College Physics 201
Graphing Review and Analysis

Materials: Graph paper, ruler, pencil and calculator.

1 Purpose

The goal of these laboratory exercises is to review graphing methods and learn how graphing is used to analyze data in the physics laboratory. Methods and skills learned in this laboratory will be employed throughout the semester. Keep this laboratory handout nearby for future reference.

2 Introduction

One of the most powerful tools in the data analysis toolbox is the use of a graph to reveal the relationship between variables within a system. It will be necessary for the student to understand both the concept and application of the graphing methods used in this laboratory in order to successfully complete future work and thereby succeed in the course. There are many different types of graphs and with today's powerful computers and sophisticated mathematical algorithms it is possible to fit a wide variety of data sets with many different functional forms. Yet it doesn't mean a hill of beans if the person plotting and fitting the data has no idea what the fit means and how to interpret the results. The interpretation or, better yet, the questioning of the results is far more important than the results themselves in many cases. Experiments don't have "right" answers all the time. In this type of course, we choose experiments with well known outcomes to aid in the teaching process but that doesn't mean your results will be exactly the same as your neighbors. This is where the final "questioning" by the data analysts comes into play and the goal is that by the end of the semester you will have more skills than when you started and a greater appreciation for the work necessary to reduce even a simple set of data down to a few quantities.

In this laboratory you will learn about three types of graphs and the methods for extracting the desired quantities from the resulting linear fits to the data. In each method the quest is to get the data into a linear form so that a straight-line (a.k.a. linear) fit can be performed and the slope and intercept determined. This laboratory will be done with paper and pencil but in the future you will use a computer program to plot the data and determine the fit parameters. The next three sections will outline the methods we are going to use to analyze several sets of simulated data. The goal of the exercises will be to introduce you to these techniques so don't be surprised that you are not an expert at the end of the day. As they say Rome wasn't built in a day but with hard work and dedication you can improve significantly by the end of the semester.

3 Rules of Graphing

***Please note that this is NOT an optional section to read**

Graphing is something you have done many times before and can be taken for granted. Several points should be made at this junction in time and rules need to be laid down before you begin with the exercises. Throughout the semester you should refer to these rules whenever you are constructing graphs. These rules plus any other details given by your instructor MUST be followed or points will most likely be deducted.

1. Each graph must have a title.

2. "widgets vs. gizmos" means widgets on the vertical axis and gizmos on the horizontal axis.
3. The axes must each be labeled with the name of the variable (eg. speed) and the units of the variable (eg. m/s). Please note that Log graphs have unit labels as well. They differ slightly.
4. The scale of each axis must be chosen so that it's easy to plot the points and to read of the coordinates of points. Use only multiples of 2 or 5 to label the bold grid lines on the graph paper. (eg. 8,10,12,14, ... but not 8,10.5,13,15.5, ...)
5. The axes don't have to start at zero. Note that EXCEL sometimes chooses poorly.
6. The range of values for each axis must be selected so that data points used **at least half of the extent of the each axis**. This means that at times your best choice for the origin may well not be (0,0). Choose appropriate axes scaling that will result in an efficient spreading of the data.
7. Keep the data points as small as possible.
8. Never use "connect-the-dots" in hand drawn or computer generated plots. Use only linear fits when generating best fit lines unless otherwise requested.
9. When writing and equation on the graph replace the generic "x" and "y" symbols with more theory specific symbols.

4 Generating the Best Fit Line

The generation of the best fit line is very important. In this laboratory you will generate the best fit line with a pencil and a ruler. The best fit line is the line which "best represents the data points". Do this by lining up the ruler with the data points keeping approximately an equal number above and below the ruler (transparent rulers work the best). Also, avoid "torquing" the line by having points at one end of the line above the line while the other end of the line has points below the line. This generates a systematic error in the slope. The key is that the best fit line should be an interpolation of the ALL the data points. It's a judgment call that is done with experience and care.

The following procedure is used to calculate the slope and y-intercept for the line your have just drawn.

- Select two **new** "slope points" on the best fit line. Indicate them by putting a square around each and number them 1 and 2. They must NOT be data points.
- The slope points must be near the ends of the best fit line — ie. far apart. Don't select them by looking for a point where the best fit line intersects with grid lines.
- On the graph, label the slope points with their coordinates and units.
- The slope formula is:

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

- Show the slope calculations (on the graph or on the back of the graph) by first writing the above equation, then substituting the slope-point coordinates with units, and only then giving the answer with units.
- The slope will be wrong it you leave out the units. This is a common mistake!

- Calculate the y-intercept by substituting into the point-slope formula the slope you just calculated and one of the points you used in the slope calculation.

$$b = y - mx$$

- Make sure you have units on the final answer for the y-intercept.

5 Interpretation of Linear Fits

A common mistake made by students starting out in this course is that they believe the slope and/or y-intercept are the only numbers desired by the experimenter. This leads to the student wanting to interpret the slope or the y-intercept to be the quantity we are attempting to measure. This is far from the truth. In your previous experiences this may have been true only due to the fact that the chosen experiments were very simple in their relationships between different variables in the system being studied. More complex relationships are found as we make our experiments more sophisticated and dig deeper into the natural world attempting to reveal is fundamental underpinnings.

In this course you will work with experiments which have both simple and more complex relationships so it will be necessary for you to learn how to work with very simple and more complex systems. Whenever we are doing data plots in introductory physics, our goal in the long run is to reduce the data down to a plot which can be fitted with a straight-line. Other fitting algorithms exist but their interpretations are more complex and are better suited for a more advanced course. Therefore, whenever we run into data that by itself is non-linear in nature then we will employ data analysis tools to "linearize" the data. Namely we will use powers and logarithms to manipulate the data prior to plotting it and performing the linear fit. Of courses as with all algebraic manipulations there is a price to be paid. The price in this case is that by performing the algebraic manipulations on the data makes the data plot linear but then the extraction of the constants from the slope and y-intercept becomes more challenging. The details of these methods will be discussed in the next few sections.

Before moving on to the specific techniques, a few points about linear fitting should be reviewed. Linear fits are a best fit to a set of data points relating two variables. The slope and intercept of the best fit line then **represent** all of the data points simultaneously and the data points are no longer necessary. All further calculations and conclusions will be done using only the slope and y-intercept. It is in the drawing of the conclusion that the most important work is done. It should also be noted that many times the theory we are attempting to test is not directly related to the slope or y-intercept. A bit of mathematical manipulation may need to be done in order to **extract** the desired value from the resulting slope and y-intercept of the fit. The examples that are given in the discussions below will illustrate this.

5.1 Graphing Linear Data

Linear data is the "easiest" to handle in terms of graphing and fitting the data to a straight-line. There can still be hang-ups in the interpretation of what the slope and y-intercept mean but at least the data can be plotted in a straight forward manner.

Data that is believed to be linear in relationship is typically plotted with the dependent variable on the vertical axis and the independent variable on the horizontal axis. From time-to-time, we will not follow this plotting convention and reverse the axis. Once the plot is generated then a linear fit is performed on the data and the equation of the fit is displayed on the plot. This equation takes on the general form of

$$y = mx + b \tag{1}$$

which is the general form of a straight-line. Here the m is the slope of the line and b is the y-intercept. The units on the slope and y-intercept are such as to satisfy the units of the independent and dependent variables and the linear equation. The "x" and "y" are replaced with the appropriate symbols for the theory being investigated.

5.2 Using Powers to Linearize non-linear Data

Often the relationship between variables is not linear. In order to handle this data and extract the interesting constants, the data needs to be manipulated algebraically before it is plotted. Using powers on either the independent or dependent data will often result in a plot that is now linear in form and can be fitted with a straight-line. However, this means the experimenter must have a good handle on the theory to give the experimenter an idea of which power should be used on the data.

Relationships between two variables in the natural world which follow a power law take on the mathematical relationship

$$y = kx^n \tag{2}$$

where k and n are constants. In this relationship the value of n is the power of the relationship. The k value is a proportionality constant which basically controls the strength of the function. If you were to graph this function the k value would control the overall height of the function as x varies.

When dealing with powers the 'power' must be a dimensionless and unit-less value. In this law, n is just a pure number. The k value, on the other hand, must have dimensions which when combined with the power dimensions of the independent variable x will result in the right-hand side having the same dimensions as the left-hand side.

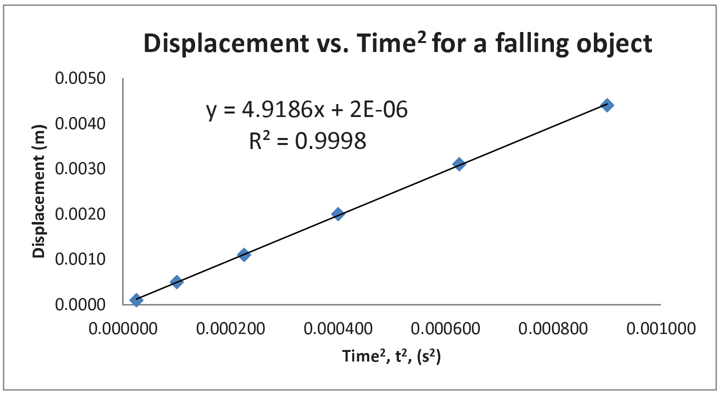
How do we handle data we believe fits a power law? Well, there are two ways to handle the data. First, the experimenter could use logarithm analysis (see next section) to determine what the power should be. The other way to do this is to use a theory which **predicts** what power you should use. A word of caution at this point. You have to be careful not to depend too heavily on the source for obtaining the choice of power for the power-law form. Both the logarithm and the theory methods have weaknesses and only experience will keep you from going astray. In this class, the experiments have been chosen such that the power law forms are well understood so you won't be "hunting in the dark".

Once the 'power' (n value) has been chosen then the experimenter can manipulate the data using algebra to linearize the data. I shall use an example. Suppose we collect data on an object falling as a function of time. The object starts from rest and is falling under the influence of gravity alone. The theoretical algebraic expression which predicts the displacement vs. time relationship is

$$d = \frac{1}{2}gt^2. \tag{3}$$

here the value of g is the acceleration of earth's gravity. If in the experiment the time is controlled and the displacement is recorded then a linearization of the data can be done. The linearization is done by squaring the data for the variable which is squared in the theory. In this case that is the time variable t.

Displacement (m)	Time (s)	Time Squared (s^2)
0.0001	0.0050	25.0×10^{-6}
0.0005	0.0100	$100.\times10^{-6}$
0.0011	0.0150	$225.\times10^{-6}$
0.0020	0.0200	$400.\times10^{-6}$
0.0031	0.0250	$625.\times10^{-6}$
0.0044	0.0300	$900.\times10^{-6}$



Now it can be seen from a plot of displacement versus time squared that the modified data is linear. The experimenter can now fit the data with a best fit line and determine the slope and intercept. Once this has been done the slope and/or intercept can be used to determine an experimental value for a particular constant. This can be done by comparing the theory equation side-by-side with the general expression for a straight line from the fit.

$$d = \frac{1}{2}gt^2 + 0 \tag{4}$$

$$y = 4.92\ x + 2.0 \times 10^{-6} \tag{5}$$

Note here that I did not convert the "y" and "x" symbols to "d" and "t^2", respectively, and I added a "0" to the end of the first equation in order to illustrate a point. You have to become familiar with the direct comparison you are making in the analysis. You have plotted d versus t^2 so you must make the connection that "y" represents "d" and "x" represents "t^2". This means that the theoretical slope equals $\frac{1}{2}g$ and the theoretical y-intercept is zero. We make the direct connection to the experimental results from the best fit of the plot. The experimental y-intercept is very small (2 parts in 1 million) as predicted and the slope is equal to $\frac{1}{2}g$. An experimental value for g can be determined by solving the expression

$$m = \frac{1}{2}g \tag{6}$$

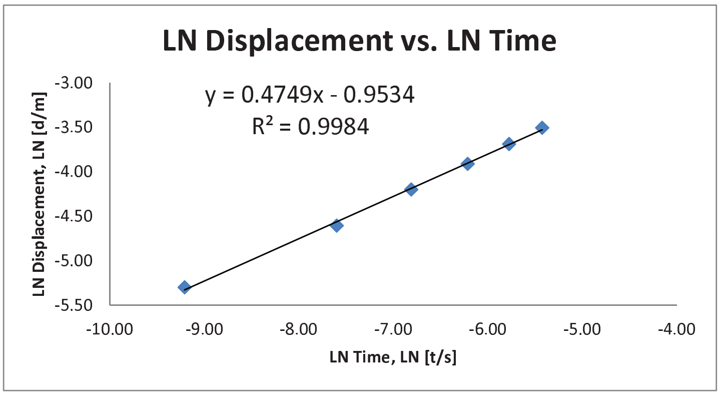
and using the measured value of the slope of the best fit line for m.

5.3 Using Logarithms to Linearize non-linear Data

The second way to analyze power law relationship is the use of logarithms. This method is used heavily in determining what the best 'power' is for the power law relationship when a trustworthy theory is not available or you are testing this aspect of the theory. In this section, the same example as above will be used in the logarithm analysis to follow.

The method requires the taking of the natural log of both the independent and the dependent columns of data. This is done in the following data table. Then the data is plotted as *Natural Log of Displacement versus Natural Log of time*. Note that the data points now show a linear relationship after the manipulation using the logarithms (see next graph). Look closely at how the axis and title of the graph are labeled. You should use similar labels on future graphs.

Displacement (m)	Time (s)	Natural Log of displacement [d/m]	Natural Log of time [t/s]
0.0001	0.0050	-9.01	-5.30
0.0005	0.0100	-7.62	-4.61
0.0011	0.0150	-6.81	-4.20
0.0020	0.0200	-6.23	-3.91
0.0031	0.0250	-5.79	-3.69
0.0044	0.0300	-5.42	-3.51



Now that the data is plotted in a linear form the slope and intercept of the best fit line can be determined. Once this is done values can be extracted from the slope and intercept for the different constants of the relationship. This is not a simple direct comparison. Why do you ask? The problem lies with the fact that the natural log function was used on the data. Linearizing the data costs us something. That cost is the fact that the constants of interest are buried in the natural logs. The following will illustrate this fact.

Modifying the data requires us to modify the theory. Let us first look at the general power law relationship before tackling our specific example. Taking the natural log of both sides of Equation (2) yields the following.

$$Ln(y) = Ln(kx^n) \tag{7}$$

Using the log rules at the end of this handout, the right-hand side can be broken into

$$Ln(y) = Ln(k) + Ln(x^n) \tag{8}$$

and then another rule brings the power out front.

$$Ln(y) = nLn(x) + Ln(k) \tag{9}$$

We can now see that the constant k is in the y-intercept term and the constant n is the multiplicative factor in front of the independent variable. The correct interpretation here is that the y-intercept of the Log-Log graph is equal to the $Ln(k)$ and the slope of the graph is the value n. Thus the theory is directly related to the graph once this algebra has been done. The experimenter can now "extract" the values of n and k from the slope and intercept by setting the theory equal to the experimental values.

$$b = Ln(k) \tag{10}$$

$$m = n \tag{11}$$

Now focusing on the data and theory for our falling object, we can perform the same algebraic manipulations with Equation (3) that we performed on Equation (2) above.

$$Ln(d) = Ln(\frac{1}{2}gt^2) \tag{12}$$

$$Ln(d) = Ln(\frac{1}{2}g) + Ln(t^2) \tag{13}$$

$$Ln(d) = 2 * Ln(t) + Ln(\frac{1}{2}g) \tag{14}$$

$$y = m * x + b \tag{15}$$

Now we do a 1-to-1 comparison between the general expression for a straight line fit (Eq. 13) and the final result of the manipulation of our theory (Eq. 12). The graph had $Ln(t)$ plotted on the x-axis and $Ln(d)$ plotted on the y-axis so in the above comparison the 2 in front of $Ln(t)$ be our theoretical slope. The trailing $Ln(\frac{1}{2}g)$ then must be the y-intercept. Since our experimental fit spits out values for slope and y-intercept we can do direct comparisons with theory and even extract values such as g. Thus the data analysis can be summarized as the following.

$$m = 2 \tag{16}$$

$$b = Ln(\frac{1}{2}g) \tag{17}$$

6 Graphing Exercises

In these exercises the student will take three data sets and perform the three primary analysis discussed above. This will all be done by hand (calculator and hand drawn graphs). In the future you will use a computer to produce your graphs. **Note: The follow data sets are DIFFERENT than the previous so although the methodology is the same the specifics are different. You need to work through ALL the details.**

6.1 Linear Data

In the following data table is a set of data which follows a linear relationship. The relationship is between the potential energy of an object in the gravitational field and the height it is above the ground. The relevant model for this physical system relates the object's gravitational potential energy (U), the object's mass (m), gravitational acceleration of an object due to the gravitational force acting on an object near the Earth's surface (g) and the object's height (h) measured from the surface; that is

$$U = mgh \tag{18}$$

Exercise: Generate a hand drawn graph plotting Gravitational Potential Energy versus height (see table for data). Fit the resulting plot with the "best fit" line and calculate the slope and intercept. Follow the rules and guidelines of the previous sections when making your graphs and calculating the slope and intercept. Is there a data point which must be treated differently? If so, circle the data point on your plot and make a note as to what should be done.

Height (m)	Gravitational Potential Energy (J)
0.20	0.49
0.40	1.04
0.60	1.33
0.80	1.96
1.00	2.49
1.20	3.21
1.40	3.43
1.60	3.93
1.80	4.54
2.00	1.73

6.2 Linearization by Powers

The data in the following table follows a power law form. The theoretical equation which fits this data is

$$T = 2\pi\sqrt{\frac{L}{g}} \tag{19}$$

where T is the period of a pendulum's swing, L is the length of a pendulum string and g is the acceleration of earth's gravity at the surface of the earth.

Exercise: We will assume the "power" of the power law form is $\frac{1}{2}$ thus you need the square root of the length of the pendulum. Fill in the rest of the table and generate the plot of *Period* vs. $\sqrt{length}$. Then draw the best fit line and calculate the slope and intercept as done before. Don't forget units and correct significant figures.

Length (m)	Oscillation Period (s)	Square Root of Length ($m^{1/2}$)
0.15	0.816	
0.30	1.13	
0.40	1.32	
0.50	1.44	
0.60	1.59	
0.77	1.77	

6.3 Linearization by Logarithms

To illustrate the power of the logarithm, the student will analyze the data from the previous section using the natural logarithm.

Exercise: Fill in the remaining two columns of the next data table. Remember to watch your significant figures. By the way, some of the numbers will be negative since you are taking the natural log of a value less than 1. Then graph the data as "Natural log of oscillation period versus natural log of length". Generate the slope and intercept of the best fit line as before.

Length (m)	Oscillation Period (s)	Natural Log of Oscillation period [T/sec]	Natural Log of length [l/m]
0.15	0.816		
0.30	1.13		
0.40	1.32		
0.50	1.44		
0.60	1.59		
0.77	1.77		

7 Analysis of the Slope and intercept

7.1 Linear Graph

- What is the slope (with units) of the line from the linear data graph? Show *all* of your work.
- Solve for the mass of the object. You may find Eq. 18 to be helpful. Assume $g = 9.80$ [m/s]. [Show all of your work by solving for the unknown quantity first algebraically and then your numeric substitutions should be the very last step.]

7.2 Linearization by Powers

- What is the slope of the line from this graph?
- What does the slope represent? [Hint: Compare the line fit with the theory Eq. 19.] This requires algebraic manipulation similar to the examples.
- From the comparison of the theory slope to the experimental slope, determine the value for g. This process is called "extraction".

7.3 Linearization by Logarithms

- What is the slope of the line from the graph? What is the y-intercept of the graph?
- Using the log rules in the next section, explicitly show that Eq. 19 can be reduced to

$$Ln(T) = \frac{1}{2}Ln(L) + Ln\left(\frac{2\pi}{\sqrt{g}}\right). \tag{20}$$

- Using Eq. 20 and the results for the y-intercept of your graph, determine a value for g. Remember to include units.
- Finally, compare this value for g to the value for g from the previous section (Linearization by Powers). Do this by taking the difference between the two values and dividing by the average of the two values. Then multiply that result by 100 to get the percent difference.

8 Useful Logarithm Expressions

$$Ln(AB) = Ln(A) + Ln(B) \tag{21}$$

$$Ln\left(\frac{A}{B}\right) = Ln(A) - Ln(B) \tag{22}$$

$$Ln(A^n) = nLn(A) \tag{23}$$

$$Ln(e^x) = x \tag{24}$$

$$e^{Ln(x)} = x \tag{25}$$

9 References

- OpenStax, *Physics*, Chapter 2, section 8
- Cutnell and Johnson, *Physics*, 5th Edition, Chapter 2, section 7
- A good algebra or calculus book

Data Tables

Height (m)	Gravitational Potential Energy (J)
0.20	0.49
0.40	1.04
0.60	1.33
0.80	1.96
1.00	2.49
1.20	3.21
1.40	3.43
1.60	3.93
1.80	4.54
2.00	1.73

Length (m)	Oscillation Period (s)	Square Root of Length ($m^{1/2}$)
0.15	0.816	
0.30	1.13	
0.40	1.32	
0.50	1.44	
0.60	1.59	
0.77	1.77	

Length (m)	Oscillation Period (s)	Natural Log of Oscillation period [T/sec]	Natural Log of length [l/m]
0.15	0.816		
0.30	1.13		
0.40	1.32		
0.50	1.44		
0.60	1.59		
0.77	1.77		