on the spreadsheet and follow the procedure above to create the second *vertical asymptote*. Finally, enter the points $(-8, 0)$, $(0, 0)$, and $(8, 0)$ elsewhere on the spreadsheet and follow the procedure above to create the *horizontal asymptote*, which will also be taken to be a *dashed black line*.

The remainder of the work to be done on the graph is the usual adding *Chart* and *Axes Titles* and inserting *Gridlines*. To be able to view the asymptotes clearly, the user should *right click* on the gridlines and format them to become *gray lines*. The *Text Box* under *Chart Tools* is used to create and format all the labels for the functions and asymptotes. All text in this graph uses the usual *Times New Roman* font. The end result after these many steps is the graph shown on the right of Fig. 3.7.

Problem 3.9 (Rational and Linear Functions)**.** *(Lab121_D1_ratl.pg) This problem explores a line and ratio-*

nal function. The points of intersection are found, and a graph is created.

Consider the functions

$$f(x) = 3.7 - x \qquad \text{and} \qquad g(x) = \frac{1.8x}{6.5 - x - x^2}.$$

a. Find the x and y-intercepts for both of these functions. Find the slope of the linear function, $f(x)$. List any asymptotes (vertical and horizontal) for the rational function, $g(x)$.

For the linear function, $f(x)$:

x-intercept = __________, y-intercept = __________, and Slope = __________

For the rational function, $g(x)$:

x-intercept = __________ and y-intercept = __________

Vertical Asymptotes: $x =$ __________, __________

Horizontal Asymptote: $y =$ __________

b. Find all points of intersection between the graphs of $f(x)$ and $g(x)$.

Points of intersection for $f(x)$ and $g(x)$,

(__________, __________), (__________, __________), and (__________, __________).

c. In the Lab Report, create a graph of these functions for $x \in [-10, 10]$ with the range restricted so that $y \in [-10, 10]$. On the graph, use points to highlight the points of intersection, and clearly show and label the vertical and horizontal asymptotes with dashed lines. Briefly discuss how the vertical and horizontal asymptotes are found. Describe how the points of intersection are found.

Problem 3.10 (Rational and Quadratic Functions)**.** *(quadrat.pg) This problem explores properties of a quadratic and rational function. The points of intersection are found, and graphs of these functions are created.*

Consider the quadratic function

$$f(x) = x^2 - 3x - 5$$

and the rational function

$$g(x) = \frac{20x}{1.4 + x}.$$

a. Find the x and y-intercepts for both of these functions. Find the vertex of the quadratic function, $f(x)$. List any asymptotes (vertical and horizontal) for the rational function, $g(x)$.

For the Quadratic Function:

x-intercepts = __________, __________ and y-intercept = __________

The vertex $(x_v, y_v) =$ (__________, __________)

For the Rational Function:

x-intercept = __________ and y-intercept = __________

Vertical Asymptote: $x =$ ____________

Horizontal Asymptote: $y =$ ____________

b. Find all points of intersection between the graphs of $f(x)$ and $g(x)$.

Points of intersection for $f(x)$ and $g(x)$,

(__________, __________), (__________, __________), and (__________, __________).

c. In the Lab Report, create a graph of these functions for $x \in [-10, 10]$ with the range restricted to $y \in [-20, 50]$. On the graph, use circles to highlight the points of intersection, and clearly show and label the vertical and horizontal asymptotes with dashed lines.

6.3 MICHAELIS-MENTEN ENZYME KINETICS

Proteins are considered the primary building blocks of life, and *enzymes* are an important class of proteins that catalyze many of the reactions occurring inside the cell. An enzyme facilitates a biochemical reaction with both speed and specificity to break down or build up key biochemical components. The most basic enzymatic reaction, known as the *Michaelis-Menten* mechanism[13], has an enzyme, E, combine with a substrate, S, to form a product, P:

$$\text{E+S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{ES} \xrightarrow{k_2} \text{E+P}.$$

The rates described in the biochemical equation above can be described in a mathematical equation known as a differential equation.

It has been observed that the enzyme-substrate forms very rapidly, while the forward reaction (also known as *turnover number*), k_2, occurs more slowly. This allows a simplifying assumption for a Michaelis-Menten (MM) enzyme to describe the rate of the forward reaction to create the product, P, as

$$V([S]) = V = \frac{V_{max}[S]}{K_m + [S]},$$

where $[S]$ is the concentration of the substrate. This rate function is a rational function with important constants, V_{max} and K_m, which characterize a MM enzyme. V_{max} is the maximum rate of the reaction and depends primarily on the enzyme concentration and temperature, while K_m is the Michaelis-Menten constant, which characterizes a particular enzyme, and depends primarily on temperature. Measuring these constants is very important for understanding the efficiency of a MM enzyme. In this form, a nonlinear least squares fit to the data is necessary for finding the constants, V_{max} and K_m.

Lineweaver-Burk plot

The rational function describing the forward reaction of a MM enzyme is a nonlinear function, which complicates finding V_{max} and K_m from experiments. However, Lineweaver and Burk[8] discovered an easy transformation to change the nonlinear relationship into a linear function. Previously, when logarithms of the data appeared linear, then an allometric model was appropriate. In this case, the transformation of the data uses inverses. We define the change of variables:

$$x = \frac{1}{[S]} \qquad \text{and} \qquad y = \frac{1}{V([S])} = \frac{1}{V},$$

so x is the inverse of the substrate concentration and y is the inverse of the forward reaction rate. The transformation gives:

$$\frac{1}{y} = \frac{V_{max}(1/x)}{K_m + (1/x)},$$

$$y = \frac{K_m}{V_{max}}x + \frac{1}{V_{max}},$$

which is a linear relationship between x and y. It follows that if the plot of these transformed variables, the *Lineweaver-Burk plot*, is approximately linear, then the enzyme is likely a MM enzyme and a *linear least squares best fit* provides a means to obtain V_{max} and K_m. This method has long been used by biochemists to find the parameters V_{max} and K_m to characterize an enzyme.

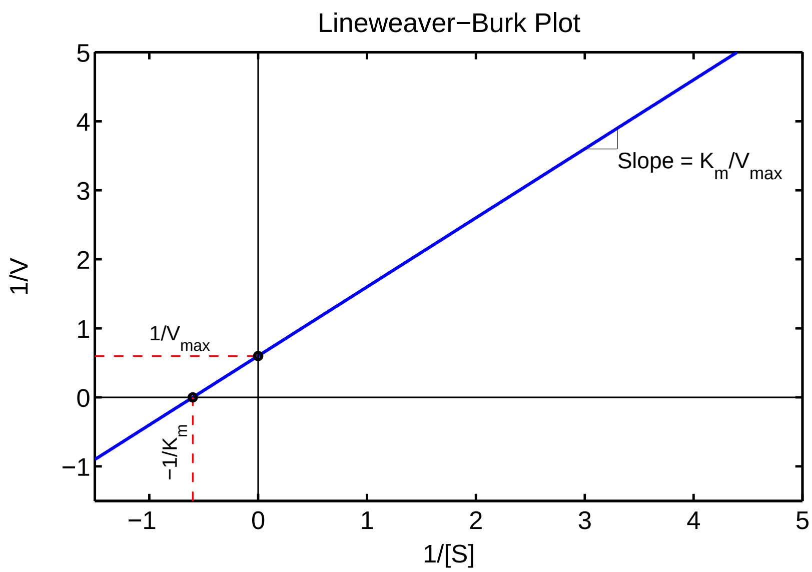

Figure 3.8: Graph for the Lineweaver-Burk function obtained from the Michaelis-Menten rate function.

Problem 3.11 (Michaelis-Menten Enzyme Kinetics)**.** *(mm_enzyme.pg) Many enzymes follow Michaelis-Menten kinetics, and this problem takes data from an enzymatic reaction (demethylase) to fit the classic rational reaction formula for enzymes.*

[AMI] (μM)	N formation (nmol/min/mg)	[AMI] (μM)	N formation (nmol/min/mg)
5	0.20	100	2.17
10	0.45	200	2.68
15	0.60	260	2.89
25	0.80	330	3.05
50	1.35	500	3.12

Table 3.8: Data for formation of nortriptyline (N) by human liver microsomes from cytochrome P450 mediated demethylation of amitriptyline (AMI)[15].

Table 3.8 provides data from Schmider *et al.*[15] for cytochrome P450 mediated demethylation of amitriptyline (AMI) to nortriptyline (N) by human liver microsomes. From this table, let $[S] = [\text{AMI}]$ and $V([S]) =$ N formation. We want to find the best values of V_{max} and K_m to characterize the enzyme cytochrome P450.

a. The next step in this analysis is to take the data in the table above and create a new table with the values $x = 1/[S]$ and $y = 1/V([S])$. Create this new table, then plot the data y vs x. Use a linear least squares technique, such as Excel's *Linear Trendline*, to find the straight line that best fits these transformed data.

x-intercept = __________ and y-intercept = __________

The slope = __________

Use this information to find V_{max} and K_m.

V_{max} = __________ (nmol/min/mg) and K_m = __________ (μM)

b. Use the values of V_{max} and K_m calculated above to obtain the model predications from the MM reaction formula for the production of nortriptyline (N) when [AMI] is $[S] = 15, 50, 100, 200$, and 500. Assuming that the actual data are the best values, then calculate the percent error between the model and the actual data. Fill in the following table:

$[S]$	$V([S])$	Percent Error
15		%
50		%
100		%
200		%
500		%

c. In the Lab Report, create a Lineweaver-Burk plot showing the transformed data and the best fitting linear model through these data. Write a brief paragraph to describe how well the linear model fits the data.

d. The transformation results in certain biases of the data. In the section for *allometric modeling*, techniques were developed using Excel's *Solver* to find the nonlinear least squares best directly to the model. (See Example 3.4.) On an Excel spreadsheet, create *Named variables* for the parameters V_{max} and K_m, guessing the values from Part a, then evaluate the nonlinear function with these parameters:

$$V = \frac{V_{max}[S]}{K_m + [S]},$$

at each of the data values $[S]$ in Table 3.8. Compute the sum of square errors between the data and this nonlinear MM reaction model. Use Excel's *Solver* to minimize this sum of square errors and find the best fitting parameters V_{max} and K_m. Record both this sum of square errors with the parameters V_{max} and K_m from both Part a and the *Least Sum of Square Errors* found by Excel's *Solver*.

$V_{max} =$ ____________ (nmol/min/mg) and $K_m =$ ____________ (μM)

Sum of Square Errors (from Part a) = ____________

Least Sum of Square Errors (*Solver*) = ____________

Find all intercepts and any asymptotes for this rational function for $[S] \geq 0$. (If no asymptote exists, then write 'None,')

$[S]$-intercept = ____________ and V-intercept = ____________

Vertical Asymptote: $[S] =$ ____________

Horizontal Asymptote: $V =$ ____________

e. Use the values of V_{max} and K_m calculated in Part d to obtain the model predications from the MM reaction formula for the production of nortriptyline (N) when [AMI] is $[S] = 15, 50, 100, 200$, and 500. Assuming that the actual data are the best values, then calculate the percent error between the model and the actual data. Fill in the following table:

$[S]$	$V([S])$	Percent Error
15		%
50		%
100		%
200		%
500		%

f. In the Lab Report, create a plot showing the original data. On this plot add the two MM models, which use the the parameters from Parts a and d. (Be sure to label which plot goes with which model.) Write a brief paragraph to describe how well each of the models fits the data and decide which model is better. Include the information from comparing the sum of square errors. Discuss some of the advantages and disadvantages of each of the models.

CHAPTER 4:

DISCRETE DYNAMICAL MODELS

Keywords: Discrete dynamical model, Malthusian growth, exponential growth, nonautonomous growth, logistic growth

This chapter introduces basic discrete dynamical models. The behavior of these models varies in complexity from simple exponential solutions to chaotic dynamics. From a computational perspective, Excel's *Updating* or *Pull Down* feature is excellent for modeling of discrete systems. The models begin with simple linear models, then extend to both nonlinear discrete dynamical models and ones that include time-varying dependence. Analysis of these models can be quite complex, but finding solutions numerically is relatively easy. This chapter extends methods from before to fit model parameters to data using Excel's *Solver*.

1 DISCRETE MALTHUSIAN GROWTH

This section begins with the simplest of discrete dynamical models. Discrete dynamical models are iterations of functions at discrete time steps. *Malthusian growth*, which is often called *exponential growth* of a population, uses the concept that over a fixed period of time, each individual has a fixed probability, r, of giving birth to a new individual. That is, the new population at time $n+1$ equals the old population plus r times the population at time n. If we let the population at time n be denoted P_n, then the discrete Malthusian growth model satisfies:

$$P_{n+1} = P_n + rP_n = (1+r)P_n.$$

For this model the solution is easy to obtain. Assume that the initial population is given by P_0, then

$$\begin{aligned} P_1 &= (1+r)P_0 \\ P_2 &= (1+r)P_1 = (1+r)^2P_0 \\ &\dots \\ P_n &= (1+r)P_{n-1} = (1+r)^nP_0. \end{aligned}$$

It follows that given an r and P_0, then the population is known for all times n with

$$P_n = (1+r)^nP_0.$$

This expression clearly shows why Malthusian growth is also known as exponential growth. Malthusian growth occurs for most organisms when there are plentiful resources and other conditions are relatively unchanged. We note that few discrete models of the form $x_{n+1} = f(x_n)$ have a closed form solution like this.

Example 4.1 (Early U. S. Population)**.** *The population of the U. S. in its early years follows a pattern of Malthusian growth very well.*

Let us apply the Malthusian growth model to the population in the U. S. over the first 50 years where census data are available, then see how this model works on the subsequent data. Table 4.1 of the U. S. population in millions is based on the census over the last 220 years.

Since the discrete Malthusian growth model has a solution, we will demonstrate **two** methods of fitting the data, which are equivalent. The first method finds the nonlinear least squares best fit to the solution of this model. The second method takes advantage of Excel's *Updating* feature to simulate the discrete

Year	Population	Year	Population	Year	Population
1790	3.93	1870	39.82	1950	150.70
1800	5.31	1880	50.19	1960	179.32
1810	7.24	1890	62.95	1970	203.30
1820	9.64	1900	76.21	1980	226.55
1830	12.87	1910	92.23	1990	248.71
1840	17.07	1920	106.02	2000	281.42
1850	23.19	1930	122.78	2010	308.75
1860	31.44	1940	132.16		

Table 4.1: The history of the U. S. census from 1790 to 2010. Population in millions[17].

Malthusian growth model, then fit the model to the data. In both cases the parameters r and P_0 are found that best fit the data over 50 years using Excel's *Solver*. In addition, since the discrete Malthusian growth has an exponential solution, we demonstrate an alternate model fit using Excel's *Trendline* and compare this fit to the nonlinear least squares fit.

The first method follows the techniques developed in the previous chapter for finding the nonlinear least squares best fit to the data with the model given by

$$P_n = (1+r)^n P_0,$$

where n is in decades. From Table 4.1, we copy the data for the U. S. population and rearrange the data into Columns A and B. Next we insert a column between A and B, so that we have a column for the number of decades, n, after the starting date, 1790. At the top of Columns A-E, we create the labels: 'Year,' 'n,' 'Population,' 'Model,' and 'SSE.' Column B has the entries, 0, 1, 2, ...

In Cells $G2$ and $G3$ enter the labels P_0 and r. In Cell $H2$, use the 1790 census $P_0 = 3.93$ as an initial estimate. The growth in the first decade is computed by solving

$$P_1 = (1+r)P_0 \qquad \text{or} \qquad r = \frac{P_1}{P_0} - 1 = \frac{5.31}{3.93} - 1 = 0.351.$$

It follows that a reasonable estimate for $r = 0.351$, which is entered into Cell $H3$. Apply the techniques from before to *name the variables* in Cells $G2$-$H3$. (We note that Excel will name r as `r_`, since r is a protected variable in Excel.) Since Column D contains the discrete Malthusian growth model, we enter in Cell $D2$ the formula, '`= P0*(1 + r_)^B2`' and *pull down* this formula to $D13$, as we are going to simulate the model to 1900 for the graph.

The problem asks for fitting the model with the first 50 years of data. In $E2$, the square error between the Model and the census data is found by typing '`= (C2 - D2)^2`.' This formula is *pulled down* to $E7$ to give the square errors for the first 50 years of data. In $E8$ we sum the square errors by clicking on the $\sum$ symbol under the *Home* or *Formulas* tabs from the *Main menu*. Thus, $E8$ will have the formula '`=SUM(E2:E7)`.'

To find the least sum of square errors, we use Excel's *Solver*. As a reminder, *Solver* is found under the *Data* tab in the *Main* menu on the right. Highlight the cell with the *sum of square errors*, $E8$, then click on *Solver*. A window pops up with the *Set Objective:* being the cell just selected. Check *Min* (short for minimize) to find the minimum of this Objective Cell, then click in the window for *By Changing Variable Cells:* and highlighting the Cells $H1:H3$, which contain the model parameters, P_0 and r. Finally, choose the *Solve* option, and Excel automatically adjusts the model parameters to find the *least sum of squares error*. In this case, the sum of square errors drops to 0.024099. The resulting best fitting parameters are $P_0 = 4.0179$ and $r = 0.33639$. This is the *nonlinear least square best fit* to the data.

Excel for Simulating Malthusian Growth Model

The discrete Malthusian growth model is an iterative model, which depends only on the previous population. It follows that a discrete population model can be written:

$$P_{n+1} = F(P_n),$$

where for the Malthusian growth model uses the linear function $F(P_n) = (1+r)P_n$. The function, $F(P_n)$, is called an *updating function*. Computers are particularly well-suited for this type of iterative simulation. For Excel, the *pull down* or *updating* function is automatic in Excel spreadsheet operations, and this function lends itself naturally to these discrete models.

The method of using Excel to simulate the discrete Malthusian growth model begins by setting up the Excel spreadsheet as described above. All of the entries are the same except for the ones in Column D, *Model*. Instead of entering the solution, which depended on $P0$, r, and n, we enter the initial value in Cell $D2$ by typing '`= P0.`' From the Malthusian growth model, the population at the next time period depends only on the previous population. This is simulated by typing '`= (1 + r_)*D2`' in Cell $D3$, then *pulling down* this formula. Since Excel *updates* the cells using this process, the operation above simulates the discrete Malthusian growth model. (Note that Column B is not needed for this simulation method.)

After simulating the model, we find the *least sum of square errors* using Excel's *Solver* in exactly the way described above. Once again, the sum of square errors drops to 0.024099 with the resulting best fitting parameters being $P_0 = 4.0179$ and $r = 0.33639$. Though it is not obvious to the user, this method requires significantly more computations by Excel's *Solver*. Notice that we could easily replace the line with the Malthusian growth model to satisfy any functional growth, including nonlinear growth, which may have no closed form solution. Thus, this method provides more generality to fitting parameters to discrete dynamical models.

Excel for Exponential Fit

The *Malthusian* or *exponential growth model* does have the special form of an *exponential function*. It follows that an alternative method can be used to fit the data. Since

$$\begin{aligned} P_n &= (1+r)^n P_0, \\ \ln(P_n) &= n\ln(1+r) + \ln(P_0), \end{aligned}$$

which shows that the *logarithm of the population* is a *linear function* of n. Thus, an alternative means of finding the parameters P_0 and r is using Excel's *Trendline* with the *Exponential fit*. This procedure uses the *linear least squares best fit* to the data, where we use the logarithm of the population, $\ln(P_n)$. This technique is very similar to the method used in finding parameters for a *power law* model in the previous chapter.

The technique to find this version of the discrete Malthusian growth model begins with the same Excel spreadsheet as described above. In this case, we are interested in Columns B and C with the variable n and the *census data*, respectively. In order to fit an *exponential function* to the U. S. population data over the first five decades, we highlight the Cells $B2:C7$ and insert a *Scatter* plot. On this plot, we *right click* on one of the data points and choose *Add Trendline* or alternately, select *Trendline* under the *Chart Tools: Layout*. In the window labeled *Format Trendline*, we select the *Trendline Option: Exponential* with the *Display Equation on chart*. The resulting model is

$$P_n = 3.9669e^{0.29388n}.$$

By solving the equation $1+r = e^{0.2939}$, we obtain the best parameters, which are $P_0 = 3.9669$ and $r = 0.3416$, which are slightly different from the nonlinear least squares best fit parameters. If the nonlinear least squares is computed with these parameters, then the sum of square errors is 0.045246, which is about double the value from Excel's *Solver* solution.

As we found for the allometric model, the nonlinear least squares and exponential fits give slightly different parameter values because the weighting of the data is different, with the exponential fit working better for early in the simulation and the nonlinear least squares fit working better over the entire range of the data. Table 4.2 shows the best fitting Malthusian and exponential growth models until 1870. This period of history satisfies the appropriate conditions for Malthusian growth. Thus, the population is expected to grow at a fairly constant rate until the onset of the Industrial Age. The small errors between the models and the census data confirm this.

Year	Census	Model 1	% Error	Model 2	% Error
1790	3.93	4.018	2.24	3.967	0.94
1800	5.31	5.370	1.12	5.322	0.23
1810	7.24	7.176	-0.88	7.140	-1.38
1820	9.64	9.590	-0.52	9.579	-0.63
1830	12.87	12.816	-0.42	12.851	-0.15
1840	17.07	17.127	0.34	17.241	1.00
1850	23.19	22.889	-1.30	23.131	-0.26
1860	31.44	30.589	-2.71	31.032	-1.30
1870	39.82	40.879	2.66	41.633	4.55

Table 4.2: This table shows the values of the Malthusian growth model (Model 1) and the exponential fit model (Model 2) with their percent error compared to the actual census data. The models were fit through the data in 1850 and can be seen to start diverging by 1870.

Graphing Discrete Dynamical Models

Fig. 4.1 shows graphs of the models for the discrete Malthusian growth for the U. S. population. The graph on the left is similar to ones prepared previously. The Excel spreadsheet needs to have columns with the years and census data from 1790 to 1900. Two other columns are needed with the discrete Malthusian growth models with the two sets of parameters described above. These four columns are highlighted and graphed by inserting a *Scatter plot*. As always, the domain is adjusted, and graph and axis titles are added. The data are made as points, and legends are modified to label the data and models. For discrete models, the models are actually only defined at the discrete points, so points would be appropriate. However, convention uses lines connecting points for discrete dynamical models (or simply lines if many points are needed). On the left of Fig. 4.1, we see the completed graph after adjustments are made to make the graph look good, which only requires skills acquired earlier in this manual. This graph shows that there is little variation between the *nonlinear* and *exponential fits* of the parameters to the *discrete Malthusian growth model*, largely because the model fits so well.

One method of determining if an *exponential growth* model is appropriate is to graph the population on a *logarithmic scale*. This produces what is called a *semilog plot*. If the data in a *semilog plot* appear almost *linear*, then an *exponential growth* model is appropriate. The graph on the right of Fig. 4.1 shows a *semilog plot* in Excel. This graph is very easy to obtain. *Copy* the normal graph created above, then use *Paste Special* with the option: *Microsoft Office Drawing Object*. Under the *Chart Tools: Layout* menu, select *Axes: Primary Vertical Axis*, then select *Show Axis with Log Scale*. Often the gridlines will be widely separated, so under the *Gridlines: Primary Horizontal Gridlines* options, we select the *Major & Minor Gridlines* option. This plot will show the exponential growth models as straight lines.

The graph shows that the data and the model are very close from 1790 to 1850 (error $< 2.3\%$). The errors in 1860 and 1870 are only off by 2.7% (low then high). During this period of time, society was primarily agrarian with plenty of land into which to expand (forgetting the impact on Native Americans) and stable family size, so the Malthusian growth model with its assumptions fits quite well. Growth rates of approximately 30% per decade are fairly typical of agrarian societies. Following 1870, the errors rapidly rise to over 28% by 1900. Thus, as the U. S. enters the industrial revolution, the population growth starts declining, which means that taking a constant value for r is inappropriate. We will model corrections to the growth term later, adding a time dependency or a nonlinearity in the population.

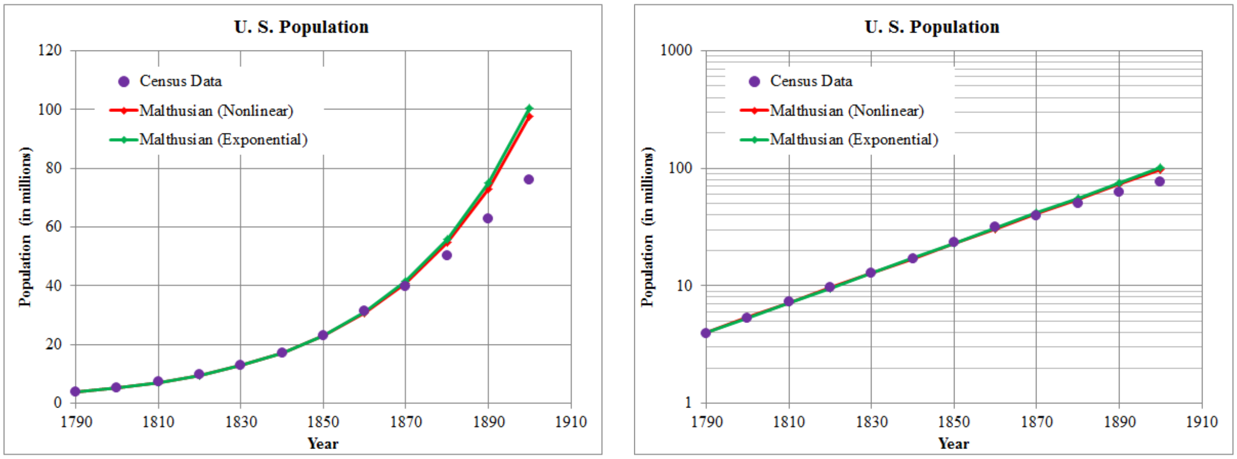

Figure 4.1: Plots of U. S. population with nonlinear least squares best fit and exponential fit. Data fit over the first 50 years, then simulated for 100 years. The graph on the right is a semilog plot.

Problem 4.1 (Malthusian Growth for Two Countries)**.** *(Lab121_F2_mal_country.pg) The discrete Malthusian growth model is applied to two different countries. Parameters are fitted to the model, and graphs are made to compare the solutions to the census data.*

The population of Japan [18] was $P_0 = 116.79$ million in 1980, while in 1990 it was $P_{10} = 123.53$ million. The population in Bangladesh [18] was $Q_0 = 87.97$ million in 1980, while it was $Q_{10} = 111.45$ million in 1990.

a. Over a limited range of years, the population of most countries can be estimated using the Malthusian growth law, which is given by:

$$P_{n+1} = (1+k)P_n \qquad \text{and} \qquad Q_{n+1} = (1+r)Q_n,$$

where n is the number of years since 1980 with P_n the population of Japan and Q_n the population of Bangladesh. Use the data above to find the annual growth rate of Japan

$k =$ __________

and annual growth rate of Bangladesh

$r =$ __________.

Find the Malthusian growth model for the population of Japan,

$P_n =$ ____________________ million.

(The expression above uses the values of P_0 and k and depends on n in years.)

Find the Malthusian growth model for the population of Bangladesh,

$Q_n =$ ____________________ million.

(The expression above uses the values of Q_0 and r and depends on n in years.)

Use the Malthusian growth model to predict the population of Japan in 2000,

$P_{20} =$ __________ million.

If the actual population of Japan was 126.76 million in 2000, then the percent error between the population and the model is

Percent Error = ____________%.

Use the Malthusian growth model to predict the population of Bangladesh in 2000,

$Q_{20} =$ ____________million.

If the actual population of Bangladesh was 136.66 million in 2000, then the percent error between the population and the model is

Percent Error = ____________%.

b. Use the Malthusian growth models to determine how long it takes for each country's population to double.

Population of Japan doubles in ____________years

Population of Bangladesh doubles in ____________years

According to the Malthusian growth models, when is the population of Japan equal to the population of Bangladesh.

Populations are equal in ____________years.

c. In your Lab Report, create a graph of the Malthusian growth models for both countries from 1980 to 2030. Include the data values of the populations for 1980, 1990, and 2000 on your graph. Briefly discuss how well you believe the models track the populations of these countries over the 50 years of simulation. Include some discussion of the strengths and weaknesses of using the Malthusian growth model for this simulation.

Problem 4.2 (Malthusian Growth and U. S. Population)**.** *(Lab121_F1_US_Malthus.pg) The discrete Malthusian growth model is applied to two different periods of U. S. history. Parameters are fitted to the model, and graphs are made to compare the solutions to the census data.*

In this problem we investigate Malthusian growth models applied to the U. S. population over various time intervals. The discrete Malthusian growth model is given by

$$P_{n+1} = (1+r)P_n, \qquad P_0 \quad \text{given},$$

where we assume that r is the **annual growth rate** and P_0 is the initial population at $n = 0$. The general solution to this model is given by

$$P_n = (1+r)^n P_0,$$

and our goal is to find the best fitting parameters r and P_0 to this Malthusian growth model.

The models are fitted to the data in two different ways. One method of fitting uses an exponential fit to the data (linear fit to the logarithm of the population data), and the other uses a nonlinear least squares best fit to the data.

a. For our first modeling effort we consider the U. S. population in Table 4.1 from 1790 to 1870. The initial year, 1790, is associated with $n = 0$. Plot the U. S. population data from 1790 to 1870 with $n = 0$ corresponding to 1790 and n measured in years, then apply Excel's *Trendline* to the data to create a model. This model has an exponential form, so use your knowledge of algebra to convert this exponential model into the same form as the general solution of the Malthusian growth model given above and identify the best fitting values of the initial population, P_0, and the annual growth rate, r. The best values for the initial population and best **annual growth rate** are:

$P_0 =$ ____________ $r =$ ____________.

Determine the sum of square errors between the Malthusian growth model and the data.

Sum of Square Errors = ____________.

Use this model to find the population in 1790 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1790 = ____________ Percent Error in 1790 = ____________%.

Also, use this model to find the population in 1840 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1840 = ____________ Percent Error in 1840 = ____________%.

Finally, project this model to find the population in 1890 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1890 = ____________ Percent Error in 1890 = ____________%.

b. For our second modeling effort we again consider the U. S. population in Table 4.1 above from 1790 to 1870. However, this time we use Excel's *Solver* to find the nonlinear least squares best fit to the data. This directly finds the best fitting parameters P_0 and r to the general solution of the Malthusian growth model to the given data. Using this method, give the best values for the initial population and **annual growth rate**

$P_0 =$ ____________ $r =$ ____________.

Determine the sum of square errors between the Malthusian growth model and the data.

Sum of Square Errors = ____________.

Use this model to find the population in 1790 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1790 = ____________ Percent Error in 1790 = ____________%.

Also, use this model to find the population in 1840 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1840 = ____________ Percent Error in 1840 = ____________%.

Finally, project this model to find the population in 1890 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1890 = ____________ Percent Error in 1890 = ____________%.

c. In your Lab Report, plot the data for 1790 to 1890. Include the graph of both Malthusian growth models found in Parts a and b. In addition, create a graph with a logarithmic scale on the population axis, as in Fig. 4.1. Write a brief paragraph describing how well the models fit the data. Which model fits the data better? Briefly describe any discrepancies that you observe on the graph between the models and the data and put these errors in the context of what you know about U. S. history. Are there significant differences in these two models? What are the advantages and disadvantages of using one model over the other?

d. We repeat the modeling effort above but use the the U. S. population in Table 4.1 from 1840 to 1920. Now let the initial year, 1840, be associated with $n = 0$. Plot the U. S. population data from 1840 to 1920 with $n = 0$ corresponding to 1840 and n measured in years, then apply Excel's *Trendline* to the data to create a model. This model has an exponential form, so convert this exponential model into the same form as the general solution of the Malthusian growth model given above and identify the best fitting values of the initial population, P_0, and the growth rate, r. The best values for the initial population and annual growth rate are:

$P_0 =$ ____________ $r =$ ____________.

Determine the sum of square errors between the Malthusian growth model and the data.

Sum of Square Errors = ____________.

Use this model to find the population at 1840 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1840 = ____________ Percent Error in 1840 = ____________%.

Also, use this model to find the population at 1890 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1890 = ____________ Percent Error in 1890 = ____________%.

Finally, project this model to find the population at 1940 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1940 = ____________ Percent Error in 1940 = ____________%.

e. Now consider the U. S. population in Table 4.1 from 1840 to 1920. This time we use Excel's *Solver* to find the nonlinear least squares best fit to the data. This directly finds the best fitting parameters P_0 and r to the general solution of the Malthusian growth model to these data. Using this method, give the best values for the initial population and **annual growth rate**

$P_0 =$ ____________ $r =$ ____________.

Determine the sum of square errors between the Malthusian growth model and the data.

Sum of Square Errors = ____________.

Use this model to find the population at 1840 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1840 = ____________ Percent Error in 1840 = ____________%.

Also, use this model to find the population at 1890 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1890 = ____________ Percent Error in 1890 = ____________%.

Finally, project this model to find the population at 1940 and determine the percent error from the population given in Table 4.1 (assuming that the value in the Table is the most accurate):

Model Population in 1940 = ____________ Percent Error in 1940 = ____________%.

f. In your Lab Report, plot the data for 1840 to 1940. Include the graph of both Malthusian growth models found in Parts d and e. Also, create a graph with a logarithmic scale on the population axis, like in Fig. 4.1. Write a brief paragraph describing how well the models fit the data. Which model fits the data better? Briefly describe any discrepancies that you observe on the graph between the models and the data and put these errors in the context of what you know about U. S. history over this period of time. What are the differences and similarities of these modeling efforts compared to the growth models over the earlier period studied at the beginning of this Computer Lab? Which range of data provides the better information using these models for predicting next two decades and why?

2 OTHER GROWTH MODELS: NONAUTONOMOUS, LOGISTIC, AND IMMIGRATION

The growth of the population in the U. S. has not continued at a constant rate. In fact, most populations have a decline in growth rate as populations increase. For human populations the decline in growth rate often results from improved health care and education or from severe crowding and disease. Studies have shown that an increased level of education for women results in the largest declines in growth rates because of better family planning and delayed reproduction. Technological advances, especially in medicine, cause reproduction rates to decline as families are more assured that their offspring will survive. It follows that a simple Malthusian growth rate cannot track populations for extended periods of time.

This section introduces three very important discrete dynamical population models. For human populations, the growth rate often depends on the time in history that the population is being studied. This leads to the time dependent discrete dynamical model, which has the form:

$$P_{n+1} = F(t_n, P_n).$$

This very general form implies that the population at the next discrete time depends on both the time, t_n, and previous population, P_n. This population model is known as a *nonautonomous discrete population model*. Our study below examines the simplest nonautonomous model, where $F(t_n, P_n)$ is the product of a linear time varying growth rate and the current population or

$$F(t_n, P_n) = (a + bt_n)P_n.$$

Animal populations often depend only on the existing population with limited time dependence. This leads to the general *autonomous discrete population model*:

$$P_{n+1} = F(P_n),$$

which states that the population at the next time only depends on the current population. One of the most important autonomous discrete population models is the *logistic growth model*. This model is very commonly used in animal studies, where the animal population is limited by resources. This model often demonstrates a saturation effect, where the population levels off at a *carrying capacity*. The logistic growth model assumes a Malthusian growth at low population density, then a decline in the growth rate for higher densities using a quadratic form. This logistic growth model has the form:

$$F(P_n) = aP_n - bP_n^2 = P_n + rP_n\left(1 - \frac{P_n}{M}\right),$$

where the linear term $a = 1 + r$ as in the Malthusian growth model and the quadratic term $b = \frac{r}{M}$. The functional form on the right above assumes a Malthusian growth rate of r with a carrying capacity of M.

An important feature of *autonomous growth models* is the existence of population *equilibria*. A population, P_e, is at *equilibrium* if successive populations remain the same, so $P_n = P_{n+1} = P_e$. The equilibria for a discrete dynamical model reduce to solving the algebraic equation

$$P_e = F(P_e).$$

In a *closed ecological model* (no immigration or emigration), there is almost always the *extinction equilibrium*, $P_e = 0$, The logistic growth model has two equilibria, the extinction equilibrium and the equilibrium at carrying capacity. The equilibria are found by solving:

$$P_e = P_e + rP_e\left(1 - \frac{P_e}{M}\right).$$

It follows that

$$0 = rP_e\left(1 - \frac{P_e}{M}\right) \qquad \text{or} \qquad P_e = 0, M,$$

assuming $r \neq 0$. Thus, the logistic growth model has a non-extinction equilibrium at the carrying capacity, $P_e = M$.

Example 4.2 (U. S. Population)**.** *This example studies the population of the U. S. from 1790 to present by fitting Malthusian, nonautonomous, and logistic growth models to the census data. The models are compared and examined to see what predictions result for the future.*

This example begins with the same techniques and analysis of Example 4.1. In this case, we fit the entire U. S. census data with the discrete Malthusian growth model:

$$P_{n+1} = (1+r)P_n, \qquad P_0 \quad \text{given},$$

where we assume that r is the **decade growth rate** and P_0 is the initial population at $n = 0$. The spreadsheet is set up with the data (Year and Population) in the first two columns, which are obtained from Table 4.1. Between these columns, we insert a column (*right click* Column B and accept *Insert*) for a variable t, which is measured in years after 1790. In Columns K and L, we insert our parameters P_0 and r for the Malthusian growth model and *create names*, say $p0m$ and rm, with reasonable guesses in the Cells $K1\!:\!L2$. In Column D, the Malthusian growth model is simulated. This is readily done by starting with '`= p0m`' in $D2$, then typing '`= (1 + rm)*D2`' in $D3$ and *pulling this down* for the period of simulation time desired. In Column E, the *square error* between the data and model is computed by typing '`= (C2 - D2)^2`' and *pulling this down* to the end of the data. In $E25$, the sum of square errors is computing by using Excel's AutoSum, $\sum$, or typing '`=SUM(E2:E24)`.' As before, we invoke Excel's *Solver* to minimize the sum of square errors *By Changing Variable Cells:* for $p0m$ and rm. The result is $P_0 = 16.3458$ and $r = 0.1460$ with a sum of square errors being $J = 2875.2$, so the best fitting Malthusian growth model over the entire U. S. census data is

$$P_{n+1} = 16.3456\,(1.1460)^n,$$

where n is in decades after 1790. This model starts with an initial population, which is not achieved in the actual census data until almost 50 years later. The growth rate is substantially below the average growth rate ($r_{ave} = 0.2233$). Thus, this model tracks the U. S. census data rather poorly, which is seen in the sum of square errors.

A nonautonomous growth model can be expressed as

$$P_{n+1} = (1 + k(t_n))P_n, \qquad P_0 \quad \text{given},$$

where $k(t_n) = a + b\,t_n$ and $t_n = 10\,n$ with n and t_n being decades and years after 1790, respectively. Note that this model has one more parameter than the simple Malthusian growth model. The function $k(t_n)$ is a linear function, so we need to find the linear least squares best fit to the growth rate of the U. S. population data. The growth rate (per decade), g, between successive census dates is computed by

$$g = \frac{P_{n+1}}{P_n} - 1.$$

To implement the procedure for finding the best nonautonomous growth model, we start a new Sheet (tab at bottom) for the spreadsheet described above. Copy the first **3** columns from the original page to the first **3** columns of the new Sheet. In Column D, the growth rate is computed for each decade by typing in $D2$, '`= C3/C2 - 1`' and *pulling down* this formula to $D23$, which is the decade before the last entry of the census data. Fig. 4.2 shows the graph of this growth rate data (Columns B and D). By *right clicking* on the data in the graph, we select *Add Trendline ...* and proceed to choose the *Linear Trendline Option*, including the option *Display Equation on Chart.* Excel produces the best fitting line given by

$$k(t_n) = -0.0014387\, t_n + 0.37436,$$

with $t_n = 10\,n$ and n being the decades after 1790. Fig. 4.2 shows a graph of $k(t_n)$ with the growth rate data and the mean growth rate. (The mean growth rate is $r_{ave} = 0.2233$ and is easily computed in Excel by typing '`=AVERAGE(F2:F23)`.')

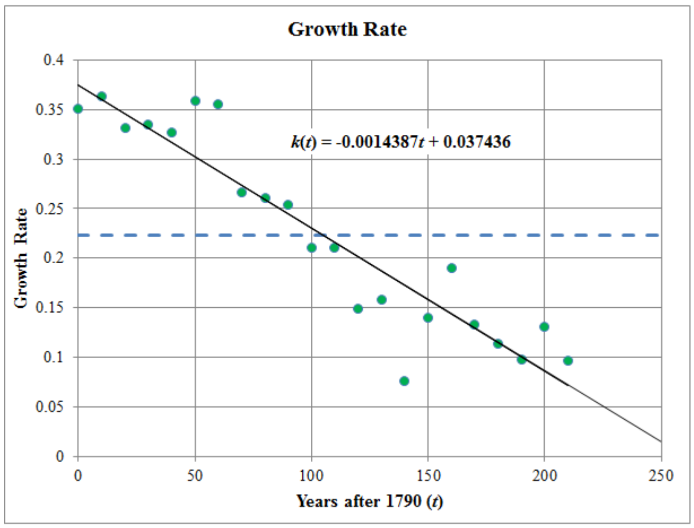

Figure 4.2: Graph showing the growth rate for the U. S. for each decade after 1790. The best linear fit is shown along with the mean growth rate (dashed line).

From the information above, the best fitting nonautonomous model with the growth data satisfies the equation:

$$P_{n+1} = (1.37436 - 0.014387\, n)P_n.$$

There remains the initial population, P_0, to be fit to the data with this nonautonomous model. We return to the original spreadsheet described above. In Columns F and G, the nonautonomous model and its sum of square errors with the data are formulated. If the *named variable*, $p0n$, is created in Cells $K4\!:\!L4$, then in $F2$, we enter '`= p0n`.' In $F3$, the formula '`= (1.37436 - 0.0014387*B2)*F2`' is entered, where $B2$ contained the time information and $F2$ is the population from the previous time. This formula is *pulled down* to simulate the nonautonomous model for whatever period of time desired. In $G2$, the *square error* between the data and model is computed. In Column G, the *square error* between the data and model is computed by typing '`= (C2 - F2)^2`' and *pulling this down* to the end of the data. Cell $G25$ contains the sum of square errors using Excel's AutoSum, $\sum$, or typing '`=SUM(G2:G24)`.' As before, we invoke Excel's *Solver* to minimize the sum of square errors *By Changing Variable Cells:* for $p0n$ only. The result is $P_0 = 3.7590$ with a sum of square errors being $J = 739.92$, which is substantially smaller than for the Malthusian growth model. Since this model has one additional parameter, it is expected to fit better, and the sum of square errors shows a significantly better fitting of this model to the census data.

The *logistic growth model* is the final model fit to the U. S. census data in this example. This model has the two parameters, growth, r, and carrying capacity, M, and the initial population, P_0. As in the previous two models, these parameters and the initial condition are fit with Excel's *Solver*. On the original spreadsheet, create the *named variables*, $p0l$, rl, and M in the Cells $K6\!:\!L8$ with reasonable initial guesses, like $p0l = 4$,

$rl = 0.25$, and $M = 400$. The model is simulated in Column H by entering '`= p0l`' in $H2$ and the formula '`= H2 + rl*H2*(1 - H2/M)`.' By *pulling down* this formula, the *logistic growth model* is simulated for any desired time period. In Column I, the *square error* between the data and model is computed by typing '`= (C2 - H2)^2`' and *pulling this down* to the end of the data. The sum of square errors is calculated in Cell $I25$ using Excel's AutoSum, $\sum$, or typing '`=SUM(I2:I24)`.' As before, we invoke Excel's *Solver* to minimize the sum of square errors *By Changing Variable Cells:* for $p0l$, rl,and M. The nonlinear least squares best fitting parameters for the U. S. census data are found to be $p0l = 8.5758$, $rl = 0.22445$, and $M = 451.75$. The sum of square errors is $J = 557.37$. This implies that the logistic growth model fits the census data best of these three models. The census data and the models are presented in Fig. 4.3.

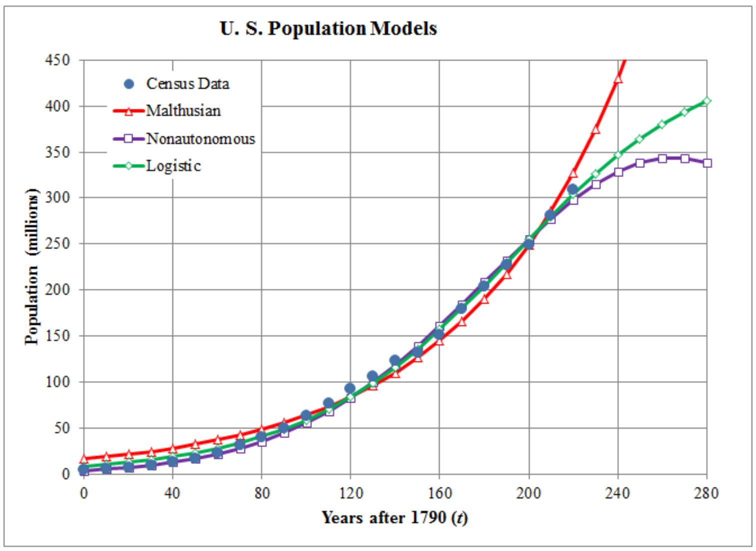

Figure 4.3: Graph showing the population for the U. S. for each decade after 1790. The best fitting Malthusian, nonautonomous, and logistic models are shown.

The models in the graph are simulated for an additional 60 years to show the differences in the projected future populations for the U. S. Not surprisingly, the Malthusian growth model grows rapidly to expand beyond the range of the graph with its exponential growth. This model was not expected to predict the future population of the U. S. very well. Both of the other models are shown to fit the census data quite well over the range of the simulations, including very close fits to the actual censuses taken in 2000 and 2010. The nonautonomous model fits the early census data very well, while the logistic growth model does better in the 20^{th} century. However, the nonautonomous and logistic growth models predict very different futures over the next century. Consider the nonautonomous growth function $k(t_n)$ and solve when $k(t_n) = 0$. This is the predicted year after 1790 when the growth rate reaches zero and the population starts declining. The solution is $t_n = 260.21$, which would be shortly after 2050. It follows that the nonautonomous model predicts that the population of the U. S. will achieve a maximum of a little over 343.4 million people between 2050 and 2060, then begin to decline. A number of developed countries around the world have seen this maximum population occur recently, with Italy being the first country to achieve this milestone in 2001. Fig. 4.3 shows the nonautonomous model peaking at this predicted population 260-270 years after 1790.

The logistic growth model shows the population of the U. S. continuing to rise, but the growth rate declining to zero. The best fitting parameters give the carrying capacity, $M = 451.7$, which means that this model predicts the population of the U. S. to level off in the future around 450 million people. Fig. 4.3 shows the population at 405 million in 2070 with significantly higher predictions than the nonautonomous model for most of this century. As noted earlier, it is generally expected that a nonautonomous model for human growth is better than the logistic growth model. However, neither model includes a term for immigration, which is significant for the U. S. population. Hence, better models are needed to obtain an accurate prediction for the future, which is one of the goals of the U. S. Census Bureau. Their combination of numerous models with careful statistical analysis, including the age structure of the population, provides the best future predictions. However, the nonautonomous and logistic growth models do give some reasonable estimates of the future.

Problem 4.3 (Nonautonomous Growth and Vietnam's Population). *(Lab121_F4_nonauto.pg) A discrete nonautonomous growth model is applied to census data for Vietnam from 1950 to 2000. The model assumes a declining linear growth, and it is used to predict future populations for Vietnam.*

Using data from the U. S. census bureau [18], the table below presents the population (in millions) for Vietnam. Human populations tend to have growth rates vary with time because of advancement of living conditions from technology. Below we develop a nonautonomous model for the population growth of Vietnam based on the following data.

Year	Population	Year	Population
1950	25.34	1980	53.66
1960	31.66	1990	66.66
1970	42.59	2000	78.53

a. We begin our modeling effort by finding the growth rate of Vietnam for each decade. Let 1950 correspond to $t = 0$. Find the growth rate for each decade using the data above by dividing the population from one decade by the population of the previous decade and subtracting 1 from this ratio. Associate each growth rate with the earlier value of t from the two census dates. Find the growth rate at $t = 30$ (the decade 1980-1990).

Growth at $t = 30$ is ____________.

Determine the average (mean) growth rate, r_{ave}, (per decade) from the data above.

Average growth rate $r_{ave} =$ ____________.

Find the best straight line through the growth data

$$k(t) = a + bt,$$

where

$a =$ ____________ and $b =$ ____________.

Assuming that the data is the best value, find the percent error between the model growth rate, $k(t)$, and the actual growth rate at $t = 30$.

Percent Error = ____________%

b. In your Lab Report, plot the growth data for Vietnam as a function of t (years after 1950). Include on this graph a line representing the average growth rate, r_{ave}, over the 50 years of data and the best fitting line through the data, $k(t)$. Briefly discuss what you observe about the growth rate for Vietnam. How well does $k(t)$ fit the growth rate data?

c. The Discrete Malthusian growth model is given by

$$P_{n+1} = (1 + r_{ave})P_n, \qquad P_0 = 25.34.$$

where r_{ave} is the average growth rate (computed above) and P_0 is the population in 1950, $t = 0$. Write the general solution to this model, where n is in decades. Your answer should include the values for r_{ave} and P_0, but depend on the variable n.

$P_n =$ ______________________

Find the sum of square errors between this discrete Malthusian growth model and the data in the table.

Sum of Square Errors for discrete Malthusian model = __________

Use this model to estimate the populations in 1960, 1980, and 2000. Then with the actual population being the best value, find the percent error from this Malthusian growth model

$P_1 =$ __________ Percent error = __________%

$P_3 =$ __________ Percent error = __________%

$P_5 =$ __________ Percent error = __________%

Use the model to predict the population in 2020 and 2050.

$P_7 =$ __________

$P_{10} =$ __________

d. The nonautonomous Malthusian growth model is given by

$$P_{n+1} = (1 + k(t_n))P_n, \qquad P_0 = 25.34.$$

where $k(t_n)$ is computed above (with $t_n = 10n$) and P_0 is again the population in 1950. Simulate this nonautonomous discrete dynamical model from 1950 to 2050 ($t_n \in [0, 100]$). Find the sum of square errors between this nonautonomous Malthusian growth model and the data in the table.

Sum of Square Errors for nonautonomous Malthusian model = __________

Use this model to estimate the populations in 1960, 1980, and 2000. Then with the actual population being accurate, find the percent error from this nonautonomous Malthusian growth model

$P_1 =$ __________ Percent error = __________%

$P_3 =$ __________ Percent error = __________%

$P_5 =$ __________ Percent error = __________%

Use the model to predict the population in 2020 and 2050.

$P_7 =$ __________

$P_{10} =$ __________

e. The nonautonomous Malthusian growth model has a declining linear growth rate. Hence, its growth rate will eventually go to zero, then become negative. According to this model, the population of Vietnam achieves a maximum during the decade following the time when the growth rate is zero. Use the growth rate function $k(t)$ to predict when the growth rate becomes zero.

$k(t_M) = 0$ when $t_M =$ __________

Simulate the nonautonomous Malthusian growth model until it achieves a maximum, and list the maximum value from the simulation along with the value of t_n when that maximum occurs.

Maximum population = ____________ with t_n = ____________

f. In your Lab Report, create a graph of the population data for Vietnam. Show the Malthusian growth and the nonautonomous Malthusian growth models on this graph with modeling projections for the period from 1950 to 2050. Briefly discuss how well these models predict the population over this period. How believable are the modeling projections for the future? List some strengths and weaknesses of each of the models and how you might obtain a better means of predicting the future population of Vietnam.

Problem 4.4 (Discrete and Logistic Growth Models). *(Lab121_G1_bac_grow) Discrete Malthusian and Logistic growth models are simulated and analyzed.*

This problem examines Discrete Malthusian and Logistic growth models, which are appropriate for studying simple organisms over limited time periods. The Malthusian growth model is given by the equation:

$$B_{n+1} = B_n + r\,B_n = (1+r)B_n,$$

where n is the time in minutes and r is the rate of growth. The Logistic growth equation is given by

$$B_{n+1} = B_n + r\,B_n\left(1 - \frac{B_n}{M}\right),$$

where M is the carrying capacity of the population.

a. Begin with a simulation of the Malthusian growth model starting with 1000 bacteria (or $B_0 = 1000$). Assume that the growth rate $r = 0.023$/min. Write an expression for the number of bacteria at each min, n.

$B_n =$ ______________________

Simulate this dynamical system for $n = 350$ min, then find the number of individuals at $n = 60$, 180, and 300 min (1, 3, and 5 hr).

$B_{60} =$ ____________ $B_{180} =$ ____________ $B_{300} =$ ____________.

How long does it take for this population to double? (Use your formula to solve exactly.)

Doubling time, $t_d =$ ____________ min.

b. Next we examine a population of bacteria that satisfies the Logistic growth law. Start again with $B_0 = 1000$ bacteria, but use a growth rate of $r = 0.026$/min. Assume that $M = 1,500,000$. Simulate this model for 350 min, then find the number of individuals after $n = 60$, 180, and 300 min (1, 3, and 5 hr).

$B_{60} =$ ____________ $B_{180} =$ ____________ $B_{300} =$ ____________.

Since we cannot solve the logistic growth model exactly, we cannot determine the exact doubling time as is possible to do for the Malthusian growth model. Use the simulation to determine the value of n when the logistic growth model first exceeds twice the initial population.

Approximate doubling time, $n_d =$ ____________ min.

The logistic growth model has a carrying capacity, which is the maximum sustainable population. Find how long it takes for this model to first exceed 0.5 times the maximum population of this model.

Approximate time for achieving 0.5 of carrying capacity, $n_M =$ ____________ min.

c. It is clear that the growth rate of the discrete Malthusian growth model is less than the growth rate for the discrete logistic growth model, so initially the discrete logistic growth model grows more rapidly. However, the simulations should show that eventually the discrete Malthusian growth model's population exceeds that of the discrete logistic growth model. Use your data to determine the first time that the population from Malthusian growth model exceeds the one growing according to the Logistic growth model.

Approximate time for discrete Malthusian population exceeding discrete logistic growth population, $n_e =$ ____________ min.

d. In your Lab Report, create a single graph with the populations of both bacterial cultures (Malthusian and Logistic) for n from 0 to 350. (Be sure to use lines to represent these simulations and not points, labeling which line represents which model.) Write a brief paragraph describing the shapes of the two population growth curves. Which model better represents a culture of bacteria on a fixed medium? Give a brief explanation for why the population from the discrete Malthusian growth model must eventually exceed the population of the discrete logistic growth model, despite the growth rate of the discrete logistic growth model being larger.

Logistic Growth Behavior

The examples above have shown the *logistic growth model* in the classic model behavior of populations growing exponentially at low density, then leveling off at a carrying capacity as the population grows. The discrete logistic growth model exhibits this behavior whenever the growth rate is not too large. Robert May[11] found that the discrete logistic growth model had some very interesting behaviors when the growth rate, r, is increased. As noted above, when this model has a low growth rate, then the population monotonically grows to the carrying capacity. As the growth rate increases, the model next exhibits a stable oscillatory behavior with the population oscillating around, but converging to the carrying capacity. Further increases in the growth rate show the population going through a series of *period doublings*. With increasing r, the population first oscillates with a stable period of two distinct populations. Next it oscillates with a stable period of four distinct populations, then eight, sixteen, etc. Eventually, at a certain value of r, the discrete logistic growth model goes into *chaos*. The behavior becomes very unpredictable. This is a very complex behavior and was quite unexpected for such a simple model. It also helps explain why one would not expect a closed form solution to the discrete logistic growth model. The problem below explores this behavior.

Problem 4.5 (Logistic Model)**.** *(Lab121_H2_log_grow.pg) Simulations are performed to observe the behavior of the logistic growth model as it goes from stable behavior to chaos.*

This problem studies the behavior of the discrete logistic growth model as the growth parameter varies. For certain parameter values, it is possible for this discrete model to exhibit chaotic behavior. The discrete logistic growth model satisfies:

$$P_{n+1} = f(P_n) = P_n + r\,P_n\left(1 - \frac{P_n}{M}\right).$$

This problem explores some of the complications that can arise as the parameter r varies.

a. Let $M = 10000$. The first step in studying this model is to find all equilibria (where the population stays the same). Determine the equilibria, P_e, by solving

$$P_e = f(P_e).$$

The equilibria are

$P_{1e} =$ ____________ and $P_{2e} =$ ____________.

b. Let $r = 0.89$ with $P_0 = 400$. Simulate the discrete logistic growth model for $n = 50$ generations. Give the values from your simulations at $n = 2, 5, 10, 20, 35$, and 50.

$P_2 =$ ____________, $P_5 =$ ____________, $P_{10} =$ ____________,

$P_{20} =$ ____________, $P_{35} =$ ____________, $P_{50} =$ ____________.

c. Now let $r = 1.69$ with $P_0 = 400$. Simulate the discrete logistic growth model for $n = 50$ generations. Give the values from your simulations at $n = 2, 5, 10, 20, 35$, and 50.

$P_2 =$ ____________, $P_5 =$ ____________, $P_{10} =$ ____________,

$P_{20} =$ ____________, $P_{35} =$ ____________, $P_{50} =$ ____________.

d. Next let $r = 2.15$ with $P_0 = 400$. Simulate the discrete logistic growth model for $n = 50$ generations. Give the values from your simulations at $n = 2, 5, 10, 20, 35$, and 50.

$P_2 =$ ____________, $P_5 =$ ____________, $P_{10} =$ ____________,

$P_{20} =$ ____________, $P_{35} =$ ____________, $P_{50} =$ ____________.

e. Next let $r = 2.51$ with $P_0 = 400$. Simulate the discrete logistic growth model for $n = 50$ generations. Give the values from your simulations at $n = 2, 5, 10, 20, 35$, and 50.

$P_2 =$ ____________, $P_5 =$ ____________, $P_{10} =$ ____________,

$P_{20} =$ ____________, $P_{35} =$ ____________, $P_{50} =$ ____________.

f. Next let $r = 2.57$ with $P_0 = 400$. Simulate the discrete logistic growth model for $n = 50$ generations. Give the values from your simulations at $n = 2, 5, 10, 20, 35$, and 50.

$P_2 =$ ____________, $P_5 =$ ____________, $P_{10} =$ ____________,

$P_{20} =$ ____________, $P_{35} =$ ____________, $P_{50} =$ ____________.

g. Place the letter of the behavior of the logistic growth model next to each r value listed below:

______ Model behavior $r = 2.15$	A. Chaos
______ Model behavior $r = 1.69$	B. Stable, Monotonic
______ Model behavior $r = 2.51$	C. Stable, Oscillatory
______ Model behavior $r = 0.89$	D. Periodic, Period 4
______ Model behavior $r = 2.57$	E. Periodic, Period 2
	F. None of the Above

h. Find a parameter value, r, that gives an oscillation with period 3. It has been shown that if the discrete logistic growth model gives a period 3 solution, then this discrete dynamical system has gone through chaos.

Period 3 oscillation when $r =$ ____________.

i. In your Lab Report, create a graph showing the simulations for $r = 0.89$ and 1.69, then write a brief description of what you observe in these simulations. Create another graph showing the simulations for $r = 2.15$ and 2.51, then write another brief description of what you observe in these simulations. Finally, create a graph showing the simulation for $r = 2.57$ and your period 3 simulation, then write a brief description of what you observe in these simulations.

Population Model with Immigration

The models described above are all considered *closed* population models. A *closed population model* means that the model accounts for all of its population from within the model, and it must have a *zero* or *extinction equilibrium*. However, when considering the populations of countries, then an important source of population change comes about from immigration or emigration. The U. S. population has a long history of immigration affecting its population demographics. The problem below examines the simplest population model with immigration.

Problem 4.6 (Immigration Model)**.** *(Lab121_H3_imm_log.pg) The population of the U. S. in the 20th century is fit with a discrete Malthusian growth model, a Malthusian growth model with immigration, and a logistic growth model. These models are compared for agreement with data and used to project future behavior of the population.*

We have examined the discrete Malthusian, a nonautonomous Malthusian, and the logistic growth models for several countries. This question examines three population models for studying the population of the U. S. during the 20th century. For comparison we include the discrete Malthusian growth model. The two other models are a Malthusian growth model with immigration (or emigration) and the logistic growth model. Use Table 4.1 for census data for the U. S. from 1910 to 2010.

a. Consider the census data for the U. S. from 1910 to 1990 in Table 4.1. (**Note** that the calculation to find the best fitting parameters does not include the data from 2000 or 2010.) The discrete Malthusian growth model is given by

$$P_{n+1} = (1 + r)P_n,$$

where r is the decade growth rate and P_0 is the initial population in 1910. Use Excel's *Solver* to find the best fitting parameters, r and P_0. For initial estimates of the parameters, consider r to be the average growth rate over the years and P_0 to be the actual population in 1910. Compute the least sum of square errors between the census data and the discrete Malthusian growth model for the years 1910 to 1990. The best values for the initial population and growth rate (per decade) are:

$P_0 =$ __________ and $r =$ __________.

Sum of Square Errors = __________.

Use this model to find the population at 1950, 2000, and 2010, then determine the percent error from the population given in Table 4.1 (assuming that the value in Table 4.1 is the most accurate):

Model Population in 1950 = __________ and Percent Error in 1950 = __________.

Model Population in 2000 = __________ and Percent Error in 2000 = __________.

Model Population in 2010 = __________ and Percent Error in 2010 = __________.

Finally, project this model to find the population at 2050.

Model Population in 2050 = __________.

b. In the U. S., immigration has played an important role. During the 20^{th} century, it has been more or less tightly regulated and maintained a relatively constant value. The discrete Malthusian growth model with immigration is given by

$$P_{n+1} = (1 + r)P_n + \mu,$$

where we assume that r is the decade growth rate, μ is the immigration rate (assumed constant), and P_0 is the initial population in 1910. Use Excel's *Solver* to find the best fitting parameters, r, μ, and P_0. For initial estimates of r, μ, and P_0, take r to be the value computed in Part a, $\mu = 4.5$, and P_0 to be the actual population in 1910. Compute the least sum of square errors between the census data and the discrete Malthusian growth model with immigration for the years 1910 to 1990. The best values for r, μ, and P_0 are:

$P_0 =$ __________ and $r =$ __________ and $\mu =$ __________.

Sum of Square Errors = __________.

Use this model to find the population at 1950, 2000, and 2010, then determine the percent error from the population given in Table 4.1 (assuming that the value in Table 4.1 is the most accurate):

Model Population in 1950 = __________ and Percent Error in 1950 = __________.

Model Population in 2000 = __________ and Percent Error in 2000 = __________.

Model Population in 2010 = __________ and Percent Error in 2010 = __________.

Finally, project this model to find the population at 2050.

Model Population in 2050 = __________.

c. The two previous models grow without bound. One question is the value at which the U. S. population will level off, and there are many estimates on what this might be. The logistic growth model is a classic model of population growth that grows like a Malthusian growth model for low population densities. However, the growth slows with increasing density, and the logistic growth model has the property of leveling off at the carrying capacity of the population. The logistic growth model is given by the discrete dynamical model of the form:

$$P_{n+1} = P_n + rP_n\left(1 - \frac{P_n}{M}\right),$$

where we assume that r is the growth rate, M is a significant constant relating to density dependent decreases in growth, and P_0 is the initial population in 1910. Use Excel's *Solver* to find the best fitting parameters, r, M, and P_0. For initial estimates of r, M, and P_0, take r to be the value computed in Part a, $M = 700$, and P_0 to be the actual population in 1910. Compute the least sum of square errors between the census data and the logistic growth model for the years 1910 to 1990. The best values for r, M, and P_0 are:

$P_0 =$ ____________ and $r =$ ____________ and $M =$ ____________.

Sum of Square Errors = ____________.

Use this model to find the population at 1950, 2000, and 2010, then determine the percent error from the population given in Table 4.1 (assuming that the value in Table 4.1 is the most accurate):

Model Population in 1950 = ____________ and Percent Error in 1950 = ____________.

Model Population in 2000 = ____________ and Percent Error in 2000 = ____________.

Model Population in 2010 = ____________ and Percent Error in 2010 = ____________.

Finally, project this model to find the population at 2050.

Model Population in 2050 = ____________.

The carrying capacity is an equilibrium population. The equilibria of a discrete population model are found by substituting $P_e = P_n$ and $P_e = P_{n+1}$. Make this substitution into the logistic growth model, leaving you with a quadratic equation. Solve that quadratic equation to find two equilibria.

$P_{e1} =$ ____________ and $P_{e2} =$ ____________.

The larger of these is the carrying capacity.

d. In your Lab Report, create a single graph with all three models and the census data on the interval 1910 to 2050. Looking at the three models above, determine which model you believe best predicts the population for the year 2050. Which model do you believe is the best and why? What are the advantages and disadvantages of using one model over the other? Describe two ways that you could improve the best model to make a better prediction for 2050. Does the immigration rate in the second model seem realistic for the U. S.? Is the carrying capacity that you computed believable for the U. S.? How might you obtain a better estimate of the immigration rate and carrying capacity?

3 BREATHING MODEL

There are numerous other applications of discrete dynamical models, other than the population models illustrated above. The physiological process of breathing is certainly a discrete process[10]. *Pulmonary ventilation* or breathing is the first step to bringing oxygen to the cells of the body and removing the metabolic waste product, carbon dioxide. Contracting the muscles of the diaphragm results in an inflow of fresh air or *inspiration*, while relaxation of these muscles or contraction of the abdominals causes *expiration* of air with the waste product CO_2. During normal respiration, the lungs exchange about 500 ml of air 12 times a minute. This is the *tidal volume* of air inspired or expired. In young adult males, there is an *inspiratory reserve volume* of about 3000 ml that can be inspired above the tidal volume, while the *expiratory reserve volume* is about 1100 ml, which can be forcefully expired. The *vital capacity* includes all of the above yielding about 4600 ml. Well-trained athletes may have values 30-40% higher, while females generally have 20-25% less for the quantities listed above. The *residual volume* represents the amount of air that cannot be expelled even by forceful expiration and averages about 1200 ml. The *functional residual capacity* is the amount of air that remains behind during normal breathing, which amounts to 2300 ml. Fig. 4.4 diagrams these physiological quantities.

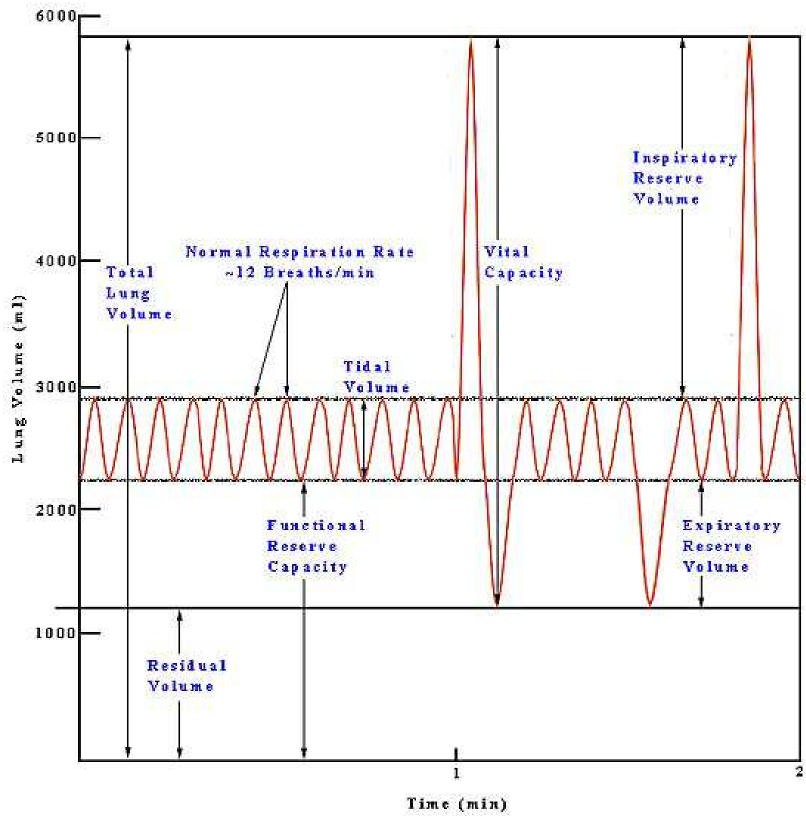

Figure 4.4: Diagram showing the graph of different lung air volumes and capacities.

There are several respiratory diseases. The respiratory muscles can be damaged by spinal paralysis or poliomyelitis, which can decrease the vital capacity to as low as 500 ml, barely enough to maintain life. Pulmonary compliance reduces vital capacity in diseases like tuberculosis, emphysema, chronic asthma, lung cancer, chronic bronchitis, cystic fibrosis, or fibrotic pleurisy. Several of the diseases above and heart disease can cause pulmonary edema, which decreases vital capacity from fluid build up in the lungs. It is important to know about the physiological state of the lungs to determine what diseases are affecting any particular patient.

One non-invasive method of determining the tidal volume and the functional residual capacity is for the subject to breathe a mixture including an inert gas, which for this experiment is Argon (Ar). The subject breathes the mixture until the lungs are essentially filled with this mixture, then the physiologist measures the amount of the inert gas in a series of breaths after the subject is removed from the gas mixture to normal air.

Let c_n be the concentration of Ar in the n^{th} breath of the subject. The physiological parameters needed for the discrete dynamical model are V_i for the tidal volume (air normally inhaled and exhaled), V_r for the functional residual volume, and γ for the concentration of Ar in the atmosphere. Define the fraction of atmospheric air exchanged in each breath, q, and fraction of air remaining in the lungs, $1-q$, as

$$q = \frac{V_i}{V_i + V_r} \qquad \text{and} \qquad 1 - q = \frac{V_r}{V_i + V_r},$$

respectively. Upon exhaling, there remains behind the functional residual volume, which contains the amount of Ar given by $V_r c_n$. The inhaled air during this cycle contains the amount of Ar given by $V_i\gamma$. (Quantities or amounts of Ar are given by the volume times concentration, and it's the amounts that are conserved.) Thus, the amount of Ar in the next breath is given by

$$V_r c_n + V_i \gamma.$$

To find the concentration in the next breath we divide by the total volume, $V_i + V_r$, then use the definition of q. Thus, the linear discrete dynamical model for breathing an inert gas is:

$$c_{n+1} = \frac{V_r c_n}{V_i + V_r} + \frac{V_i \gamma}{V_i + V_r} = (1-q)c_n + q\gamma. \tag{4.1}$$

Problem 4.7 (Breathing Model)**.** *(Lab121_G2_breathing.pg) A linear discrete model is examined for determining vital lung functions of normal and diseased subjects following breathing an enriched source of argon gas.*

If c_n represents the concentration of the inert gas argon (Ar) in the lungs, then a mathematical model for breathing is given by the linear discrete dynamical model, (4.1), where q is the fraction of the lung volume exchanged with each breath and $\gamma = 0.0093$ (fraction of Ar in dry air) is the concentration of Ar in the atmosphere. Normal breathing usually exchanges a volume of air, known as the *tidal volume*, V_i. The space remaining in the lung after exhaling from a normal breath is known as the *functional residual volume*, V_r. The fraction of air exchanged is $q = \frac{V_i}{V_i+V_r}$.

a. Assume that a normal subject breathes an enriched mixture of air that contains 10% Ar, so that $c_0 = 0.1$ (fraction of Ar in dry air). Suppose that the tidal volume is measured at $V_i = 505$ ml for this subject, while another measurement gives the functional residual volume, $V_r = 2350$ ml. Determine the fraction of the lung volume q exchanged for this subject.

$q =$ __________

Find the concentration of Ar in the first, third, fifth, and tenth breaths

$c_1 =$ __________ $c_3 =$ __________ $c_5 =$ __________ $c_{10} =$ __________

What is the equilibrium concentration, c_e, of Ar in this subject?

$c_e =$ __________.

Determine how many breaths are required until the concentration of Ar drops below 0.01.

Number of Breaths = __________.

b. In your Lab Report, create a graph showing the concentration of Ar in the first 10 breaths.

c. A patient with emphysema is given the same mixture of Ar (so again $c_0 = 0.1$ (fraction of Ar in dry air). The tidal volume for this patient is measured at $V_i = 205$ ml. The concentration of Ar in the first breath in found to contain 0.0899 (fraction of Ar in dry air) for this patient or $c_1 = 0.0899$. Find the fraction of the lung volume exchanged q and the functional residual volume, V_r.

$q =$ ____________ $V_r =$ ____________.

Find the concentration of Ar in the second, third, fifth, and tenth breaths for this subject

$c_2 =$ ____________ $c_3 =$ ____________ $c_5 =$ ____________ $c_{10} =$ ____________.

What is the equilibrium concentration, c_e, of Ar in this subject?

$c_e =$ ____________.

Determine how many breaths are required until the concentration of Ar drops below 0.01.

Number of Breaths = ____________.

d. In your Lab Report, create a graph showing the concentration of Ar in the first 10 breaths of this subject with emphysema. Write a brief paragraph discussing the differences between the breathing of a normal subject and a patient with emphysema based on the results above.

CHAPTER 5:

DERIVATIVE AND APPLICATIONS

Keywords: Derivative, tangent line, maximum and minimum, point of inflection

This chapter begins with examples and problems designed to motivate the derivative. Geometrically, the derivative is the slope of a tangent line to a curve. Maple commands are introduced to symbolically differentiate functions. From a biological perspective the derivative often arises when an application considers growth, velocity, or kinetic rates of reactions, so several problems examine these phenomena. Derivatives also prove useful for finding a maximum or minimum, hence help locate optimal solutions to problems. Maple provides an easy and accurate method for finding these critical points.

1 INTRODUCTION TO THE DERIVATIVE

This section begins with a geometric interpretation of the *derivative*. Subsequently, the formal definition of the *derivative* is provided. Maple commands are shown to help with the algebra symbolically and easily differentiate functions accurately. Several problems for growth and velocity are presented where approximations of the derivative are used.

1.1 THE DERIVATIVE AS A TANGENT LINE

Perhaps the easiest way to understand a derivative is to think of the derivative as the value of the slope of the *tangent line* for a curve. However, we have seen that the slope of a line is computed with two points, while a *tangent line* intersects a curve at exactly one point. Hence, it is best to consider a sequence of *secant lines*, which approach the *tangent line* for a point on a curve. This introduces us to a *limiting* process necessary for obtaining the tangent line.

Example 5.1 (Tangent Line)**.** *Consider the function*

$$y = x^2.$$

Find the equation of the **tangent line** *at the point* $(1, 1)$ *on the graph. The* **derivative** *of* $y(1)$ *is the value of the slope of this tangent line.*

A *secant line* is found by taking two points on the curve and finding the equation of the line through those points. The Fig. 5.1 starts with the secant line through the points $(1, 1)$ and $(2, 4)$. This line has a slope of $m = 3$, and its equation is $y = 3x - 2$. The next pair of points on the curve $y = x^2$ is $(1, 1)$ and $(1.5, 2.25)$. This line has a slope of $m = 2.5$, and its equation is $y = 2.5\,x - 1.5$. The secant line through the points $(1, 1)$ and $(1.1, 1.21)$ has a slope of $m = 2.1$, and its equation is $y = 2.1\,x - 1.1$. To which line equation are we approaching as the "mobile" point gets closer to the fixed point $(1, 1)$?

There is a more formal way to answer this question. Suppose we take $x = 1 + h$ for some small h. With $y = x^2$, the corresponding $y-$value is $y = (1 + h)^2 = 1 + 2h + h^2$. The slope of the secant line through this point and the point $(1, 1)$ is

$$m(h) = \frac{(1 + 2h + h^2) - 1}{(1 + h) - 1} = \frac{2h + h^2}{h} = 2 + h,$$

and the formula for this secant line is

$$y = (2 + h)x - (1 + h).$$

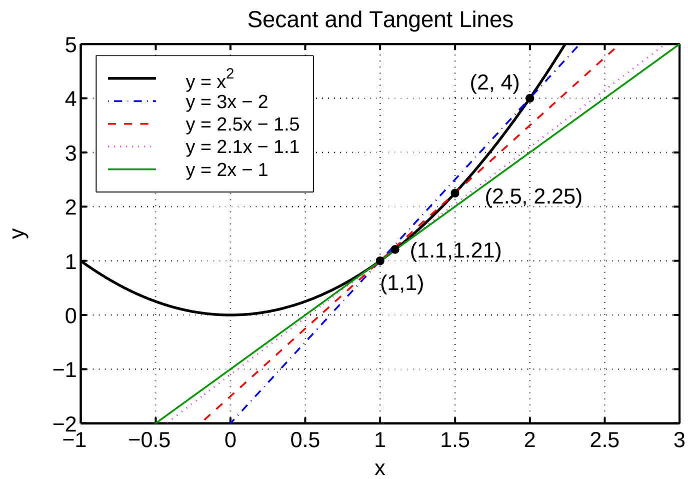

Figure 5.1: Graph showing .

As h gets very small, the secant line gets very close to the tangent line. Letting $h = 0$ in the equation above gives the tangent line for $y = x^2$ at $(1, 1)$, and it is

$$y = 2\,x - 1.$$

The slope of the tangent line, $m = 2$, **is the value of the derivative** of $y = x^2$ at $x = 1$.

Maple for Computing Derivatives

The definition of the *derivative* of a function, $f(x)$, at a point, $x = a$, is given by the limit:

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h},$$

provided this limit exists. This definition is the limit of finding the slope of the secant lines through the points $(a, f(a))$ and $(a+h, f(a+h))$ for different values of h. The slope of the secant line is given by

$$m(h) = \frac{f(a+h) - f(a)}{h}.$$

Geometrically, as $h \to 0$, the *secant line* approaches the *tangent line*. The slope of the tangent line at $x = a$ gives the value of the *derivative* of $f(x)$ at $x = a$ or $f'(a)$, which is reflected in the definition of the derivative.

Needless to say, the algebra for evaluating the slope of the secant line can be cumbersome. Maple's ability as a symbolic algebra package simplifies this computational aspect of the problem.

Example 5.2 (Maple for Secant and Tangent Lines)**.** *Consider the function*

$$f(x) = \frac{1}{(x+3)^3}.$$

Find the slope, $m(h)$, of the secant line between the points on the graph with $x = -2$ and $x = -2 + h$. The **derivative** *of $f(-2)$ (or $f'(-2)$) is the limit as $h \to 0$ of $m(h)$.*

For this example, we begin by entering the function into Maple:

$f := x \to \frac{1}{(x+3)^3}$

The slope function, $m(h)$, for $f(x)$ at $x = -2$ is found below and simplified by Maple:

$m := h \to \frac{(f(-2+h)-f(-2))}{h}$; $\quad m(h)$; $\quad simplify(\%)$;

This results in Maple giving the expression

$$m(h) = -\frac{(h^2 + 3h + 3)}{(1+h)^3}.$$

In this case it is easy to see what is the limit as $h \to 0$, but Maple does solve the limit problem with the command:

$limit(m(h), h = 0)$;

which gives $f'(-2) = -3$.

Maple *diff* Command

Using the definition to find the derivative for all functions would be onerous. However, Maple has a simple command to find the derivative, the *diff* command. If we consider the function in Example 5.2, then the Maple command is:

$diff(f(x), x)$;

which yields

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

Maple *subs* Command

Often we want the specific value of a derivative, such as a growth rate at a particular time. In this case, we need to substitute a particular t-value into our expression for the derivative. Maple uses the substitution command, *subs*. Alternately, Maple has the unapply command, *unapply*, which converts an explicit expression into a function.

As an example, suppose we have the length of a fish satisfying:

$$L(t) = 172.5(1 - e^{-0.5243t}).$$

The growth rate is given by the derivative, which we'll store as an expression, dL. The Maple commands for entering the function and finding its derivative are:

$L := t \to 172.5 \cdot (1 - exp(-0.5243 \cdot t))$;
$diff(L(t), t)$;

Evaluation of the derivative at $t = 7.5$ is found with either of the two lines below:

$subs(t = 7.5, dL)$; $evalf(\%)$;
$Lp := unapply(dL, t)$; $Lp(7.5)$

Either line produces the result that $L'(7.5) = 1.772615$. The first line gives the single point substitution, while the result with the *unapply* command produces the function $L'(t)$, which could be used for other operations.

Problem 5.1 (Tangent Lines and Derivative)**.** *(Lab121_J1_tan_lines.pg) This problem investigates the concept of a derivative from the geometric perspective of limiting secant lines going to a tangent line. It also allows an experimental exploration of the rules of differentiation using Maple.*

a. Consider the function:

$$f(x) = \frac{3}{(4-x)^2}.$$

We want to investigate the derivative of $f(x)$ at $x = 3$. This line always passes through the point $(3, f(3))$. Let $h = 0.5$ and find the slope of the secant line, $m(h) = \frac{f(3+h)-f(3)}{h}$.

Slope of secant line $m(0.5) =$ ____________.

The equation of the secant line passing through $(3, f(3))$ and $(3 + 0.5, f(3 + 0.5))$ satisfies:

$y =$ __________________________.

Let $h = 0.2$ and find the slope of the secant line

Slope of secant line $m(0.2) =$ ____________.

The equation of the secant line passing through $(3, f(3))$ and $(3 + 0.2, f(3 + 0.2))$ satisfies:

$y =$ __________________________.

Let $h = 0.1$ and find the slope of the secant line

Slope of secant line $m(0.1) =$ ____________.

The equation of the secant line passing through $(3, f(3))$ and $(3 + 0.1, f(3 + 0.1))$ satisfies:

$y =$__________________________.

Let $h = 0.05$ and find the slope of the secant line

Slope of secant line $m(0.05) =$ ____________.

The equation of the secant line passing through $(3, f(3))$ and $(3 + 0.05, f(3 + 0.05))$ satisfies:

$y =$ __________________________.

b. Find the general formula for the slope of the secant line through $x = 3$ and $x = 3 + h$. (Your answer below will include the variable h.)

$m(h) =$ ____.

The slope of the tangent line is found by taking the slope of the secant line and letting h tend toward zero. This becomes the derivative of $f(x)$ at $x = 3$. Find the derivative of $f(x)$, $f'(x)$, at $x = 3$.

$f'(3) =$ ____________.

The equation of the tangent line passing through $(3, f(3))$ satisfies:

$y =$ __________________________.

c. In your Lab Report, create a graph of $f(x)$. To this graph of the function, add the four secant lines and the tangent line found above for $x \in [0, 3.95]$. Limit the range, so that $y \in [-10, 20]$.

d. Maple is a powerful symbolic algebra program. This problem allows one to explore the rules of differentiation by looking for patterns from the output of Maple's *diff* command. This gives an experimental approach to finding derivatives. Differentiate the following functions using Maple.

If $f(x) = x^2 + 6x - 11$, then its derivative is

$f'(x) =$ ____.

If $g(x) = (x^2 + 6x - 11)^5$, then its derivative is

$g'(x) =$ ____.

If $h(x) = e^{-6x}$, then its derivative is

$h'(x) =$ ____.

If $k(x) = \ln(x + 4)$, then its derivative is

$k'(x) =$ ____.

e. In your Lab Report, write a brief description of what you observe upon taking the derivative in each of the above cases.

1.2 AVERAGE GROWTH RATE OR VELOCITY

The classical use of the derivative is for velocity. It is also commonly used for finding growth rates. If $h(t)$ is the position of an object, then the average velocity between times t_1 and t_2 satisfies:

$$v_{ave} = \frac{h(t_2) - h(t_1)}{t_2 - t_1}.$$

A similar expression can be formed to determine an average growth rate. This expression is very similar to the expression above for the slope of a secant line.

Problem 5.2 (Weight and Height of Girls)**.** *(Lab121_I2_ht_wt.pg) Data on the age, height, and weight of normal girls are presented. An allometric modeling compares the relationship between height and weight. Graphs are created for the rate of growth in both height and weight for girls ages 0 to 18.*

Pediatricians monitor for normal growth of children by the annual measurement of height and weight. These are expected to increase annually, the growth curve paralleling a standardized curve. In the material introducing the idea of a derivative, there are data on juvenile heights from birth to age 18. Table reftable:HtWt of both heights and weights for American girls in the 90^{th} percentile. (http://www.kidsgrowth.com/resources/articledetail.cfm

a. We begin this problem by using allometric modeling to compare early childhood with later childhood. In particular, use the data from age 0 to 5 to find an allometric relationship between the height of a child and her weight. The best fitting power law model of the form

$$w = ah^k.$$

age (years)	Height (cm)	Weight (kg)	age (years)	Height (cm)	Weight (kg)
0	52	3.8	8	133	33.2
0.25	62	6.4	9	141	37.3
0.5	70	8.4	10	147	43.2
0.75	75	9.8	11	153	50
1	79	10.7	12	159	55.5
1.5	86	12.3	13	165	61.4
2	91	13.6	14	168	65.9
3	99	15.9	15	170	69.1
4	107	18.2	16	171	70.9
5	114	21.4	17	172	71.8
6	121	25	18	172	71.8
7	128	28.2			

Table 5.1: The heights and weights for American girls in the 90^{th} percentile.

The best fit coefficients, a and k, found by Excel are

$a =$ __________ and $k =$ __________.

Next use the data for height and weight to find an allometric model for girls between ages 4 and 18. If again,

$$w = ah^k.$$

The best fit coefficients, a and k, found by Excel are

$a =$ __________ and $k =$ __________.

These models predict that a girl in the 90^{th} percentile at age 2 with a height 91 cm should weigh

$w =$ __________ kg,
which gives a percent error of (compared to the actual data, the best value)

Percent error = __________%.

These models predict that a girl in the 90^{th} percentile at age 9 with a height 141 cm should weigh

$w =$ __________ kg,
which gives a percent error of

Percent error = __________%.

These models predict that a girl in the 90^{th} percentile at age 1 weighing 10.7 kg should measure

$h =$ __________ cm,
which gives a percent error of (compared to the actual data, the best value)

Percent error = __________%.

These models predict that a girl in the 90^{th} percentile at age 12 weighing 55.5 kg should measure

$h =$ __________ cm,

which gives a percent error of

Percent error = ____________%.

b. In your Lab Report, create the graphs for each of your allometric models over the given age ranges. Give a physiological explanation for the coefficient k that you obtained for the model spanning the ages 4-18. Why doesn't this value of k hold for the earlier years of child development? Give a brief discussion of how well your model fits the data.

c. In biology, one of the most common uses of the derivative is a measurement of growth. We want to examine the data in Table reftable:HtWt to determine the changes in growth rate at different ages.

The average rate of growth in height over a time period $t \in [t_0, t_1]$ is easily computed by the formula:

$$g_{ht}(t_0) = \frac{h(t_1) - h(t_0)}{t_1 - t_0},$$

where we associate the growth rate, g_{ht}, with the earlier time $t = t_0$. Use Table 5.1 to compute the average rate of growth in height (in cm/yr) for each of the following age ranges:

Fill in the following table:

Growth from age 0 to 3 months: $g_{ht} =$ ____________cm/yr	Growth from age 0 to 6 months: $g_{ht} =$ ____________cm/yr
Growth from age 0 to 1 year: $g_{ht} =$ ____________cm/yr	Growth from age 1 to 2 years: $g_{ht} =$ ____________cm/yr
Growth from age 0 to 5 year: $g_{ht} =$ ____________cm/yr	Growth from age 2 to 5 years: $g_{ht} =$ ____________cm/yr
Growth from age 5 to 7 year: $g_{ht} =$ ____________cm/yr	Growth from age 10 to 12 years: $g_{ht} =$ ____________cm/yr
Growth from age 12 to 15 year: $g_{ht} =$ ____________cm/yr	Growth from age 15 to 18 years: $g_{ht} =$ ____________cm/yr

d. In your Lab Report, create a graph of height versus age, then create another graph of rate of change in height versus age (much like the graphs seen in the text). Find each growth rate (change of height) from successive values of height in Table 5.1. Associate the rate of change of height with the earlier age. What happens with the rate of change in height? Describe the graph for the rate of height gain over the early years (0-3), the ages 3-12, then adolescence (13-18).

e. Similarly, The average rate of growth in weight over a time period $t \in [t_0, t_1]$ is easily computed by the formula:

$$g_{wt}(t_0) = \frac{w(t_1) - w(t_0)}{t_1 - t_0},$$

where we associate the growth rate, g_{wt}, with the earlier time $t = t_0$. Use Table 5.1 to compute the average rate of growth in weight (in kg/yr) for each of the following age ranges:

Fill in the following table:

Growth from age 0 to 3 months: $g_{wt} =$ __________ cm/yr	Growth from age 0 to 6 months: $g_{wt} =$ __________ cm/yr
Growth from age 0 to 1 year: $g_{wt} =$ __________ cm/yr	Growth from age 1 to 2 years: $g_{wt} =$ __________ cm/yr
Growth from age 0 to 5 year: $g_{wt} =$ __________ cm/yr	Growth from age 2 to 5 years: $g_{wt} =$ __________ cm/yr
Growth from age 5 to 7 year: $g_{wt} =$ __________ cm/yr	Growth from age 10 to 12 years: $g_{wt} =$ __________ cm/yr
Growth from age 12 to 15 year: $g_{wt} =$ __________ cm/yr	Growth from age 15 to 18 years: $g_{wt} =$ __________ cm/yr

f. In your Lab Report, create a graph of weight versus age, then create another graph of rate of change in weight versus age (much like the graphs seen in the notes). Find each growth rate (change of weight) from successive values of weight in Table 5.1. Associate the rate of change of weight with the earlier age. What happens with the rate of change in weight? Describe the graph for the rate of weight gain over the early years (0-3), the ages 3-12, then adolescence (13-18) and compare these rate of changes to the ones you found for the rate of change in height in Part d.

Problem 5.3 (Flight of a Ball)**.** *(Lab121_I1_vert_ball.pg) Data for a vertically thrown ball is fit and analyzed. Average velocities are computed to provide insight into the understanding of the derivative.*

A ball is thrown vertically and data are collected at various times in its flight.

Time (sec)	Distance (m)	Time (sec)	Distance (m)
0.5	7.7	2	16.3
1	13	2.5	14.3
1.5	16	3	9.9

Table 5.2: The heights and times for the flight of a ball.

Assume that air resistance can be ignored, then the height of the ball satisfies the quadratic equation:

$$h(t) = v_0 t - \frac{g}{2} t^2,$$

due to gravity. (Note: There is no constant term as we are assuming that the height of the ball is zero at $t = 0$.)

a. Use the Excel's *Trendline* to find the best constants v_0 and g that fit the data in the table. (Remember that when you are using *Trendline*, you must decide if your graph passes through the origin. Does this one? (Yes or No) __________)

$v_0 =$ __________ m/sec and $g =$ __________ m/sec^2

Find the time that your model predicts the ball will hit the ground.

Time ball hits the ground = ____________sec

Find how high the ball goes, and find the time that it reaches this highest point.

Maximum height = ____________m and Time of Maximum height = ____________sec

b. In your Lab Report, create a graph of the quadratic function and the data. How well does the model fit the data?

c. The average velocity between two times t_1 and t_2 is given by the formula:

$$v_{ave} = \frac{h(t_2) - h(t_1)}{t_2 - t_1}.$$

Compute the average velocity with the function $h(t)$ that you found above between each of the following pairs of times:

Between $t_1 = 1$ and $t_2 = 2$: $v_{ave} =$ ____________m/sec	Between $t_1 = 1$ and $t_2 = 1.5$: $v_{ave} =$ ____________m/sec
Between $t_1 = 1$ and $t_2 = 1.1$: $v_{ave} =$ ____________m/sec	Between $t_1 = 1$ and $t_2 = 1.01$ $v_{ave} =$ ____________m/sec
Between $t_1 = 2$ and $t_2 = 3$: $v_{ave} =$ ____________m/sec	Between $t_1 = 2$ and $t_2 = 2.2$: $v_{ave} =$ ____________m/sec
Between $t_1 = 2$ and $t_2 = 2.05$: $v_{ave} =$ ____________m/sec	Between $t_1 = 2$ and $t_2 = 2.002$: $v_{ave} =$ ____________m/sec
Between $t_1 = 2.8$ and $t_2 = 3$: $v_{ave} =$ ____________m/sec	Between $t_1 = 2.9$ and $t_2 = 3$: $v_{ave} =$ ____________m/sec
Between $t_1 = 3$ and $t_2 = 3.01$: $v_{ave} =$ ____________m/sec	Between $t_1 = 2.999$ and $t_2 = 3$: $v_{ave} =$ ____________m/sec

d. The velocity of the ball at a given time is the derivative of the height function at that time. Use the differentiation techniques or Maple to compute the derivative of $h(t)$,

$$h'(t) = v(t).$$

$v(t) =$ ____m/sec

Evaluate the velocity at $t = 1, 2, 3$

$v(1) =$ ____________m/sec, $v(2) =$ ____________m/sec, and $v(3) =$ ____________m/sec

e. In your Lab Report, write a brief paragraph describing how the computed velocities in Part d compare to the average velocities computed in Part c.

f. For this part of the problem, you again compute some average velocities:

$$v_a(t_m) = \frac{h(t_2) - h(t_1)}{t_2 - t_1} \quad \text{where} \quad t_m = \frac{t_1 + t_2}{2}.$$

Create the following coordinate pairs, $(t_m, v_a(t_m))$, given values of t_1 and t_2.

For $t_1 = 0$ and $t_2 = 0.02$, then $(t_m, v_a(t_m)) = ($__________, __________$)$.

For $t_1 = 0.99$ and $t_2 = 1.01$, then $(t_m, v_a(t_m)) = ($__________, __________$)$.

For $t_1 = 1.99$ and $t_2 = 2.01$, then $(t_m, v_a(t_m)) = ($__________, __________$)$.

For $t_1 = 2.99$ and $t_2 = 3.01$, then $(t_m, v_a(t_m)) = ($__________, __________$)$.

g. In your Lab Report, use the coordinate pairs found above to create a graph of $v_a(t)$ versus t using these data. Use Trendline (or any other method) to find the equation of this graph.

$v_a(t) =$ ______________________________ m/sec

The v_a-intercept = __________ m/sec and the t-intercept = __________ sec

In your Lab Report, describe the graph that you have produced. Compare this equation to the equation of the derivative $h'(t)$ that you obtained above.

Problem 5.4 (Growth of a Fish). *(Lab121_I4_fish_cm.pg) The von Bertalanffy's equation is used for estimating the length of fish with some fish data to find growth in length of a fish.*

The growth of fish has been shown to satisfy a model given by the von Bertalanffy equation:

$$L(t) = L_\infty(1 - e^{-bt}),$$

where L_∞ and b are constants that fit the data.

a. Below are growth data for the Albacore (*Thunnus alalunga*)[6, 16].

Age (yr)	Length (cm)	Age (yr)	Length (cm)
1	41	6	110
2	63	7	113
3	82	8	116
4	98	9	118
5	101	10	119

Table 5.3: The age and lengths of Albacore.

Find the least squares best fit of the data to the von Bertalanffy equation above. Give the values of the constants L_∞ and b and write the model with these constants. Include the value of the least sum of squares error fitting the data.

$L_\infty =$ __________ cm and $b =$ __________

$L(t) =$ __________ cm and SSE = __________

Find the L-intercept and the horizontal asymptote for the length of the Albacore.

L-intercept = ____________cm and Horizontal Asymptote $L =$ ____________cm

Give the model prediction at age 5 and 10 and find the percent error at each of these ages from the actual data given (assuming the actual data is the more accurate value):

Length at age 5 = ____________cm and Percent Error at 5 = ____________%

Length at age 10 = ____________cm and Percent Error at 10 = ____________%

b. In your Lab report, create a graph with the data and the von Bertalanffy model for $t \in [0, 15]$. Create a short paragraph that briefly describes how well the model simulates the data and what the maximum size of this fish can be.

c. For this part of the problem, we use the data to compute average growth rates for Albacore. The average growth rate is found by the following formula:

$$g_a(t_m) = \frac{L(t_2) - L(t_1)}{t_2 - t_1} \quad \text{where} \quad t_m = \frac{t_1 + t_2}{2},$$

where $L(t_1)$ and $L(t_2)$ are two successive data measurements of length and t_m is the midpoint between the ages that the fish was measured. We determine all the average growth rates from successive measurements from the Table above.

For $t_1 = 1$ and $t_2 = 2$, then $g_a(1.5) =$ ____________cm/yr.

For $t_1 = 2$ and $t_2 = 3$, then $g_a(2.5) =$ ____________cm/yr.

For $t_1 = 3$ and $t_2 = 4$, then $g_a(3.5) =$ ____________cm/yr.

For $t_1 = 4$ and $t_2 = 5$, then $g_a(4.5) =$ ____________cm/yr.

For $t_1 = 5$ and $t_2 = 6$, then $g_a(5.5) =$ ____________cm/yr.

For $t_1 = 6$ and $t_2 = 7$, then $g_a(6.5) =$ ____________cm/yr.

For $t_1 = 7$ and $t_2 = 8$, then $g_a(7.5) =$ ____________cm/yr.

For $t_1 = 8$ and $t_2 = 9$, then $g_a(8.5) =$ ____________cm/yr.

For $t_1 = 9$ and $t_2 = 10$, then $g_a(9.5) =$ ____________cm/yr.

We have noted that growth rates are derivatives. Use Maple to find the derivative of the von Bertalanffy model, $L'(t)$, found in Part a.

$L'(t) =$ ____cm/yr.

Use this formula to find the growth rate at the times listed below, then determine the percent error using the model growth rate from the actual growth rate determined from the data (the better value).

At age $t = 2.5$, $L'(2.5) =$ ____________cm/yr and Percent error = ____________%

At age $t = 4.5$, $L'(4.5) =$ ____________cm/yr and Percent error = ____________%

At age $t = 6.5$, $L'(6.5) =$ ____________cm/yr and Percent error = ____________%

At age $t = 8.5$, $L'(8.5) =$ ____________cm/yr and Percent error = ____________%

d. In your Lab Report, graph as data points the average growth rates that you computed above. To this graph add the model growth rate computed from the derivative of the von Bertalanffy equation for $t \in [0, 12]$. Briefly discuss how well the derivative of the model simulates the actual measured growth rates.

2 MAXIMA AND MINIMA WITH DERIVATIVES

One of the most important applications of the derivative is finding *maxima* and *minima* of a function. If a function is continuous and differentiable, then the derivative is zero at any *relative maximum* or *minimum.* Maple provides a valuable tool for finding maxima and minima.

Example 5.3. *Consider the cubic function:*

$$f(x) = 0.3x^3 + 1.2x^2 - 3.7x + 42.2.$$

Find the relative maximum and minimum for this function using Maple.

For this example, we enter the function and find its derivative, storing this derivative in the expression df. The following Maple commands satisfy this:

$f := x \to 0.3 \cdot x^3 + 1.2 \cdot x^2 - \cdot 3.7 * x + 42.2;$
$df := diff(f(x), x);$

The derivative produces

$$f'(x) = 0.9x^2 + 2.4x - 3.7.$$

Maple's *fsolve* command for a polynomial finds all roots of the polynomial and in this case stores it in the variable (vector), xm,

$xm := fsolve(df = 0, x);$
$f(xm[1]); \quad f(xm[2]);$

The result gives a relative maximum at $(-3.76004, 57.1299)$ and a relative minimum at $(1.09337, 39.9812)$.

Note: Maple only gives all solutions with *fsolve* if the function is a polynomial. Otherwise, it gives the first zero that it finds. To find a different zero of $f(x)$, say $x_m \in (a, b)$, then one issues the Maple command

$xm := fsolve(f(x) = 0, x = a..b);$

Problem 5.5 (Triatoma O_2)**.** *(Lab121_J2_triatoma.pg) This problem studies the oxygen consumption of the bug Triatoma phyllosoma after ingestion of a blood meal. A cubic polynomial is fit through the data, and maximum and minimum levels of oxygen consumption are found.*

This problem was given to me by Professor Boyd Collier from his research in Ecology. The oxygen consumption of the bug *Triatoma phyllosoma* or "kissing bug" after ingestion of a blood meal is measured over 15 hours. In this experiment the bug matures through its fifth instar stage of development and begins its molt into the adult phase. This is a bug that causes major problems throughout South America. It spreads the deadly disease chagas when it obtains a blood meal at night from its sleeping victims. The poor rural houses may be so infested with thousands of these bugs that some of the occupants become anemic by losing up to 500 ml of blood per month.

a. Below are the data for the time, t (in hours), and the oxygen consumption, y (in ml of O_2/hr) after a blood meal for *Triatoma phyllosoma*).

t (hr)	ml O_2/hr	t (hr)	ml O_2/hr	t (hr)	ml O_2/hr
1	109.3	6	105.2	11	77.1
2	111.2	7	92.6	12	54.9
3	112.1	8	94.2	13	52.7
4	122.6	9	91.7	14	68.4
5	115.3	10	81.3	15	75.2

Table 5.4: The times and oxygen consumption for *Triatoma phyllosoma*.

These data are most reasonably fit by a cubic equation. Use Excel's *Trendline* to find the best fit to the data using the equation:

$$y(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3.$$

The best fitting coefficients are

$a_0 =$ ____________

$a_1 =$ ____________

$a_2 =$ ____________

$a_3 =$ ____________

Use the model to estimate the oxygen consumption at $t = 3$, and assuming the data give the best values, determine the percent error.

$y(3) =$ __________ and Percent error = __________%

Use the model to estimate the oxygen consumption at $t = 7$, and assuming the data give the best values, determine the percent error.

$y(7) =$ __________ and Percent error = __________%

Use the model to estimate the oxygen consumption at $t = 13$, and assuming the data give the best values, determine the percent error.

$y(13) =$ __________ and Percent error = __________%

b. In your Lab Report, show the plot of the data points and the best cubic polynomial through the data for $t \in [0, 15]$. Write a brief description of how well this model fits the data.

c. Differentiate the polynomial that you found in Part a.

$y'(t) =$ ____

Find the value of the derivative at $t = 3$.

$y'(3) =$ ____________

Find the value of the derivative at $t = 7$.

$y'(7) =$ ____________

Find the value of the derivative at $t = 13$.

$y'(13) =$ ____________

The derivative is zero at two values $t = t_1$ and $t = t_2$ with $t_1 < t_2$. Find t_1 and determine $y(t_1)$.

$t_1 =$ ______________ and $y(t_1) = =$ ______________

Is this a MAXIMUM or MINIMUM? ______________

Find t_2 and determine $y(t_2)$.

$t_2 =$ ______________ and $y(t_2) = =$ ______________

Is this a MAXIMUM or MINIMUM? ______________

d. In your Lab Report, describe what is happening physiologically to this bug. State when the oxygen consumption is increasing and when it is decreasing. Recall that at $t = 0$, the bug has just finished eating a blood meal, and at $t = 15$, it molts to become an adult. With this information briefly present a biological explanation for why you see the regions of increasing oxygen consumption and decreasing oxygen consumption. (Hint: You might want to think about your own physiology after eating a meal. The molting stage could be related to you deciding to exercise at some time after your meal.)

Problem 5.6 (Circadian)**.** *(Lab121_J4_circadian.pg) Over the period of a day, the body temperature of a typical human varies about* $1^\circ C$*. The variation is part of the human circadian or internal clock. These variations of body temperature are correlated with the normal wake and sleep cycle and controlled by the brain. The clock is regularly reset by the pineal gland, which responds to external light.*

a. Below is a table of the body temperature of a young adult, showing the variation with time. The time $t = 0$ corresponds to midnight, and t is in hours.

t (hr)	°C	t (hr)	°C	t (hr)	°C
0	36.91	9	36.64	18	37.22
1	36.67	10	36.78	19	37.24
2	36.52	11	36.91	20	37.25
3	36.4	12	37.03	21	37.24
4	36.35	13	37.07	22	37.2
5	36.32	14	37.1	23	37.11
6	36.35	15	37.15	24	36.92
7	36.39	16	37.19		
8	36.48	17	37.22		

Table 5.5: The times and body temperatures over a day.

These data are most reasonably fit by a cubic equation. Use Excel's *Trendline* to find the best fit to the data using the equation:

$$T(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3.$$

The best fitting coefficients are

$a_0 =$ ______________

$a_1 =$ ______________

$a_2 =$ ______________

$a_3 =$ ______________

Use the model to estimate the body temperature at $t = 5$, and assuming the data give the best values, determine the percent error.

$T(5) =$ ____________ and Percent error = ____________%

Use the model to estimate the body temperature at $t = 13$, and assuming the data give the best values, determine the percent error.

$T(13) =$ ____________ and Percent error = ____________%

Use the model to estimate the body temperature at $t = 17$, and assuming the data give the best values, determine the percent error.

$T(17) =$ ____________ and Percent error = ____________%

b. In your Lab Report, show the plot of the data points and the best cubic polynomial through the data for $t \in [0, 24]$. Write a brief description of how well this model fits the data.

c. Differentiate the polynomial that you found in Part a.

$T'(t) =$ ____

Find the value of the derivative at $t = 5$.

$T'(5) =$ ____________

Find the value of the derivative at $t = 13$.

$T'(13) =$ ____________

Find the value of the derivative at $t = 17$.

$T'(17) =$ ____________

The derivative is zero at two values $t = t_1$ and $t = t_2$ with $t_1 < t_2$. Find t_1 and determine $T(t_1)$.

$t_1 =$ ____________ and $T(t_1) = =$ ____________

Is this a MAXIMUM or MINIMUM? ____________

Find t_2 and determine $T(t_2)$.

$t_2 =$ ____________ and $T(t_2) = =$ ____________

Is this a MAXIMUM or MINIMUM? ____________

d. In your Lab Report, assuming this is a normal young adult, write a brief discussion of how the body temperature varies with daily activities. Relate the body temperature to the normal sleep/wake cycle.

Problem 5.7 (Plankton)**.** *(Lab121_J3_plankton.pg) The Salton Sea has a variety of sea plankton that undergo seasonal blooms. This problem fits a quartic polynomial to the natural logarithm of the populations to estimate when the peak blooms occur and what levels they achieve based on a given set of data.*

Since the death of Sonny Bono, there has been renewed interest in the Salton Sea. The Salton Sea formed from 1905-1907 when an engineering mistake plus heavy rains on the watershed of the then undammed Colorado River combined to break through a levee. The lake was originally freshwater but became saltier than seawater as there is no outlet and a lot of evaporation. The water going into it is fairly salty (leached from the agricultural soils). The creatures in it are mostly marine, some introduced on purpose and some accidentally with the establishment of a sport fishery. If it weren't for the agricultural and municipal wastewater flowing into the Sea, it would have dried up long ago. This water also has a lot of fertilizers, which cause massive algal blooms, and a large biomass of invertebrates and fish. People like to fish there because it is so easy to catch fish, but there are also large fish kills, which are NOT pleasing. There are many birds at the Sea (there's lots of food for them), but they also experience large die-offs at times from avian epidemics (crowding).

There are not very many kinds of metazoan zooplankton (6) in the Salton Sea, but often there is a high density present. Dr. Debbie Dexter (retired SDSU) was studying marine invertebrates in the Salton Sea, and a presentation of some of her work can be found on the second floor of the Life Sciences building. Mary Ann Tiffany provided the background information above and the data below on some of the zooplankton. The table below lists the number/liter of various zooplankton for data averaged over depth for station S-1 (in the center of the north basin of the Salton Sea at a depth of 14 meters). The first column lists the number of days after January 1, 1997 when the data were taken. The second column represents the rotifer, *Brachionus rotundiformis*, the third column represents the nauplius form of the barnacle (*Balanus amphitrite*) larvae, and the last column is the nauplius form of the copepod, *Apocyclops dengizicus*.

Date	Rotifers	Barnacles	Copepods
21	0.045	4.466	0.06
34	0.047	2.04	0.063
53	0.073	0.7	0.102
78	0.167	0.573	0.251
106	51.785	0.295	0.093
154	182.403	0.035	45.687
175	372.655	0.031	50.25
199	295.288	0.035	14.56
225	802.128	0.039	59.539
249	532.203	0.031	21.629
277	33.723	0.031	2.992
311	9.245	0.056	4.551
329	0.93	0.149	1.65
371	0.047	0.491	0.144
402	0.081	1.618	0.159
423	0.178	1.925	0.097
454	5.826	0.408	0.08
479	299.183	1.923	0.08

Table 5.6: Days after 1/1/1997 and populations/cc of Salton Sea water.

a. In your Lab Report, create a separate graph for each of these species in Excel using a logarithmic scale for the population. Determine the season of the year when each of these species is most productive.

Peak season for rotifers is (Winter, Spring, Summer, Fall) ______________

Peak season for barnacles is (Winter, Spring, Summer, Fall) ______________

Peak season for copepods is (Winter, Spring, Summer, Fall) ______________

b. As seen in the graphs above, the populations of these zooplankton have very large changes in numbers from very low densities to relatively high densities. Without a logarithmic scale these fluctuations in population appear to jump rapidly between high and low densities. It follows that to fit a polynomial through the data we need to use the natural logarithms of the data.

Take the natural logarithm of the data for Barnacles, then plot these values against time. Use Excel's *Trendline* to find the best 4^{th} order polynomial through the logarithm of the data versus time. (Be sure to have at least 5 significant figures for your coefficients.) The polynomial has the form:

$$P(t) = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0.$$

List the values of the coefficients that *Trendline* gives you:

$a_4 =$ ____________

$a_3 =$ ____________

$a_2 =$ ____________

$a_1 =$ ____________

$a_0 =$ ____________

Use this polynomial to approximate the population at $t = 106$, then determine the absolute error. Find the polynomial value and the corresponding population (which requires conversion from the natural logarithm).

$P(106) =$ __________ Population at $t = 106$ is __________ Absolute Error is __________

Find the polynomial value, population, and absolute error at $t = 199$.

$P(199) =$ __________ Population at $t = 199$ is __________ Absolute Error is __________

Find the polynomial value, population, and absolute error at $t = 311$.

$P(311) =$ __________ Population at $t = 311$ is __________ Absolute Error is __________

c. In your Lab Report, create the graph of the natural logarithm of the data for Barnacles and show the *Trendline* passing through this graph. Describe how well the polynomial fits the data. Briefly discuss the seasonal variation in the population for Barnacles.

d. Differentiate the polynomial found in Part b. The derivative will have the form:

$$P'(t) = b_3t^3 + b_2t^2 + b_1t + b_0.$$

List the values of the coefficients for this derivative:

$b_3 =$ ____________

$b_2 =$ ____________

$b_1 =$ ____________

$b_0 =$ ____________

A quartic polynomial has three minima and maxima. Find when these minima and maxima occur and state the value of P(t) and the populations at these extrema. Assume the minima and maxima occur at times $t_1 < t_2 < t_3$.

$t_1 =$ __________ $P(t_1) =$ __________ and Population at t_1 is __________

What type of extremum is this? (Maximum or minimum) __________

$t_2 =$ __________ $P(t_2) =$ __________ and Population at t_2 is __________

What type of extremum is this? (Maximum or minimum) __________

$t_3 =$ __________ $P(t_3) =$ __________ and Population at t_3 is __________

What type of extremum is this? (Maximum or minimum) __________

The Salton Sea provides a valuable place for birds to congregate on their migrations. A maximum would be a good time for birds to feed at the Salton Sea. If January 1 corresponds to $t = 0$, then find the date that the maximum population of Barnacles occur (in the form MM/DD). Use only the maximum that occurs within the range of the data and assume there are 365 days in the year.

Date of the Maximum population = __________

e. In your Lab Report, write a brief paragraph discussing how well the derivative of the polynomial finds the actual maximum and minimum populations both in time and number. Does this provide a reasonable way to model this complex ecosystem? Explain.

Appendix A:

APPENDIX – LIST OF PROBLEMS

This appendix provides a complete listing of the problems in the Lab Manual.

BIBLIOGRAPHY

[1] C. A. Bessey and E. A. Bessey. Further notes on thermometer crickets. *American Naturalist*, 32:263–264, 1898.

[2] C. A. Beuchat, W. A. Calder III, and E. J. Braun. The integration of osmoregulation and energy balance in hummingbirds. *Physiological Zoology*, 63:10591081, 1990.

[3] T. Carlson. Über Geschwindigkeit und Größe der Hefevermehrung in Würze. *Biochem. Z.*, 57:313–334, 1913.

[4] AIDS cases in the U. S.: HIV/AIDS Surveillance Report. U. S. Department of Health and Human Services, Centers for Disease Control, Division of HIV/AIDS, Atlanta, GA, 1993.

[5] A. E. Dolbear. The cricket as a thermometer. *American Naturalist*, 31:970–971, 1897.

[6] M. G. Hinton. Status of Blue Marlin in the Pacific Ocean. http://isc.ac.affrc.go.jp/pdf. (Accessed 8/14/2015).

[7] David N. Holvey, editor. *The Merck Manual of Diagnosis and Therapy*. Merck Sharp and Dohme Research Laboratories, Rahway, NJ, 15th edition, 1987.

[8] H. Lineweaver and D. Burk. The determination of enzyme dissociation constants. *J. Am. Chem. Soc.*, 56 (3):658666, 1934.

[9] J. M. Mahaffy and A. Chávez-Ross. *Calculus: A Modeling Approach for the Life Sciences*, volume 2. Pearson Custom Publishing, Upper Saddle River, NJ, 2006.

[10] J. M. Mahaffy and A. Chávez-Ross. *Calculus: A Modeling Approach for the Life Sciences*, volume 1. Pearson Custom Publishing, Upper Saddle River, NJ, 2nd edition, 2009.

[11] R. M. May. Deterministic models with chaotic dynamics. *Nature*, 256:165–166, 1975.

[12] J. Mazumdar. *An Introduction to Mathematical Physiology and Biology*. Australian Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1989.

[13] L. Michaelis and M. L. Menten. Die Kinetik der Invertinwirkung. *Biochem Z*, 49:333369, 1913.

[14] T. A. Ryan, B. L. Joiner, and B. F. Ryan. *The Minitab Student Handbook*. Duxbury Press, 1976.

[15] J. Schmider, D. J. Greenblatt, J. S. Harmatz, and R. I. Shader. Enzyme kinetic modelling as a tool to analyse the behaviour of cytochrome P450 catalysed reactions: application to amitriptyline N-demethylation. *Brit. J. Clin. Pharmoco.*, 41:593–604, 1996.

[16] J. H. Uchiyama and T. K. Kazama. Updated Weight-on-length relationships for pelagic fishes caught in the central north Pacific Ocean and bottomfishes from the Northwestern Hawaiian Islands. http://www.pifsc.noaa.gov. (Accessed 8/14/2015).

[17] United States Census Bureau: Historical Census Data. http://www.census.gov. (Accessed 2/05/2014).

[18] United States Census Bureau: International Data Base. http://www.census.gov. (Accessed 12/19/2013).

Index