Practice Guide (knowledge and skills in graphing).
Second Lab: Deflection of a beam.

Tabulating Data in tables with appropriate column headers.

We are used to labeling things with words, aren't we? OK, then, use all the words you like. But in the end I need to see a variable at the top of each column. You need to have defined this variable in the Introduction section, so I know where to find it. This is pretty standard fare, and is a very good thing to do as well. Later you will be substituting this data into equations. These equations will have the exact same variables as the column headers, and so the substitutions will work easily and smoothly. You could almost write a program to do it for you.

But what goes into the Data Table, and when do you use one? You use a Data Table if you are going to repeat a measurement, or one very much like it, many times over. In experiment 2 you are measuring deflections ['D' is the deflection of the loaded beam] for a bunch of different lengths ['L' is the length of the unsupported beam from the edge of the support to the position of the load].

So you would be forgiven if you thought the Data Table would have two columns: L, and D. In the manual Procedure section, however, you are given different instructions. It is suggested that you include two other columns, and a result column too. This is because you will be calculating D instead of measuring it directly. [$D = L_2 - L_1$, if L_2 and L_1 are defined correctly, which you should figure out on your own if possible]

This means that the table you will make is really not a Data Table at all. Data tables contain Data. Results tables contain Results. In this case the table is a 'Mostly Data, with One Result' Table.

Plotting a rough graph as data is taken.

At first this may seem straight forward, and it can be. You might think that the first step is to get all the data you need. Wrong. If you have all the data you need, then there is no purpose in drawing a rough graph. You might as well proceed to the Results section and be done with it.

The purpose of the rough graph is two-fold:

- It lets you know that the data is reasonable <u>as you take the data</u>. This means that you must be plotting the data on your rough graph as you take it.
- It lets you know where more data is needed or desirable, thereby guiding your experiment as it progresses.

Imagine taking a data point and then plotting it on this rough graph. Wouldn't the graph need to have the two axis labeled and scaled? Of course. But how do you do this without first taking the data? That is the key question that needs a little guidance from me. The next two sections deal with the problem of creating a rough graph before you have all the data.

Determining the practical domain and range of an experiment.

If we are clamping a meter stick to the edge of a desk, we know right away that the part of the meter stick that hangs over can't be shorter than 0 cm, and it can't be longer than 100 cm. So we have a practical limit on the values for L (the overhang length of the meter stick).

In an experiment where we set one thing, then measure the other, (in this case we set L and then measure D, then deflection) then the values we set (L) are the domain of the experiment, and the values we measure (D) are the range of the experiment.

So we know the possible domain of the experiment (i.e. We know that L will be somewhere between 0 and 100 cm), but we don't know the range. If the deflections (D) were not predictable, then we would have to take all of them and then find the maximum and minimum values of these D values for us to find the range of this experiment.

However, in this case we do know the behavior of the experiment. We have the form of the equation that is supposed to govern the behavior of the meter stick, and we have our general knowledge to work with. The equation tells us that $D = AL^n$. If 'n' is positive, then the deflection will be greatest when L is greatest, and 0 when L is zero. This is consistent with our general knowledge of things sticking out over an edge.

So to find the range of values (the max and min values of D), we only need to do the experiment (roughly) for the max and min values of L. What we say is 'D is monotonic in L', which means that D always increases if L increases. [aside: It could also mean that D always decreases as L increases]

Try this: sketch a graph that you think would correspond to the deflection, D, of the meter stick as you increase L, the overhang length.

So in summary, all we do to find the domain and range of the data (and of the data graph) is to take a measurement of the deflection at L = 0, and L = maximum, where 'maximum' is the maximum amount of the meter stick we can reasonably hang over the bench. These two data pairs (L_{min},D_{min}), (L_{max},D_{max}) give us the domain and range of the experiment.

The *domain* would be [L_{min} ...L_{max}], with a lowest value of L_{min}, and a *domain size* of ($L_{max} - L_{min}$).

For example, if L_{max} = 100cm, and L_{min} = 0cm, then the *domain size* is 100cm.

Scaling a graph (scales of 1, 2, and 5 per unit only).

So now that we have the domain and range of the graph. We are ready to scale this graph. Take a ruler with a metric scale and put it in front of you on the desk. Notice it has a 'scale'. It is numbered on the 'big ticks' as 0, 1, 2, ... How big are the 'big ticks'? If you look and see on this ruler, you will see that these big ticks (the ones that are numbered), are 1 cm each. The important thing is that each of these 'big ticks' on the ruler are 1 cm in size.

I am going to define a term here: I am going to define the *tick value* as the value we assign to size of one 'big tick'. From now on I am also going to refer to a 'big tick' as simply a *tick*. On the rule, the *tick* is 1cm. Therefore the *tick value* is 1cm/tick.

Normal graphs have two axes (horizontal and vertical), and *ticks* on each axis. The *ticks* are numbered. Each axis has a *tick value*, just like the ruler. But on a graph you can cheat. The *tick value* doesn't have to be 1cm. In fact, the tick value doesn't even have to be in the units of cm, even though in your notebook the graph paper ticks are placed 1cm apart, just like on your ruler. Graphs look like the ruler (especially when you draw an axis on the graph), but they aren't a ruler.

If I have graph paper, and I put my ruler against it along the horizontal direction, I can shift the ruler to match up the *ticks* of the ruler with the *ticks* on the graph. But the graph doesn't have numbers on it yet. If I transfer the numbers from the ruler to the graph, I will have a graph with a *tick value* of 1/tick. If I label this axis and include the units *cm* for this axis, then I will have a graph with a *tick value* of 1cm/tick, which will correspond to a ruler.

For each graph axis, there are 2 things you can change that effect the usefulness and interpretation of the graph:

The *tick value*. You can assign any *tick value* to either horizontal or the vertical axis. This includes both a magnitude and a unit. If my tick value on the X axis is 200 Joules/tick, then this is an entirely different graph from one in which my tick value on the X axis is 1cm/tick.

The offset. This is a bit more tricky. It means that you can pick any ONE *tick* on the axis and assign any value to this one *tick*. Every other *tick* is then determined by the *tick value*. We simply move along the axis one *tick* at a time, and add *one tick* value each move. Try this: pick any one tick on the graph axis. Label it 439. Now use a tick value of 1 Joule/tick and finish labeling the axis. Notice that the graph axis now represents about 10 Joules of 'space' on the axis, with 439 being included in this 'space'. This 'space' is the *domain* of the graph.

So now we can get down to business. We want the domain of our graph to match the domain of our data. In the Deflection experiment, we have a *domain* of [0cm .. 100cm], with a minimum of 0 cm, and a *domain size* of 100cm. But the graph in the horizontal direction has, maybe, 14 *ticks* to play with. We want the domains to match up. The trick is to first choose a good *tick value*.

When we choose a tick value we don't want weird tick values, because weird tick values make it hard to plot the data properly. So we do a simple little trick. We choose either 1, 2 or 5 as our *tick value* magnitude (times $10^{n)}$. This means I am choosing my *tick value* magnitude from a sequence like 100, 50, 20, 10, 5, 2, 1, 0.5, 0.2, 0.1, 0.05, 0.02, 0.01, ... (with the appropriate units, of course)

This is the same as the 10, 5, 2 method that I used earlier to estimate an uncertainty. This sequence of numbers is an easy one to remember and use.

In our case, I have 14 ticks and I need to cover a domain size of about 100cm. So lets be dramatic and choose a *tick value* of 100cm/tick. That means that for every tick I cover the entire domain size. Clearly, the domain of the graph would be 14 * 100cm/tick. This is way too big. I only want a domain size of something bigger than 100cm, but not too much bigger. So I go down in *tick value* from 100cm/tick, to 50cm/tick (in the 10, 5, 2 sequence). Now my domain would be 14 * 50cm/tick, which is still way too big. I try a *tick value* of 20cm/tick, then a *tick value* of 10cm /tick(which gives my graph a domain size of 140cm), then I try a *tick value* of 5cm/tick (which gives my graph a domain size of 70cm) , which is too small. I need my graph domain size to be bigger than the data domain size. Therefore, I go back to a *tick value* of 10cm/tick, and I am happy.

With a *tick value* of 10cm/tick on the x axis (the horizontal axis) I will have enough room to plot my data. Picking the *tick value* is also called 'setting the scale of the graph'.

The next tricky part is to pick the *offset*, which means we pick a *tick* on the graph and assign a value to it. We then fill in all the rest of the ticks around it until the whole axis is numbered. If we pick a bad *offset*, then the graph domain won't contain the data domain and we won't be able to plot all of our data on our graph. Here is a simple procedure picking the *offset*: Take the lowest value in the data domain. Round it down to the nearest number evenly divisible by the *tick value*. Now assign the lowest *tick* on our graph this value. For instance, suppose the lowest length we measure is 7.5cm and our tick value is 2cm/tick. Then we round 7.5cm down, but NOT to 7cm (because it is not going to be a natural tick value in the normal sequence 0, 2, 4, 6, 8, 10, 12 ....) Instead we round it down to 6cm, and assign this to the lowest tick on our graph. Our graph would be labeled starting at 6cm, and going up the available ticks in steps of 2cm until we run out of ticks.

Try this: with a data domain of [232y ... 265y], and 12 ticks on our graph to play with, what would be the tick value and offset? Ans: The domain size is 265y – 232y = 33y. We have 12 ticks. Starting at tick size of 50y/tick, (50y/tick * 12 ticks = 600y, which is WAY bigger than 33y), we work down the list of possible tick values through 20y/tick, 10y/tick, 5y/tick, 2y/tick (which give us a graph domain size of only 2y/tick * 12 ticks = 24y), which is less than the data domain size of 33y, so we go back to 5y/tick for our tick value. Then we take the lowest value in the data domain of 232y, and round it down to the nearest 5y (because our tick value is 5y/tick) and come up with 230y. We label the first tick as 230y, and go up in steps of 5y/tick for each tick until our graph runs out. Since there are 12 ticks the highest value would be 290y.

Make up some of your own data domains, and do what I did above until you can do it easily. You will be expected to do this on the first period of Experiment 2.

Plotting a Data Graph with some error bars:

See above for scaling the graph. Measure 1 data pair. Put this data pair into a table labeled at the top with the proper symbols you are using to represent the data. Now plot this data pair on this data graph before taking any more data. This will form a 'real time' image of the work you are doing. That way you can see that your data looks ugly right away, so you know to correct it early on in the experiment. This the same reason we do rough calculations (as described in the document for experiment #1). So we take a data pair, put it into the table, then plot it. Next, we take a data pair, put it into the table, then plot it. Next, we take a data pair, put it into the table, then plot it. Next, we take a data pair, put it into the table, then plot it. Next, we.... You get the idea.

A data pair is, of course, one value selected from the Domain (the x axis usually), and its corresponding measured value in the Range (the y axis usually). In the case of the 2nd experiment, we set an overhang length (L), and measure (or determine) the deflection (D) for this L. This, then, is a data pair. For instance, I might find a deflection, D, of 7.1cm for an L of 75cm. I need to plot the value 7.1cm on the y axis directly above the value 75cm on the X axis.

Naturally, I start by finding the value 75cm on my scaled X axis. I have chosen a tick value of 2cm for the X axis, and so I look along the X axis and find ticks at ...70, 72, 74, 76, 78.... Clearly my value is greater than 74, but less than 76, so my value lies somewhere between these ticks. The only remaining problem is to find out where it lies in this interval. I subtract the lower tick value (74cm) from my value. This leaves me with a remainder of 1cm. But my tick value is 2cm. I divide this remainder by the tick value, and get 0.5 (or half). This means that my value of 75 lies ½ way along between the ticks at 74 and 76. (But you knew that already.) On the graphs you will be dealing with, there are 10 minor ticks per tick (on the millimeter, as it happens). The tricky part is that I multiply this fraction times the number of minor ticks, and get 0.5 * 10 = 5 minor ticks. So when I plot this point (75cm), I move to the 74 tick, then move over 5 minor ticks, and in pencil I make a light little mark (usually in pencil).

Now I draw a very light construction line vertically up from this mark on my X axis (L) all the way up the page.

I now do the exact same process, but this time on the Y axis (D), and produce a very light construction line horizontally. Where these two lines intersect is the data point. I mark this point visibly with pen.

An invaluable addition to this graph is made by plotting the error bars for this point. On a data graph, not all points have to have error bars, but putting a few error bars on the graph gives a good idea of what 'graph features' are important and what are not. A graph feature is any indication of behavior other than a straight line. Features need to be detailed with enough data. However, if the 'feature' is inside the bounds of the uncertainty, then it is not (yet) significant.

To plot error bars is simple: I move my pen to the plotted point. Now for uncertainties in the value on the 'L' axis, I simply draw a line to the right of the data point the distance on the graph corresponding to the uncertainty in L (which is designated ΔL). If the uncertainty is 4cm, and the tick value is 10cm, then I move 6 minor ticks. If the uncertainty ΔL is 1cm, and the tick value is 0.5cm, then I move 2 ticks. I move (ΔL / *tick value*) ticks. Now I move back to the plotted point again, and move the same distance to the right. The width of the line I have drawn in the L direction is 2 * ΔL wide. I repeat this process for ΔD in the vertical axis. This gives me a cross, with the data point in the middle. The area of the graph that is 'covered' by the cross represents the area of the graph that the 'real' data is likely to be in (with slightly better than an even money bet).

Now that I have this data point with uncertainty on the graph, I look over my graph, and decide where I would next like to take a data point to best fill in the graph. I then choose an L that corresponds reasonably closely to this value and obtain another data pair. I plot this data pair, then repeat this whole process until I run out of time or my graph is reasonably filled in.

I do not connect the data points ever. Graphing data is not a game of 'connect the dots'. I do not draw any 'fit lines' through this data either. I look at it to see that a pattern is emerging, and imagine a 'fit curve' only.

So in summary:

- I measure a data pair, and put this pair into my Data Table.
- I find the two ticks that my measured value lies between on the X axis.
- I subtract the lower tick from my measured value, then divide by the tick value. This gives me a decimal fraction of a tick I need to move over from the lower tick. I multiply by the number of minor ticks, and this tells me the number of minor ticks I have to move over from the lower tick.
- I lightly mark the X axis at the location of my X value. I lightly draw a faint construction line vertically from this tick all the way up the page.
- I repeat this process for the Y value.
- I mark in pen the intersection of my construction lines.
- I plot the uncertainties in the X and Y values, thus making a cross through the best fit point.
- I review my graph to determine what I am going to use for my next X value.
- I then measure a data pair.... [Repeat until the graph is full or time runs out]

Try this yourself by pretending, and working through the entire steps above, or do a little experiment of your own.

Propagating uncertainties through the ln function.

Before you can produce results to plot, you must be able to calculate a result based on your Data. In experiment 2 you will be required to calculate the best estimate and its uncertainty when you calculate the value $X = \ln(L)$.

Without further ado, I will present the equation for the uncertainty of X when $X = \ln(L)$:

$\Delta X := \Delta L/L$

This equation makes no sense if you compare it to the equations for the uncertainty when you add, subtract, multiply, or divide numbers. In words, it says that the absolute uncertainty of X is the relative uncertainty of L. It is a direct consequence of the strange nature of the ln function. If you wish to better understand this strange result, then you might want to look at the definition of the ln function as an integral. Now look at the stuff under the integral sign (the differential).

But regardless of the source of the equation, it is never-the-less the equation you will be using in Experiment 2. Simply plug in the numbers for ΔL and L, and you will have the size of the error bar in X. By symmetry you obtain the result for ΔD. Symmetry is easy to see when we represent things symbolically.

Plotting a results graph with error bars.

A results graph is a graph of results. This means that you must first calculate results before you can plot them. Each result that is calculated will consist of a calculated best estimate and a calculated uncertainty. The uncertainty will have two significant figures, and the result will be converted to standard form and tabulated for ease of plotting.

Again, it makes sense to determine the domain and range of our results graph, then plot the results as we calculate them just as we did for the rough graph. The reasons for doing this are the following:

- If you are making a mistake in calculating the results, then plotting them as you go will make this clear.
- If you run out of time, then you can concentrate on plotting the important points first so that you can obtain an interim result (such as the slope of the line) without all of the results being plotted.

To plot a results graph in this lab, simply follow all the steps as you have learned them for a Data Graph, as outline in the sections above starting at *Determining the practical domain and range of an experiment***.**

The only features of the results graph that are distinct from the data graph are:

- The source of the pairs of numbers you use to plot the results graph
- The results graph is required to have error bars on it on all plotted points, whereas the data graph is not required to have these error bars on all the data points.

To obtain the Domain and Range of the Results graph, you must choose the two data values that will produce the maximum and minimum points on your graph. With these two values build a Results Table and enter these values in it, then proceed to scale the graph and plot these first two points. From then pick a data pair from your data, calculate a result, then put it into the Results Table, and plot the point with its uncertainty (error bars). Repeat until your are finished all your data or time runs out.

Drawing the best fit and two worst fit lines.

Once you have your results graph, you may then attempt to draw the *best fit line*. This is the line that you feel best represents all of the results, given the distribution of the results and their uncertainties.

The first problem people have when drawing the *best fit line* is that the 'big' points look more important. They are, after all, 'bigger'. But unfortunately the 'bigger' points have larger error bars on the graph. These points are in fact less important than the points that are 'smaller' (i.e. The ones that have smaller error bars). 'Bigger' points are more uncertain, and have less importance. We must try hard to weight the smaller points more when we are judging the *best fit line*s.

The second problem people have with drawing a *best fit line* is that they assume this line must go through some particular point. Generally all points have an uncertainty, and therefore it is not reasonable to assume that the *best fit line* goes through any one of them. The *best fit line* weighs all points according to their uncertainty, with the most uncertain points weighted the least.

Having said all of that, the *best fit line* m_B is drawn by using a transparent ruler and moving it about until we feel we have the *best fit line*. This means that it is subjective. Needless to say there is a more analytical method to do this, but we must not discount the amazing ability of our visual system to estimate a *best fit line* m_B. Generally it is pretty good at doing this.

There are two *worst fit lines*. There is the *worst fit high slope line* m_+,
and the *worst fit low slope line* m-. The names are misleading. The obvious *worst fit high slope line* is a vertical line, with slope of $m = \infty$. The obvious *worst fit low slope line* is a vertical line, with slope of $m- = -\infty$. This would produce a slope $m = m_B \pm \infty$, which is a pretty safe bet. This would be equivalent to suggesting that the slope lies somewhere between $-\infty$ and $+\infty$. The value of this determination is correspondingly meaningless. We want to make the even money bet that was discussed in the document covering the first lab for estimating the uncertainty.

Therefore, with the transparent ruler we start the *worst fit high slope line* m_+ at $+\infty$ (a vertical line), then rotate the ruler (and our line) clockwise until we are willing to make an even money bet that the 'real line' has a lower slope than m_+ . The 'real line' can be thought of as the one that would be determined by repeating and averaging the experiment many, many times over. Draw m_+.

Similarly, we find by m- by starting with a vertical line at $-\infty$, then rotating the transparent ruler counterclockwise until we are below the best fit line still, and are willing to make an even money bet that the 'real line' has a higher slope than this one. Draw m-.

We now have three lines: m_B, m_+, and m-.

From here we proceed as in the lab manual, Appendix B, Uncertainties in Graphs, Slope Calculations with Worst Fits.

Working with the equation of a line.

The bad news is that we assume in this lab that you know the equation of a line, and can recognize an equation of a line in many forms. In this particular case, we show you that

$\ln(D) = n \ln(L) + \ln(A)$

and note that if you substitute $x = \ln(L)$, $y = \ln(D)$, and $b = \ln(A)$ into this equation we get

$y = nx + b$

which is very similar to the equation $y = mx + b$, the equation of a line. This will be an equation of a line if n = constant, and b = constant, both of which are true in this case.

Therefore, once we calculate y and x as above and plot them, then we should get a straight line with a slope of n (instead of m).

If this is not all clear to you, then it is necessary that you revisit your prior math to re-familiarize yourself with this. You may also check on this with other sources and the math department.

Calculating the slope and its uncertainty from the graph.

Please refer to the manual, Appendix B, Uncertainties in Graphs, Slope Calculations with Worst Fits.