Computer Data Analysis and Plotting

In this lab we will use Microsoft EXCEL® to do our calculations. This program has been written primarily for use by the business community, but it is very user friendly and easy to learn, and has in addition most of the functions and tools necessary for serious scientific calculation. It can also produce beautiful, customized charts and graphs. The exercises below introduce you to a number of EXCEL functions which will be used in the labs to come. If you do not have time to finish them all today, you may still want to come back to them for reference later on. All of these exercises can be carried out on a home computer, or on any of the computers in the University computer labs available to students.

Each lab station has an IBM PC, connected over a local network to a laser printer and to a more powerful PC where all the programs and data files are stored. When the computers are turned on, you will need to log on with the user name student and the password student in order to have access to this network.

I. Introduction to Excel (Each person should do Part I separately.)

A. Running Microsoft Excel. This program is stored on the hard disk of each computer workstation. To run Excel, use the left mouse button to double click on the Excel icon located on the desktop screen. (Below, when we say "open," "choose," and "select", that generally means to click on something with the mouse.)

B. Some simple calculations. You should now have an empty spreadsheet, titled Sheet1, open in front of you. Select cell B2, type your name and press "Enter." Notice that the contents of the selected cell appears on a line near the top of the window, as well as in the spreadsheet.

Now we will make a graph of the trigonometric functions sin(θ) and cos(θ). Start by clicking on cell B5. Enter the number 0. Now select B6 and enter the number 0.1. While holding down the left mouse button, highlight the two cells B5 and B6. Once the two cells are highlighted, point the mouse arrow at the bottom right corner of the B6 cell (there should be a small black box in this corner, over which the cursor becomes a simple cross), push and hold the left mouse button and drag the mouse down the column B to cell B65, and let up on the mouse button. Column B should now contain a series of increasing values. Make sure that the value contained in the last cell, B65, is 6.0. Click on cell B4, and enter "theta." This column contains values of the angle theta, in radians.

Next click on cell C5. Here we want to calculate the value of sin(0), the sin of the theta value in cell B5. To do this, type:

=sin(

Now click on cell B5. In the bar near the top of the screen you should see:

=sin(B5

Finish off the formula by adding a right parenthesis, and pressing the "Enter" key. The number zero should appear in C5; this is the value of sin(0).

So far this doesn't seem very efficient; it was a lot of work to calculate just one value of $\sin(\theta)$. But now, we will calculate an entire column of values at once. Select cell C5. Then move the cursor to the lower right-hand corner of cell C5, until the cursor changes to a simple cross. Hold the mouse button down, and "drag" the formula all the way down the column, to cell C65, and release the button. An entire column of values of $\sin(\theta)$ values will be calculated, almost instantly.

Now click on cells C5, C6, and C7, one after the other, and watch the formula appearing in the data bar near the top of the screen. Do you see what happened? As the formula was copied down, the cell number for theta was incremented, so that each cell calculates a different value of $\sin(\theta)$. This clever trick is what lets you do a lot of calculations fast with a spreadsheet.

What about $\cos(\theta)$? Select cell D5, type in:

=cos(

select cell B5, and continue as you did with the sine function.

C. Graphing. Now you should have three columns of numbers: values of theta in B5-B65, values of $\sin(\theta)$ in C5-C65, and values of $\cos(\theta)$ in D5-D65. We can easily graph these values. To do this, select the entire block of data: put the cursor on cell B5, and, holding the button down, drag down to D65. The data should appear as a highlighted block, in contrasting colors. To graph it, look at the toolbar across the top of the screen. Click on the ChartWizard button, the third icon from the right. To indicate where to put the chart, put the cursor on any desired location and click once. A ChartWizard dialog box should open up, indicating that you are on step one of five. You should see a statement that says:

Range = $B$5:$D$65

These match the upper left cell of your highlighted region, B5, and the lower right cell of the region, D65. Click on the Next> button to continue on to step 2. Step 2 allows you to choose the type of chart you wish to display. For now choose XY (Scatter) by selecting the appropriate box and then clicking the Next> button to go on to step 3. Here you choose the format of the chart you are going to display: select box 2, and click on the Next> button. For step 4, leave all the selections unchanged and click on Next> to go straight on to step 5. In the step 5 dialog box you can choose whether or not you wish to display a data legend by selecting yes or no, you can add a title to your chart by filling in a name in the Chart title: box, and you can name each of the coordinate axes as well. When you are done, click on the Finish button to complete your chart. At any point during this process, you can repeat any step by using the <Back button to retreat to any earlier step in the sequence.

To see this graph better, single click on any portion of it, then move the mouse arrow over any of the six small boxes around its borders where the mouse pointer becomes a double-headed arrow. Push and hold the left mouse button and move the mouse to either enlarge or reduce the graph's size along that edge. Here are some things to try now, in any order:

1. Try some other types of graphs. Repeat the process again, but make your own choices for steps 2 and 3. See how these changes affect the displayed chart.
2. Select one of the curves by double-clicking on the chart, then double-click on the curve. You will get a menu where you can change the type of points, or eliminate points and make a line instead.
3. Try moving the chart to a different location. Push and hold the left mouse button while the arrow is on the chart, then move the chart around the page by moving the mouse. Let go of the mouse button and the chart drops where you leave it.
4. Preview a printout by selecting Print Preview from the File menu, or by clicking on the print preview icon on the tool bar (the magnifying glass over a sheet of paper, fifth from the left, top row). Here you can zoom in and out to preview the document as it will be printed. When done, click on the Close box at the top of the preview screen.
5. Print the document by selecting Print… from the File menu, then selecting OK from the Print screen. Or you can print by simply clicking once on the print icon on the toolbar (the printer icon, fourth from the left on the top row). Using the toolbar icon will not prompt you again, it will print the document as is!

That's the end of the introduction. Now you will plot some data for the purpose of analysis. Close this workbook by selecting Close from the File menu, or by clicking on the box containing an X just above the toolbar. You will be asked if you wish to save changes in 'Book 1'; select no.

Now if one of your partners has still not done Part I, please turn the computer over to him or her.

II. Graphing Data for the Purpose of Analysis Using Excel

Today most data analysis is done using computers. Here are a few examples. Please make graphs of the following data, print them out and place them in your lab book, according to the accompanying instructions. In each case: (1) give each axis a title, in words; (2) put a title at the top of the graph. It doesn't have to be big. A final suggestion - when you are graphing data, arrange to have the data either on the same page as the graph for printing, or on a facing page in your lab book from where you paste in your printout.

To open a new workbook, select New… from the File menu. Under General, select the Workbook, and click on OK. You should now have a blank spreadsheet opened on your computer.

A. Voltage across a resistor as a function of current. Below is a table of voltage measured across a resistor, for various values of the current. The objective of this exercise is to plot the data, with current as the *x* axis and voltage as the *y* axis, find the best-fit line, and draw the best-fit line through the points.

measurement number	current (amps)	voltage (volts)
1	0.5	1.22
2	1.00	3.11
3	1.5	4.37
4	2.0	5.11
5	2.5	7.02
6	1.25	3.62

Making the Graph. Begin by selecting cell B2, and filling in the names of all lab partners. It is important to do this on all graphs to properly identify your work when printing in the lab. Next, fill in cells A5 through A10 with the measurement numbers 1 through 6. Then select cell B4, type in "current" and press Tab. You should now be in cell C4, where you should type in "voltage" and press Enter. Notice the difference in what happens when you press Tab and when you press Enter. Now you are ready to enter the data into your spreadsheet. For each measurement number, fill in the corresponding values for measured current and voltage from the table above into your sheet.

We want to plot current along the x-axis and voltage on the y-axis, so select cell B5, push and hold the left mouse button, then use the mouse to highlight the current and voltage columns (not including their titles in cells B4 and C4), then release the mouse button. Next, click on the ChartWizard button on the toolbar and position the graph on your spreadsheet. For step one, verify that the Range = $B$5:$C$10 and click on the Next> button. Step two allows you to select the graph type, choose the X-Y Scatter option and then go to step 3. In this example, we want to include the gridlines on our graph, so select option (3) for the graph format selection then click Next>. Do not make any changes on step 4, simply move on to step five where you should fill in a Chart Title as well as titles for each axis, do that, then select Finish. Enlarge your graph to get a better view of your data.

Best-fit line - TRENDLINE method. Now we would like to display the straight line which best fits these data points. We will use the Excel function Trendline. Select the graph, and select the data points. Then, with the mouse on the data points, right-click and select "Add Trendline." Choose linear regression (a straight line), and, under "options," choose "Display

equation on chart." The best-fit line should appear on your chart, plotted over the data points.
The law of physics describing the voltage across a resistor is Ohm's law, which is written

$$V = I R,$$

where *V* represents the voltage across the resistor, *I* represents the current through the resistor, and *R* is the resistance. For the data points which you just plotted, what is the value of the resistance?

Best-fit line - LINEST method. The preceding method plots the straight line on your graph, and gives you the value of the slope and intercept. The one thing missing is the *errors* on the slope and intercept. What is the error on the value of the resistance which you just determined?
To find this out, we will use the Excel function LINEST. Highlight the four cells B13 through C14, then type =linest(c5:c10,b5:b10,1,1), but DO NOT hit Enter when done. When you have typed this, push and hold the Ctrl and Shift buttons, and while holding these, press Enter.
The four highlighted squares should contain the following information:

	B	C
13	the slope: 2.709061	the intercept: 0.124286
14	the error on the slope: 0.195653	the error on the intercept: 0.312561

So: what is the error on the value of the resistance?
There is a standard form for stating the result of a measurement. Here is how you should state the result for this determination of the resistance *R*

$$R = 2.71 \text{ ohms} \pm 0.20 \text{ ohms}.$$

[NOTE. If you want to require the intercept of the line to be zero, as is the case for Ohm's law, you can select this option when calculating the TRENDLINE. And in LINEST, the intercept will be set to zero if you enter =linest(c5:c10,b5:b10,0,1).]
Print out your data table and graph and paste them in your lab book. HINT: First use the Print Preview option to be sure your data and graph will print out on one piece of paper. You may need to move and resize your data plot to do this. DO NOT move your data tables. You may also want to save your EXCEL file to the hard disk. To do this, click on "file," "save as," and navigate to your proper directory. If you are in section *n,* this will be
c:\labs\ph122\secn\ .
Give the file a name, and then save it.
Now start a new workbook by selecting New… from the File menu, and then selecting Workbook and clicking on OK in the New dialog box.

B. Voltage in a R-C circuit versus time: the semi-log plot. The exponential function occurs in a number of physics phenomena. In this lab, we will encounter exponential decay in the R-C circuit, and in the decay of radioactive isotopes. The general form of an exponential decay is

$$x = A e^{-\lambda t}$$

Here *A* is the amplitude of the decaying function, and the constant λ determines how rapidly it decays. There are two times which are normally of interest:

$\tau = 1/\lambda$: this is the exponential decay time or the "time constant" for this process; in time τ, the function decreases by a factor of 1/e (about 1/3).

$T_{1/2} = 0.693 \times 1/\lambda$: this is the "half-life" of the decay process; in time $T_{1/2}$, the function decreases by a factor of 1/2. The half-life is generally used to describe nuclear decay processes.

An exponentially decaying function looks simpler when its logarithm is plotted as a function of time. Taking the log of both sides of the equation above gives

$$y = \ln(x) = \ln(A) - \lambda t$$

It is easy to see that in a plot of *y vs. t,* the slope will be equal to $-\lambda$, and from that you can calculate τ or $T_{1/2.}$

In this exercise we will make two graphs, a linear graph and a log-linear, or semi-log, graph. The procedure will be to calculate the log of the voltage values, make graphs of *V* versus time and ln V versus time, and determine the slope of the semi-log graph. We will follow the methods used in the previous section.

measurement number	time (sec)	voltage V (volts)	ln V
1	0	10.13	
2	2.0	6.93	
3	5.0	4.45	
4	7.5	3.03	
5	10.0	1.81	
6	15.	0.85	

Select cell A4 and type in time, then press Enter. In cells A5 through A10, enter the time values in order from the table above. Then select cell B4 and type in voltage and press Enter. In cells B5 through B10, enter the voltage values for each measurement as given above. Select cell C4 and type in lnV and press Enter. In cell C5 type =ln(B5) and press Enter. Highlight cell C5 again, move the mouse to the lower right corner of that cell where the pointer becomes a simple cross, push and hold the left mouse button, then highlight the cells C5 through C10. The highlighted cells should now contain the natural log of the voltages entered in the B column.

Now we are ready to create our graphs. First we will graph the voltage versus time. Use the mouse to highlight the cells A5 through B10 and then click on the ChartWizard button on the toolbar. Use the pointer to place the graph and press the left mouse button. Verify the range indicated, select an X-Y Scatter, use a linear format (selection (3)) and name the chart and axes. You should now have a graph of the voltage (y-axis) versus time (x-axis). Now we want to plot the natural log of the voltage versus time. Highlight the cells A5 through A10, then while holding the Ctrl button down highlight the cells C5 through C10. Again, click on the ChartWizard button to create a graph just as you have done before. Compare the two graphs you have created. Which one most resembles a straight line?

What we now want to do is to find the slope of the ln V *vs.* time graph. Use TRENDLINE to find the equation for the best-fit straight line. What is the slope?

Use the value of the slope to calculate the time constant τ and the half-life $T_{1/2}$ for this data set.

You again have the option of saving your spreadsheet to the hard drive by using the procedure described above.

III. Error Analysis.

We will often be using EXCEL to calculate the error on quantities which have been measured more than once. We may also want to compare our measured values with an expected (or theoretical) value. Suppose that a number, say the focal length of a lens, has been measured twice. The difference between the two measurements gives an idea of the random error in the measurement. We will define some terms, and show how to do the calculation, using EXCEL or by hand.

f The measured variable. Here it represents the focal length.

$\bar{f}$ The average value of *f*. (Pronounce it "eff bar.")

σ_f The standard deviation of the set. This represents the error on a single measurement of *f*. (Pronounce it "sigma eff.")

$\sigma_{\bar{f}}$ The standard deviation of the mean. $\sigma_{\bar{f}} = \frac{\sigma_f}{\sqrt{N}}$, where *N* is the number of measurements (N=2, in our case). This is what we will refer to as the error on the measurement of *f*. (Pronounce it "sigma eff bar.")

disc The "discrepancy," or difference between the measured value $\bar{f}$ and an expected ("theoretical") value: $\text{disc} = \left|\bar{f} - f_{theor}\right|$.

# of sigmas The discrepancy, expressed in terms of the standard deviation of the mean: $\#\text{ of sigmas} = \text{disc}/\sigma_{\bar{f}}$.

Q The "quality of agreement." This is a qualitative statement about how likely it is that a value measured with the precision indicated by $\sigma_{\bar{f}}$ should disagree with the true value by a certain number of sigmas, due to random errors. Look up *Q* in the table on page 4 of the appendix on errors in this manual.

Let us suppose that the two measured values of *f* are 11.5 cm and 9.2 cm, and that the lens is expected *a priori* to have a focal length of 10 cm. We will calculate *f*, $\bar{f}$, σ_f, $\sigma_{\bar{f}}$, disc, # of sigmas, and Q

.

Open a new spreadsheet. In B3 enter f1. In C3 enter f2. In D3 enter f-bar. In E3 enter sig-f. In F3 enter sig-f-bar. In G3 enter disc. In H3 enter No. of sigmas. And in I3 enter Q. In B4 and C4 enter the two values of *f*, 11.5 and 9.2. In D4 enter =AVERAGE(B4:C4). In E4 enter

=stdev(B4:C4) . In F4 enter =E4/sqrt(2) . Here is what this should look like:

f1	f2	fbar	sig-f	sig-f-bar	disc	No. of sigmas	Q
11.5	9.2	10.35	1.626346	1.15			

Now in G4 enter =D4 - 10 . (Here "10" is the expected value for the focal length *f.)* In H4 enter =G4/F4 . Then look up the quality of agreement in Table I of the appendix on Theory of Statistics and enter it in I4. Now your results should look like this:

f1	f2	fbar	sig-f	sig-f-bar	disc	No. of sigmas	Q
11.5	9.2	10.35	1.626346	1.15	0.35	0.304347826	Very Good

You now have the numbers to state the result of the measurement of *f*. The standard way to state this result would be as follows:

f = 10.35 cm ± 1.15 cm .

Agreement with expected value is very good (0.3 sigmas).

Be sure you understand this procedure, as it will be used repeatedly in the labs to come.

Sometimes you may not have the computer available to do these calculations, and you will then need to do them by hand. Calculating the average value $\bar{f}$ is easy, and dividing by the square root of two is also easy To calculate the standard deviation by hand, you can use the following formula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \xrightarrow{\text{two values, } f_1 \text{ and } f_2} \sqrt{(f_1 - \bar{f})^2 + (f_2 - \bar{f})^2}$$

You should be able to plug in values for f_1, f_2, and $\bar{f}$ and verify that Excel calculated the right value for σ_f .