Graphical Analysis and Errors – MBL

I. Graphical Analysis

Graphs are vital tools for analyzing and displaying data. Graphs allow us to explore the relationship between two quantities -- an independent variable usually plotted on the x-axis and a dependent variable usually plotted on the y-axis. Consider an experiment measuring the pressure of a gas while its temperature is varied. The measured pressure is the dependent variable, depending on the temperature set by the experimenter.

In plotting a graph by hand, the key is to find a way to plot the two variables so that a straight line is obtained. If you use a computer, on the other hand, software can be used to fit data with a variety of functions. One purpose of this experiment is to compare how you graph by hand to how you can use a computer to help you graph.

In general, you will face two types of situations:

1) You have no theory relating the two variables.

2) You have a theory that gives the functional relationship between the variables but the theory contains parameters you need to determine.

In case 1, the relationship between variables can often be determined by investigating different types of graphs. In case 2, the slope and intercept(s) of the graph can be used to determine the values of parameters.

II. How to Draw a Good Graph

Keeping these points in mind will make for more effective graphs:

1) Determine the minimum and maximum values of the data in both the x and the y directions.

2) Choose a scale for the tick marks on the two axes. Usually you will want to use as much of the graph paper as possible and make the scale on a given axis just sufficient to cover beyond the minimum and maximum values to be plotted. On the other hand, it's also best to use simple values for the tick marks. Some compromise between these objectives is usually necessary.

3) Plot the data carefully (including error bars if applicable). It's often a good idea to plot the data as you take it instead of waiting until you finish collecting it. One of the lab partners may be able to record data in a table as well as plotting it while the data is coming in.

III. Linear Relationships

If the relationship between the two variables is linear there's a simple relationship between the values on the x- and y- axes:

$$y(x) = m_1 x + b_1 \qquad (1)$$

Simply plot the points on regular graph paper, and draw a straight line of best fit through the points. The best-fit line is your estimate of the straight line that comes closest to passing through or near all the points. The slope of this line is the value of m_1, and the y-intercept (the y-value where the line crosses the y axis) is the value of b_1.

Things to keep in mind when determining slope:

1. Use points as far apart as possible, to minimize effects of uncertainties in reading points off the graph.

2. Use convenient points on the graph that the line goes through, not two points you plotted on the graph. One problem with points you plotted is that you often know those values to greater precision than you can read points off the graph. Other issues relate to why you're using a graph in the first place:

• the graph tells us whether a linear relationship really holds between the two variables

• the graph allows us to easily average over the fluctuations between data points

• the graph can be used to determine an experimental uncertainty for the slope and y-intercept.

For a graph drawn by hand, to use the graph to estimate uncertainty in the slope and y-intercept you should draw three straight lines on the graph:

1. A line of best fit with slope m_1 and intercept b_1.

2. A max-slope line (the steepest straight line you can draw that is still a reasonable fit to the data points), with slope m_2 and intercept b_2.

3. A min-slope line (the straight line with the least slope that is still a reasonable fit to the data points), with slope m_3 and intercept b_3.

Your values of slope and y-intercept can then be specified with approximate uncertainties:

Slope: $m = m_1 \pm (\, | \, m_2 - m_3 | \, /2)$, where $| \, |$ means take the absolute value. $\qquad (2)$

Y-intercept: $b = b_1 \pm (\, | \, b_2 - b_3 | \, /2)$ $\qquad (3)$

For graphs drawn on the computer the software generally uses a more sophisticated least-squares fitting approach to determine the slope, intercept, and uncertainties.

IV. Nonlinear Relationships

Many times we encounter situations in which the relationship between variables is not linear. We will examine two examples.

IV.A Power-law Relationships

In many situations the relationship between two variables is a power-law:

$$y(x) = a\, x^b \tag{4}$$

For instance, we will see later in the course that the period T of a planet orbiting the sun varies as the radius of its orbit R to a power b. Thus:

$$T = aR^b \tag{5}$$

To test whether a relationship is a power-law, and determine the parameters a and b, we can use a log-log plot. To see how this works, we take the logarithm of both sides of Equation (4):

$$log\, y = b\, log\, x + log\, a \tag{6}$$

While y is not linear in x, $log y$ is linear in $log x$. If we computed the logarithms of the points and plotted them, we would get a straight line. To make things easier, we can use log-log graph paper, which automatically plots the logarithms without computation on your part.

Using log-log paper:

Each logarithmic axis has a few cycles, with each cycle covering a factor of 10. If the values on the x axis fall between 3 and 500, for instance, you would need three cycles. The first cycle would cover values between 1-10, the second between 10-100, and the third between 100-1000.

You do not need to take any logarithms – the paper does it for you.

As Equation (6) shows, if the cycles in the x- and y-directions have equal length, the slope of a log-log plot is equal to the exponent b. To find the slope of the best-fit line, pick two convenient points that are quite far apart along the best-fit line. If these points are (x_1, y_1) and (x_2, y_2) the slope is:

$$b = (log\, y_2 - log\, y_1)\, /\, (\, log\, x_2 - log\, x_1) \qquad . \tag{7}$$

The parameter a is the value of the y-intercept, although here the y-intercept is where $log(x) = 0$, which is where $x = 1$. Note that b has no units while a has whatever units are needed to make Equation (4) dimensionally correct.

Graphical Analysis and Errors

IV.B Other Nonlinear Relationships

Sometimes you may encounter nonlinear relationships which are neither power-law (which can be handled with log-log paper) nor exponential (which can be handled with semi-log paper). If you know the mathematical relationship between the variables, however, you may still be able to devise an approach that will yield a straight-line plot.

Suppose for instance, that you drop a ball and it accelerates due to gravity. You will see in class that the velocity of the ball v after falling a distance d is:

$$v^2 = v_0^2 + 2gd \tag{8}$$

Here g is the acceleration due to gravity and v_0 is the initial velocity of the ball. By measuring the velocity of the ball at various distances d, we can determine g. However, if we do not know v_0, then we cannot directly use any of the graphical methods we have discussed thus far because the relationship is not linear, exponential or power-law. On the other hand equation (8) shows that the graph of v^2 vs. d should be a straight line, and the slope and intercept will yield values of g and v_0. The trick is always to find some way to produce a straight line.

V. Summary of Graphical Analysis

1. Graphical analysis is a powerful tool for determining the relationship between two experimental quantities.

2. The key to obtaining quantitative results from graphical analysis is plotting the variables in such a way that a straight line is obtained.

3. If the relationship between the variables is exponential, then a plot using semi-log graph paper should be straight. The slope of the line gives the characteristic constant of the exponential.

4. If the relationship between the variables is a power-law, then a plot using log-log graph paper should be straight. The slope of the line gives the exponent.

5. If the relationship between the variables is nonlinear but not exponential or power-law, then you must search for another way of plotting the variables which yields a straight line.

6. Finally, ALWAYS plot data as you take it. This allows you to detect errors in your measurements before it is too late to correct them. It also shows where data points are most needed in order to determine the position and slope of the best-fit line. To find the correct minima and maxima for the plot, it is usually best to perform measurements at the end of the experimental ranges first. This also guarantees that the experimental apparatus is functioning properly over its full range.

VI. Error Analysis

VI.A Propagation of Errors

An important part of any experiment is evaluating the uncertainty in the final results. Conceptually, this process can be divided into two stages:

(1) Determining the uncertainty of the initial data

(2) Propagating this uncertainty into the final result calculated from that data.

The first process varies widely depending on the type of measurement being performed. Usually uncertainties in readings can be assigned based on a little experience and common sense.

The second part of the process --propagating the initial uncertainty in the data through various calculations and into the final result -- is a straightforward process. Suppose for instance that $z=sin\ \theta$, and we know the angle θ and its uncertainty $\Delta\theta$. What is the resulting uncertainty in z? Using calculus we know that the change in $sin\ \theta$ with a small change in θ is $cos\ \theta\ \Delta\theta$, where $\Delta\theta$ is measured in radians (this is very important!). Thus the uncertainty on z is $\Delta z = cos\ \theta\ \Delta\theta$.

As a second example, suppose z is the sum of x and y, which have uncertainties Δx and Δy respectively, i.e. $z=x+y$. What is the uncertainty in z? We might guess that $\Delta z = \Delta x + \Delta y$. However, if the uncertainties in x and y have opposite signs they will partly cancel and the error in z may be quite small. A simple summation of the two uncertainties is too pessimistic because it neglects this possible cancellation. The correct uncertainty is given by $\Delta z^2 = \Delta x^2 + \Delta y^2$.

The last page of this write-up has equations for error propagation in other situations.

Graphical Analysis and Errors: Exercises

1. Graph the data you obtained in the first pre-lab exercise using linear graph paper. Position should be on the y-axis and time on the x-axis. State your result completely (in other words, find values for the slope and y-intercept), including units. Using your graph, estimate the uncertainties on your slope and y-intercept.

Question 1: The data you obtained relates to the motion of an object. What does the slope tell you about the motion? What does the y-intercept tell you? (Hint: think about the units.)

2. In the second pre-lab exercise you measured the gravitational force F that one mass exerts on another a distance r away. We will now plot the force versus distance data on linear graph paper to see that there is not a linear relationship between F and r. If you are doing this lab assignment in lab classroom, you will enter your data into the *Graphical Analysis* program on the lab computer and use its plotting features to

make the graph. If you are doing this lab assignment at home, you can print your own graph paper at `http://www.incompetech.com/graphpaper/`. After plotting the data, you will see that the relationship is nonlinear. If you are doing this assignment in lab class, you can use the *Graphical Analysis* software to determine the appropriate function to the data to obtain an equation relating F and r.

Since the data is not consistent with a linear relationship, try making graphs on log-log and semi-log (x-axis is linear with y-axis logarithmic) paper. Which works best?

Question 2: What do your results tell you about the relationship between F and r? From your best fit line on the appropriate graph, obtain a mathematical relationship between force and distance. If you used the *Graphical Analysis* software, how do the two methods compare? Do your two equations relating F and r agree with one another?

3. In the third pre-lab exercise you measured distances and speeds for an object that was released from an initial height of 4.0 m with an unknown initial speed. Your job is to determine the acceleration due to gravity and the initial speed of the object.

If you make a linear plot of your values for v and d (where d is distance from a height of 4.0 m, as opposed to the actual y coordinate) either on graph paper (if you're working at home) or using the *Graphical Analysis* software (if you're working in the lab classroom) you should see that the graph of v vs. d is not linear. Now figure out what you should graph on the x- and y-axes to produce a linear graph. If you are working at home, calculate the necessary numbers for the new graph and make the plot. If you are working in the lab classroom, you can create one or more new columns in Graphical Analysis to produce the data you need for the new graph: to do this go to the Data menu, select New Column, and then Calculated. Enter a name, and select v or d from the Columns pull-down menu. You should then be able to set up a new column with just about any function of v or d you want.

Whether you've created the new graph by hand, or used the *Graphical Analysis* software, once you've got a linear graph, find its slope and y-intercept.

Question 3: What is the value of the acceleration due to gravity in the simulation? Is this a reasonable value?

Question 4: What is the initial speed of the object (i.e. the speed when it's at a height of 4.0 m) ?

Error Propagation

Question 5. Suppose you know that $F = m\,a$, where F is the force, m is the mass and a is the acceleration. If you measure $m = 2.0 \pm 0.1$ kg and $a = 5.1 \pm 0.2$ m /s^2, what is the value and uncertainty of F?

Question 6. Suppose we have a relationship $y = a + sin\,\theta$. You measure $a = 2.3 \pm 0.2$ and $\theta = 10.0 \pm 1.0$ degrees. What is the value and uncertainty of y?

Error Propagation

Below we give the equations governing the propagation of errors when combining two measurements, $x \pm \Delta x$ and $y \pm \Delta y$, or when taking a trigonometric function of an angle $\theta \pm \Delta\theta$ where $\Delta\theta$ is in RADIANS.

1. Addition: $x + y = z$

$$(\Delta z)^2 = (\Delta x)^2 + (\Delta y)^2$$

2. Subtraction: $x - y = z$

$$(\Delta z)^2 = (\Delta x)^2 + (\Delta y)^2$$

3. Multiplication: $xy = z$

$$\Delta z = xy\sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2}$$

4. Division: $x/y = z$

$$\Delta z = \frac{x}{y}\sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2}$$

5. Exponentiation: $x^n = z$

$$\Delta z = nx^{n-1}\Delta x$$

6. Sine function: $\sin\theta = z$

$$\Delta z = (\cos\theta)\,\Delta\theta \qquad\qquad \Delta\theta \text{ in RADIANS}$$

7. Cosine function: $\cos\theta = z$

$$\Delta z = (\sin\theta)\,\Delta\theta \qquad\qquad \Delta\theta \text{ in RADIANS}$$

8. Tangent function: $\tan\theta = z$

$$\Delta z = (\sec^2\theta)\,\Delta\theta \qquad\qquad \Delta\theta \text{ in RADIANS}$$