

225L Graphical Analysis

Introduction: Most of you have likely had to plot graphs sometime in your previous education. Graphs are an extremely useful way to display data, trends in the data, and help us to understand the nature of a process or proposed law. There are many different types of graphs, a sampling of which can be found under the insert menu item, charts, in the Excel spreadsheet.

In this first lab period we will focus on two dimensional graphs, and the two most commonly used in science and engineering, Cartesian plots and log-log plots. The most common type of graph called a Cartesian plot is when both of the axis scales are linear (equally spaced divisions). This type of plot is useful for displaying data that have a linear relationship, or one that is nonlinear, where the data do not range over multiple orders of magnitude. Arguably the second most common graph used to display data and find trends is the log-log plot. This is where both axis scales are ruled proportional to the log of the number¹. This type of graph is extremely useful in situations where one or both variables ranges over multiple orders of magnitude, and for extracting power law relationships and their exponents between variables.

Summary: In today's lab, you will be making two graphs by hand, one on Cartesian graph paper and one on log-log graph paper which will be supplied. In addition, you will make two graphs on the computer using the Excel spreadsheet. In the Cartesian graph section, you will plot the data given and determine the slope and intercept, assuming a linear relationship between the variables exists. You will also determine the slope and intercept using a best fit trendline and compare to the hand drawn graph results for the slope and intercept. A spreadsheet will be used to investigate the effects of random noise on the slope and intercept values. In the log-log graph section, you will plot the data for the period of the planets vs. their semi-major axis distances (effectively their mean distance from the sun) to find the mathematical relationship between them. You will find the exponent of an assumed power law relationship between the variables for both the hand drawn graph and the computer graph with a least square fit trendline, and compare the two values. The graphs with your comments, answers to questions, and data table should be stapled together and *submitted as a single lab report for each lab station*. The details of this exercise follow.

Cartesian Plots

You are given the following data of velocity v as a function of time t for an air track cart.

	no noise	noisy
t(s)	v (m/s)	v_n (m/s)
0.35	2.948	2.786
0.85	2.873	3.143
1.35	2.798	2.540
1.85	2.723	2.719
2.35	2.648	2.857
2.85	2.573	2.805
3.35	2.498	2.257
3.85	2.423	2.596
4.35	2.348	2.243
4.85	2.273	2.377

Table 1

¹ The semi-log graph, where one axis is scaled linearly, and the other proportional to the log of the number, will not be covered in this lab.

The second column in Table 1 are the exact or ideal values of v . The third column of the table labeled v_n , are one set of values with uniformly distributed random noise or error of up to $\pm 10\%$ added to the data of column 2. This is one of many possible sets of random noise values of v_n .

1.) Plot the column 3 data as a function of time, v_n vs. t from Table 1 by hand on Cartesian graph paper (v_n on the vertical scale, t on the horizontal scale). Draw data points with large markers so that they are easily seen (X is a good marker for hand graphs having relatively few data points. Computer graphs have many different markers to select from). Choose the vertical and horizontal scales on the hand drawn graph so that the data are spread out over most of the horizontal and vertical range of the plot area and the numbered values on the axis are convenient. Label the variables near their axis with their units in parenthesis, as shown on the example (computer drawn) graphs of the data in fig. 1 below. Assume constant acceleration so that $v = v_0 + at$, and draw the best “eyeball fit”² straight line through your data. Extend the line to the y axis so that you can determine the y intercept from your plot. Determine a and v_0 from the slope³ and y intercept of your graph and write their values (with units) on the graph and in Table 2 which follows.

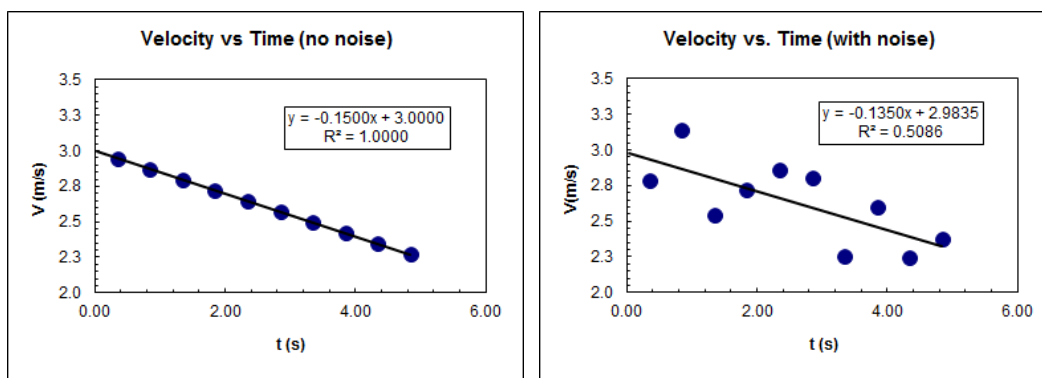


Figure 1. Velocity vs. time for the data in Table 1 with best fit trendlines and their equations shown. Note the reduced value of R^2 for the noisy data, corresponding to much larger deviations from the best fit straight line.

Example graphs in figure 1 show the labeling and choice of axis scales to display data over most of the plot area. Note that the best fit, “least squares” straight line equation is given in the inset, as well as R^2 , which is the coefficient of determination, and a measure of the goodness of the fit. The closer R^2 is to one, the better the fit of the model curve to the actual data.

2.) Enter the (t , v_n) data points in 2 adjacent columns in the sheet 1 tab of the Excel spreadsheet found on the lab website. Make a similar plot (scatter plot with no lines connecting the data points) for the data in the Excel spreadsheet and pass a linear trend line through the data (this can be done by right clicking on the data point series once you make the scatter plot). Display the graph equation and the R^2 value on the graph, which are options in the trendline box that pops up when you specify trendline. Format the trend line label so that it has 4 decimals (this is under format

² An “eyeball fit” straight line to the data places the line so that about the same number of data points are above the line as are below the line, and the areas above and below the line to the data are approximately equal. Obviously this is a somewhat subjective procedure.

³ The slope m in the linear equation $y = mx + b$ is determined by choosing 2 points (x_1, y_1) and (x_2, y_2) on the line that are conveniently separated and using the equation $m = \Delta y / \Delta x = (y_2 - y_1) / (x_2 - x_1)$.

trendline label when you right click on the trendline and number format). The default does not usually have enough significant figures to get an accurate value of the slope and intercept. Be sure to title the graph, and label the axis with title and units in parenthesis, as in the examples in fig. 1 above.

Compare the difference between your “eyeball fit” to the data, slope and intercept, and your trendline fit values from Excel for the same set of data. Enter your eyeball fit slope and intercept values into the first row of Student Table 2 tab at the bottom of the spreadsheet that follows.

Effect of errors and noise on the determination of fit parameters

3.) Use the Cartesian linear tab worksheet to see how different amounts of error or noise in the data affect the slope and intercept values by entering in the same and different values for the noise amplitude. Use the first column values in Student Table 2 below. Note how re-entering the same number for the noise amplitude gives a different set of random values for v_n and thus different slope, intercept, and R^2 values. Copy and use the paste special (select number formats and values option) to paste the 3 values for a , v_0 and R^2 at the top of the sheet labeled Cartesian linear into the corresponding rows in the Student Table 2 sheet tab (it does not work if you just use paste, since it copies the cell formulas from the Cartesian linear sheet and not the numerical values). The 0.1 noise amplitude values have been filled out for you already as a pattern. The last 2 columns of percent difference are calculated automatically in the spreadsheet.

Student Table 2						
Effect of different noise amplitude on determination of slope and intercept						
				no noise value	no noise value	
Enter your numbers in highlighted cells				$a(m/s^2)$	$v_0(m/s)$	
				-0.15	3	
	Noise Amplitude	Slope $a(m/s^2)$	Intercept $b = v_0 (m/s)$	Coeffic of Det. R^2	%dev from no noise Δa	%dev from no noise Δv_0
eyeball fit	0.1			NA	-100.0000	-100.0000
	0.1	-0.1402	2.8872	0.7492	-6.5038	-3.7611
	0.1	-0.1900	3.0856	0.8543	26.6750	2.8536
	0.1	-0.0929	2.7627	0.4953	-38.0953	-7.9111
	0.1	-0.1665	3.0055	0.6696	10.9838	0.1844
	0.1	-0.1334	3.0378	0.6093	-11.0744	1.2604
	0.05				-100.0000	-100.0000
	0.05				-100.0000	-100.0000
	0.05				-100.0000	-100.0000
	0.05				-100.0000	-100.0000
	0.05				-100.0000	-100.0000
	0.08				-100.0000	-100.0000
	0.08				-100.0000	-100.0000
	0.08				-100.0000	-100.0000
	0.08				-100.0000	-100.0000
	0.08				-100.0000	-100.0000
	0.15				-100.0000	-100.0000
	0.15				-100.0000	-100.0000
	0.15				-100.0000	-100.0000
	0.15				-100.0000	-100.0000
	0.15				-100.0000	-100.0000

Table 2

Notice for the 0.1 noise amplitude the slope and intercept vary considerably between different random configurations with the same noise amplitude. The slope varies from the no noise value between about -36% to +27%, and the intercept varies considerably less, from about -8% to about +2.9%. This shows that the determination of the slope can be much less accurate than the noise percentage, which was no more than +/-10% in the table, and the accuracy of the intercept determination is comparable to the noise amplitude. There are mathematical equations for how the noise in the data affect the determination of the parameters, but these are beyond the present introduction to the effects of noise and error on the determination of parameters.

After filling out the table for the other cases in Student table 2, compare the variations in slope and intercept for the other cases of noise amplitudes of 0.05 and 0.15 (*you can skip the 0.08 case if you are short of time*). What can you conclude about the possible accuracy of the determination of the slope and intercept vs. the variation of the noise for these other cases?

Log-Log Plots

4.) Using the 4 cycle by 3 cycle log-log graph paper that is available in the lab, plot the solar system data Table 3 below with the period on the vertical (4 cycle) axis and semimajor axis data (approximate mean distance to the sun) on the horizontal (3 cycle) axis. (each cycle is a factor of 10 change from the previous value (x_1, y_1) and (x_2, y_2)) Label the vertical axis scale for each power of 10 starting at 0.1 at the bottom, and successively increasing by factors of 10 to 1000 at the top (i.e. 0.1, 1, 10, 100, 1000). For the horizontal, semi-major axis, start with 1 on the left and go to 1000 on the extreme right (i.e. 1, 10, 100, 1000). Find the slope of the hand drawn log-log graph and thus the exponent k in the power law fit for the data, $y = Ax^k$. Also find the value of the prefactor A . Both the slope formula for the exponent k and the fact that a power law turns into a straight line when plotted on log-log graph paper, are shown in the Appendix. (see Appendix eqn. 6 for the power law exponent k , and eqn. 7 for the prefactor A).

Planet	Semimajor	Period	T^2/a^3
	axis	T (y)	$(10^{-34} y^2/m^3)$
	$(10^{10} m)$		
Mercury	5.79	0.241	2.99
Venus	10.8	0.615	3
Earth	15	1	2.96
Mars	22.8	1.88	2.98
Jupiter	77.8	11.9	3.01
Saturn	143	29.5	2.98
Uranus	287	84	2.98
Neptune	450	165	2.99
Phuto	590	248	2.99

Table 3

Solar System Planetary Period vs. Distance from sun (semi-major axis)

5.) Repeat the exercise in part 2 for the solar system data on the sheet 2 tab of the spreadsheet, but change the graph scales to log-log (this is done by right clicking on the numbers below each axis and selecting format axis. Click log scale as one of the options on the window that pops up). See

the example in the Log-Log Planets spreadsheet tab. Find the exponent k of a power law trendline fit to the data. Report this value and compare it to the value you found with the hand drawn log-log graph. Briefly comment on the goodness of the fit of the data to the power law fit that you found.

Report:

When you are done with items 1.→ 5., copy your computer drawn graphs into a word document, along with any comments you may have and print them (suggest using paste special as a bitmap). Staple your computer graphs, hand drawn graphs, and comments together, along with a cover sheet. Both lab partner names, experiment title, and date should be on the cover sheet. Your instructor may want you to turn in your lab reports online in the future. However, for this first report that contains hand drawn graphs, you will turn in a paper report.

Appendix: (you only need eqns. 6 and 7 below to do the log-log graph part of the lab)

How to find the power law exponent from the slope of a linear plot on a log-log graph

As in the case for a Cartesian graph, pick 2 widely separated and convenient points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ that lie on the “eyeball fit” straight line on a log-log plot. Call them “slope points”. You should try to pick points that are near major divisions of smaller values if possible for the best accuracy in reading the values.

The power law model, with power law exponent k and prefactor A is given by

$$y = Ax^k \quad (1.)$$

Taking the log to any base of both sides yields

$$\log(y) = \log(Ax^k) = \log(A) + \log(x^k) = \log(A) + k \log(x) \quad (2.)$$

where the second equality follows from the product property of all logarithms (to any base) and the third equality follows from the power property of all logarithms. From the third equality in eqn. 2, it is clear that a plot of $\log(y)$ vs. $\log(x)$ will yield a straight line with intercept $\log(A)$ and slope k .

Since the slope points P_1 and P_2 are on the power law curve eqn. 1, eqn. 2 applies to both of the slope points so we can write,

$$\log(y_1) = \log(A) + k \log(x_1) \quad (3.)$$

$$\log(y_2) = \log(A) + k \log(x_2) \quad (4.)$$

Subtracting eqn. 3 from eqn. 4 yields,

$$\log(y_2) - \log(y_1) = k \log(x_2) - k \log(x_1) = k[\log(x_2) - \log(x_1)] \quad (5.)$$

Solving eqn. 5 for the power law exponent k yields.

$$k = \frac{[\log(y_2) - \log(y_1)]}{[\log(x_2) - \log(x_1)]} \quad (6.)$$

Since k is the slope on a plot of $\log(y)$ vs. $\log(x)$, *eqn. 6. is the slope formula for plots on log-log paper.*

Note that eqn. 6 can be computed in any convenient logarithm base, including the base e natural log, \ln and base 10 \log , \log_{10} . This is because the steps used hold for any base logarithm.

Determination of the power law prefactor A

Once the power law exponent k has been obtained from eqn. 6, it is simple to calculate the prefactor A by using either of the slope points $P_1 = (x_1, y_1)$ or $P_2 = (x_2, y_2)$, since both of them must satisfy eqn. 1. Substituting the values into eqn. 1 and solving for A yields.

$$A = \frac{y_1}{x_1^k} = \frac{y_2}{x_2^k} \quad (7.)$$

Miscellaneous Comments:

Notice that the scales on log-log plots are never zero. This is because $\log(0) = -\infty$! You can start the scale at any integer power of 10, successive cycles will increase by a factor of 10. (i.e. $10^{-3} = 0.001$, $10^{-2} = 0.01$, $10^{-1} = 0.1$, $10^0 = 1$, $10^1 = 10$, $10^2 = 100$, etc.)

As we have seen in this lab and associated spreadsheets, the addition of relatively small amounts of noise to data (<10%), can have much larger effects on the analysis and results obtained from the data. For our example of finding constant acceleration, in finding the slope, we subtract two numbers in the numerator, and two numbers in the denominator. If the slope is small, we may have a situation where the two numbers in the numerator are very close to one another, so close that the uncertainty or noise in each number is on the order of their difference. When this happens we can have very large errors in the calculation of the slope. No such differences occur in the determination of the y intercept, so its determination is not subject to such large possible errors.

There is another method to find the exponent k on a log-log plot by using a ruler to measure the rise and run and thus find the slope from the ratio of distances. However, without going into details, this method is complicated when the log scale cycles are different distances for the x and y directions. One must take account of this difference in scale to determine the power law exponent from the slope when using the measured distances for the rise over the run from the graph to calculate the slope.