
PRESENTING DATA GRAPHICALLY

“ ‘A crowd in a little room -- Miss Woodhouse, you have the art of giving pictures in a few words.’ ”

-- *Emma*

4.1 INTRODUCTION

“Draw a picture!” is an important general principle in explaining things. Frank Churchill’s remark to Emma Woodhouse notwithstanding, “the art of giving pictures in a few words” is not nearly as useful as a good diagram or graph, because most people process visual information much more quickly than information in other forms. Graphing your data shows relationships much more clearly and quickly, both to you and your reader, than presenting the same information in a table.

Typically you use two levels of graphing in the lab. A graph that appears in your final report is a **higher-level graph**. Such a graph should be done very neatly, following *all* the presentation guidelines listed at the end of this chapter. It’s made primarily for the benefit of the person reading your report.

A **lower-level graph** is a rough graph you make for your own benefit; they’re the ones the lab assistants will hound you to construct. These lower-level graphs tell you when you need to take more data or check a data point, since any strange measurements really stand out in a graph. They’re most useful when you make them in time to act in response to what you see. This means that you should graph your data roughly *before you leave the lab* so you still have the chance to make more measurements. (That’s one reason we recommend that you leave every other sheet in your lab notebook free, so you can use that blank sheet to graph your data.) In graphing your data in the lab, you don’t need to be too fussy about taking up the whole page or making the divisions nice, but you should label the axes and title the graph to remind yourself later what it shows.

Graphing your data right after you have completed a set of measurements also flags regions in your data range where you should take more data. Typically people take approximately evenly spaced data points over the entire range of the controllable variable (the “independent” variable), which is certainly a good way to start. A graph of that “survey” data will tell you if there are regions where you should look more closely: regions where your graph is changing rapidly, going through a minimum or maximum, or changing curvature, for example. The graph helps you identify interesting sections where you should get more data, and saves you from taking lots of data in regions where little is happening.

Graphing each point as you take it, though, is *not* a good idea. Doing so is inefficient and, worse, can prejudice you about the value of the next data point. So take five or six data points and then graph them all.

For example, Figure 4.1 on the next page shows the original data taken on a phenomenon called mechanical resonance. All you need to know about resonance for our purposes right now is that the “amplitude” (a measure of the response of an oscillating system) depends on the frequency at which that system is perturbed (or “driven”) by an external oscillating force.

Notice that the experimenter initially chose driving frequencies in the first run that were approximately evenly spaced across the range shown in the graph. For this particular apparatus, the highest and lowest frequencies attainable with the equipment are easy to find, and the experimenter chose to space the frequencies evenly to get roughly 10 different frequencies over the range in fre-

quencies. You can see from the graph of the original data that the response doesn't change very much at either very high or very low frequencies, but near some intermediate frequency, between 4 and 6 cycles/s, something strange and interesting happens.

Mechanical Resonance

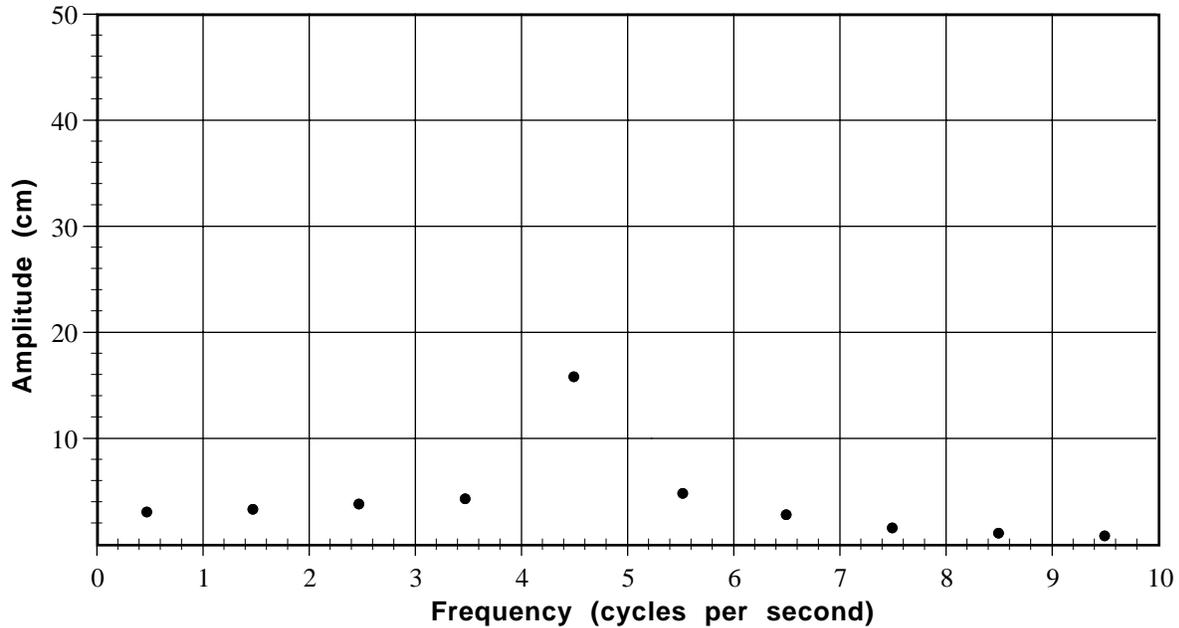


Figure 4.1: First set of data (black dots) for amplitude response versus driving frequency in an oscillating system. Notice that the data points are evenly spaced.

Mechanical Resonance

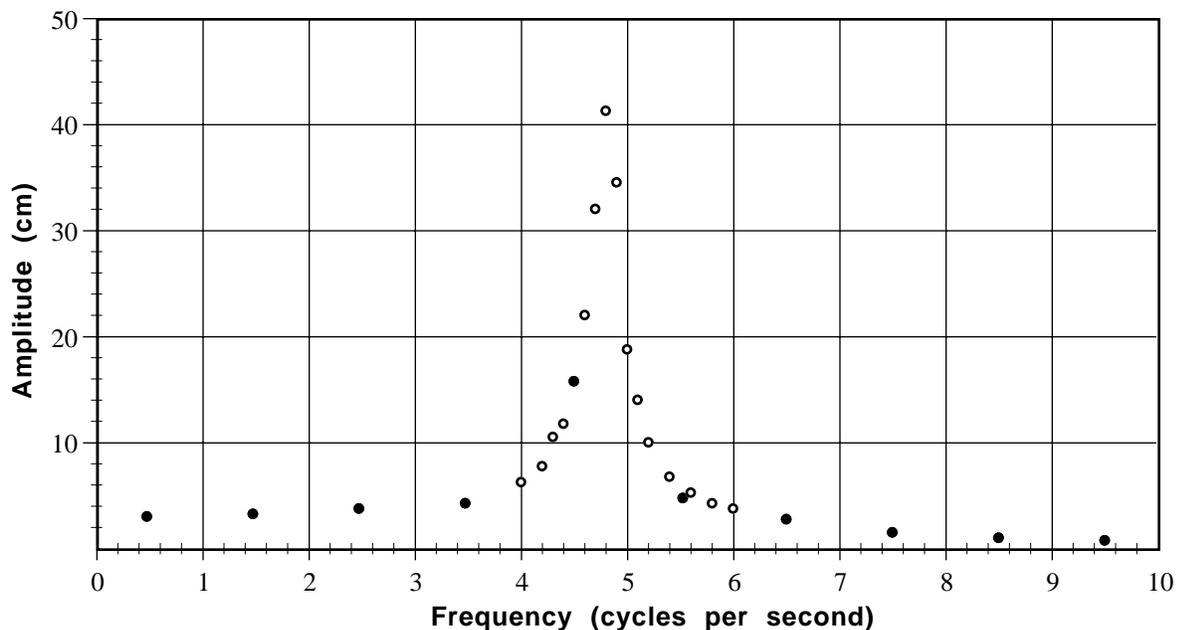


Figure 4.2: Amplitude vs. driving frequency in a resonant system, after adding more data points (white dots) between 4.0 and 6.0 cycles per second.

The experimenter noticed this, too, and went back to take more data in the interesting range of frequencies. The frequency spacing used in the second round is smaller than used in the first set by about a factor of ten, yielding 15 more measurements in the critical region. The result of adding the second set of measurements is shown in Figure 4.2. As you can see, the shape of the graph is now *much* better defined. Furthermore, the new data show that the anomalously high amplitude at 4.5 cycles/s is not a mistake (as one might think considering the other values). The experimenter *could*, of course, have taken data with the closer spacing over the entire frequency range, but that would waste time on measurements at both low and high frequencies where nothing much is happening. The strategy of taking coarsely spaced data and then backing up to take more data in interesting regions is a good compromise between completeness and efficiency. But remember that you usually can't identify the "interesting" regions if you don't graph your data to begin with!

4.2 ANALYZING YOUR GRAPH

Graphical data analysis is typically used as a euphemism for "find the slope and intercept of a line." You will find this semester that you spend a lot of effort manipulating your data so that the resulting graph is a straight line. As you will find throughout the semester, the slope and/or intercept of such a line often gives useful information about the physical system under investigation.

Determining the slope and intercept of a linear graph is such a common and important task that we have developed a computer program (described in Chapter 10 in this manual) to help you do it accurately. Many scientific calculators can also do this (although they almost never give the *uncertainty* in the slope and intercept). It is good to be able to estimate roughly the slope and intercept of a lower-level graph by hand, so that you can see if your measurements are at least roughly correct before you enter them all into the computer. Manual estimates of the slope and intercept also give you a check on the computer's results (allowing you to catch simple errors like entering the data in the wrong columns, for example). This process is so important that, although we have this fond hope that you learned how to do this in high school, we're going to review it anyway.

Imagine that you have constructed a lower-level graph of your data by hand, and it looks pretty linear. Start by drawing in by eye the line that you think best matches your data. The analytic procedure called **linear regression** (described in Chapter 10) gives the optimum result, but in fact an eyeballed "best fit" line will generally be quite close to the line found by linear regression. Your job now is to find the slope and the intercept of that line you've drawn.

The slope of a line is defined as the "rise over run," the change in the vertical coordinate value divided by the change in the horizontal coordinate value. To determine the slope, you must first choose two points on your line. They need not be actual data points but they *must* lie exactly on your line. They should be about as far apart on the graph as possible to minimize the effects of the inevitable experimental uncertainty in their position. Mark each of those points with a medium-large dot or ¥ and/or draw a circle around it. Read the coordinates (x_1, y_1) and (x_2, y_2) of each point off the graph. (As is the convention, the symbol x here represents the independent variable, plotted along the horizontal axis, and y is the dependent variable, plotted along the vertical axis.)

With these two coordinate pairs, you can calculate the slope m using the equation

$$\text{slope} = m = \frac{\text{rise}}{\text{run}} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \quad (4.1)$$

substituting your values for (x_1, y_1) and (x_2, y_2) . You can call either point (x_1, y_1) , as long as you assign the correct y -value to each x -value. Now that you have m , you can find the y -intercept from

$$y\text{-intercept} = b = y_1 - mx_1 \quad (4.2)$$

Again, you can call either point (x_1, y_1) as long as they both lie on the line. Since the y -intercept is defined as the value of y where a line intersects the y -axis (defined to be the $x = 0$ line), you can also read the intercept directly off the graph as long as the graph shows the $x = 0$ line.

The graph in the second sample lab notebook in section 1.5 illustrates the analysis of a lower-level graph. Note the use of ¥ 's to mark the points used to compute the slope.

4.3 UNCERTAINTY BARS

Individual data points plotted on any graph should include *uncertainty bars*, sometimes called *error bars*, showing the uncertainty range associated with each data point. You should show both vertical and horizontal uncertainty bars, if the uncertainties are large enough to be visible on the graph. If they aren't large enough, you should mention this in your report so we don't think you've forgotten them. You draw uncertainty bars by indicating the "best guess" value (typically either a single measured value or mean of a set of measurements) with a dot, and then drawing an "I-bar" through the dot, whose length spans the 95% confidence range of that value. An example of such an uncertainty bar is shown in Figure 3.3 below. The single data point plotted corresponds to a measured pendulum period T of $1.93 \text{ s} \pm 0.03 \text{ s}$ for an initial release angle q of $20^\circ \pm 2^\circ$. (The dotted lines are not part of the graph, but are included to show you how the point and the uncertainty bars are related to the axes. Notice also that the T -axis does not begin at $T = 0$.)

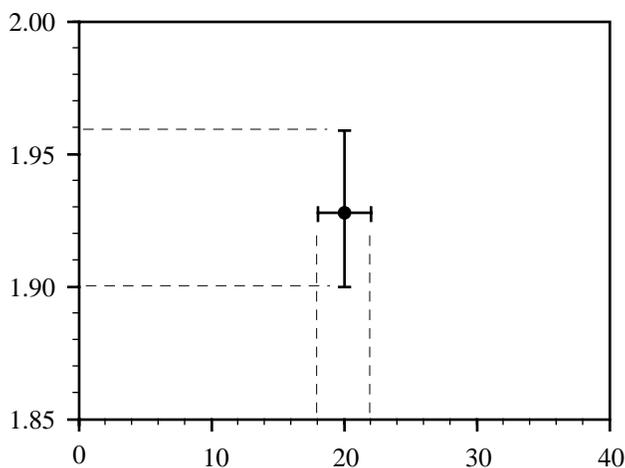


Figure 4.3. A graph of pendulum period T versus initial angle q , showing how uncertainty bars indicate the uncertainty ranges associated with the displayed measurement value.

4.4 PRESENTATION GUIDELINES for "higher level" graphs

You will create "higher-level" graphs for any written work that you submit for a project lab. These graphs should be more carefully and formally drawn and labeled than the lower-level graphs that appear in your lab notebook. Here are some guidelines for constructing these graphs:

1. Draw your graphs in pencil; mistakes are easy to make. If you wish, go back later and touch them up in ink. High-quality graph paper may be purchased from Connie Wilson, the physics department secretary for 10¢ a sheet. Computer-drawn graphs are fine as long as they comply with the remaining guidelines. The program described in Chapter 10 of this manual makes it very easy to produce graphs that automatically comply with all these guidelines, but the graphs produced other programs (such as *Excel* or *Cricket Graph*) may require extensive modification to fit the remaining guidelines.

2. Scale your axes to create as large a graph as possible consistent with the constraint that the divisions on the axes correspond to some nice interval like 1, 2, or 5 (times some power of 10). If you must make the graph smaller than full size to get nice intervals, OK, but check that you've picked the interval that gives you the largest possible graph (which will display your data in as much detail as possible). When using log-log or semi-log paper, choose paper with the number of cycles that gives the largest possible graph (see Chapters 5 and 13).
3. The lower left-hand corner need not be the point (0,0). Choose the range of values for each axis to be just wide enough to display all the data. If (0,0) does not appear on a hand-drawn graph, it is customary to mark the break in the axis or axes with two wavy lines (^a).
4. Mark the scale of each axis along each axis for the *entire length* of the axis.
5. Label both axes, identifying the quantity being plotted on each axis *and* the units being used.
6. Give each graph a title that summarizes the information contained in the axes and provides any additional information needed to distinguish this graph from other graphs in the report.
7. Give each graph a number (e.g., "Figure 2"), which you use in the body of the report or summary to refer quickly to the graph. (You can write such a number on a computer graph.)
8. Draw points and uncertainty bars as discussed in section 4.3.
9. If you calculate the slope and intercept of the graph from two points (rather than using the method of linear regression described in chapter 10), indicate the two points you used on the graph. Draw the line through the two points, label it "Best-fit line" (or something similar), and give its slope and intercept on the graph in some large clear space.

4.5 CHECKLIST FOR EACH GRAPH IN A WRITTEN REPORT

Use this checklist to make sure that any higher-level graph that you include in a submitted written report (as opposed to your lab notebook) has the correct features and format.

- The axes are scaled correctly with divisions equal to "nice" intervals.
- The graph is drawn as large as possible so that it fills the page.
- The scales on the axes have tick marks that run for their entire length.
- The axes have labels describing the variables they represent (including units).
- The measured data points are clearly plotted, including uncertainty bars.
- The graph has an appropriate title and figure number.
- The points used to calculate the slope and intercept are clearly marked (if that method is used).

EXERCISES

Exercise 4.1

Using the blank graph paper on the next page, create a *higher-level* graph of the data provided in Table 4.1. Use the checklist in section 4.5 to make sure that you have included everything.

Exercise 4.2

Draw what you think is the best possible line through the data points in the graph you just created in the last problem, and find the slope and intercept of this line.

Table 4.1: Kinetic energy per unit mass of a falling object as a function of distance fallen.

Distance fallen from rest (m)	Kinetic energy per unit mass (J/kg)
0.200 ± 0.003	2.00 ± 0.10
0.400 ± 0.003	3.78 ± 0.15
0.600 ± 0.003	5.65 ± 0.20
0.800 ± 0.003	8.01 ± 0.25
1.000 ± 0.003	9.82 ± 0.30

