

Plots, Curve-Fitting, and Data Modeling in Microsoft Excel

This handout offers some tips on making nice plots of data collected in your lab experiments, as well as instruction on how to use the built-in curve-fitting routines in Microsoft Excel. Excel is a good utility program for data recording and plotting, and is actually used a lot by practicing engineers in industry. The main reason for its popularity is simply cost and convenience (most people have it on their computers) making information sharing very easy. It is also easy to learn and use. With a little extra effort you can write your own computational routines using the built-in VBA (Visual Basic) compiler.

If you already know how to create a basic X-Y plot on Excel, then skip ahead to page 3 and the section called “Changing the Plot Appearance”.

Simple X-Y Plots

Table 1 includes measured data on the current-voltage relationship of a diode that we can use for demonstration of the plotting and curve-fitting features of Excel. Enter the data in two columns as shown in the figure below, select the two columns and then choose “Chart...” from the “Insert” menu (or just click on the Chart icon in a toolbar if it is visible). You will then see a dialog box like that shown in the figure below. Select the “XY (Scatter)” chart type as shown. In this case there are some options (sub-types) that control whether each data point is highlighted by a marker of some kind, and whether a straight- or smoothed line is shown connecting the data points. That isn’t really important at this stage because you can always change the appearance later, but let’s start by choosing the smoothed-lines with data markers (highlighted selection in the figure).

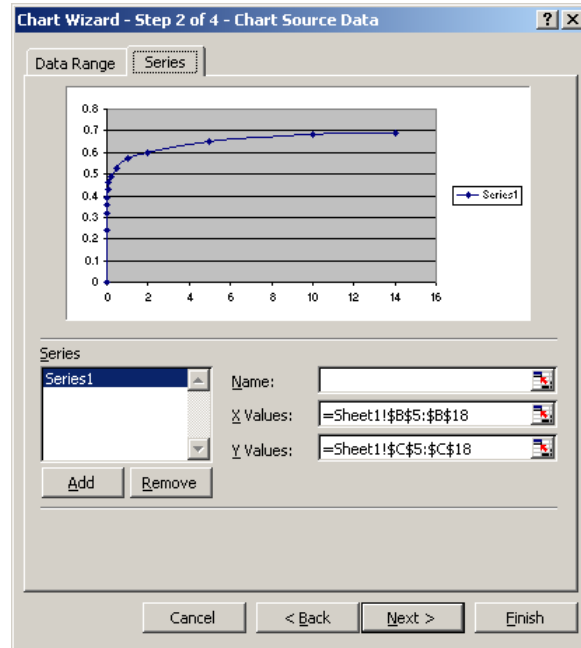
Table 1 – Sample Diode I-V Data

I [mA]	V [Volts]
0	0
0.001	0.24
0.005	0.34
0.01	0.36
0.02	0.39
0.05	0.43
0.1	0.46
0.2	0.49
0.5	0.53
1.0	0.57
2.0	0.60
5	0.65
10	0.68
14	0.69

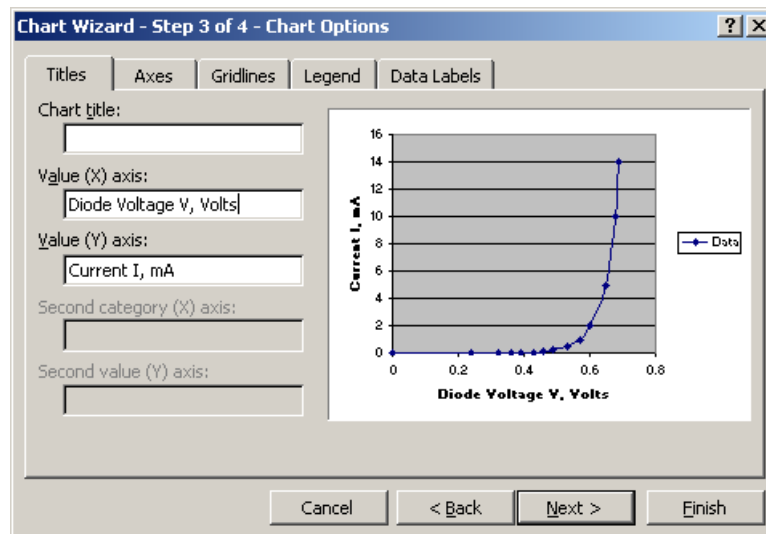
(Note: this is actual data recorded by Prof. York on a certain diode. In this experiment I expected an exponential dependence so I made a list of diode currents that would yield a nice plot, and then recorded the diode voltage that produced those currents.)

The image shows two screenshots from Microsoft Excel. The left screenshot shows a spreadsheet with data entered in columns B and C, labeled 'I [mA]' and 'V' respectively. A callout box points to the 'Chart...' option in the 'Insert' menu, with the text: "Step 2: Select 'Chart' from the Insert menu or toolbar". Another callout box points to the selected data range, with the text: "Step 1: Enter data, then select the columns you want to plot". The right screenshot shows the 'Chart Wizard - Step 1 of 4 - Chart Type' dialog box. The 'Standard Types' tab is selected, and 'XY (Scatter)' is chosen as the chart type. Under 'Chart sub-type', the option 'Scatter with data points connected by smoothed Lines' is selected. The dialog box also includes a 'View Sample' button and 'Cancel', '< Back', 'Next >', and 'Finish' buttons.

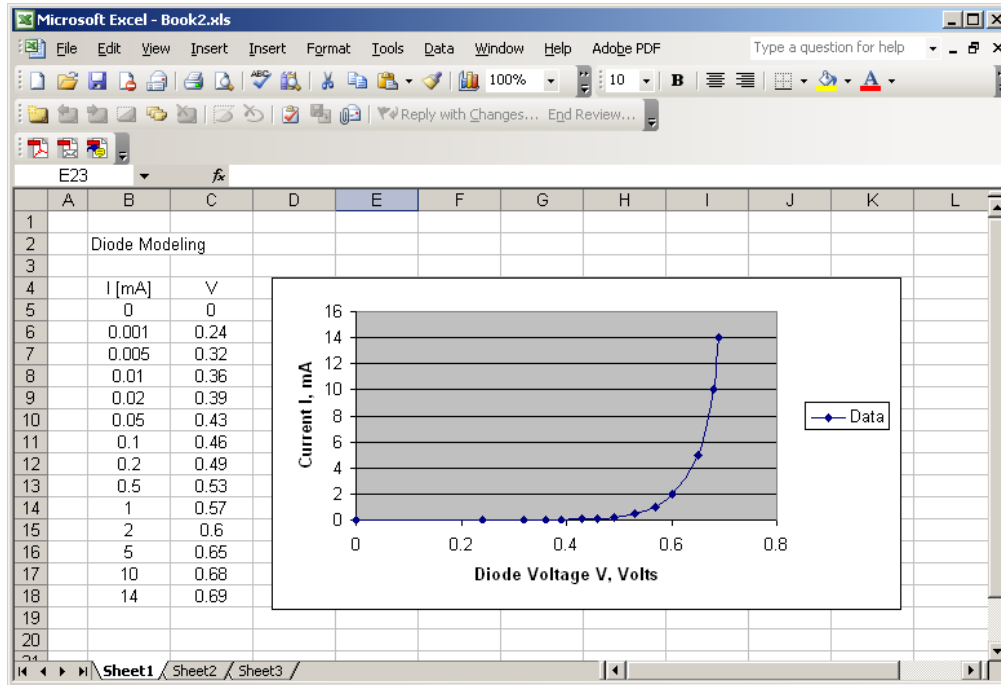
Now select “Next” and you will see the “Source Data” dialog box. Click on the “Series” Tab and you will see something like the following:



This gives a snapshot of what the graph will look like so far, and the source of data. It looks a little strange for a diode characteristic because Excel has assumed the first column represents the horizontal coordinate, so we need to switch the X and Y data range around. This is easy, just interchange the data ranges for the X and Y values. It is also useful to give the data series a name (like “Data” or “Measurements”) which will appear in the legend. Now select “Next” and you will see the “Chart Options” dialog box below. Here you can add labels for the x- and y-axes as I have already done:

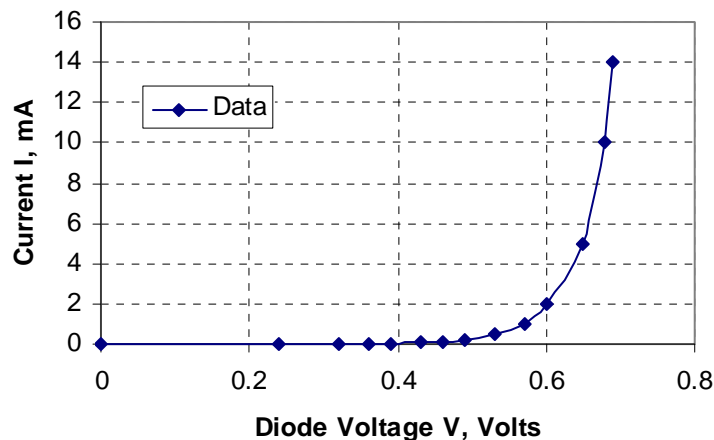


At this point you can just select “Finish”, and you will see the finished plot appear in the worksheet as shown below.



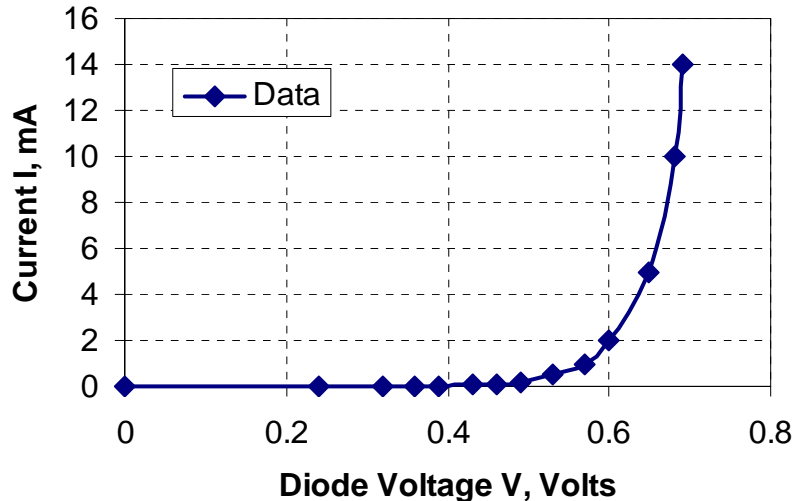
Changing the Plot Appearance

The plot above is the default Excel format for plots, which looks okay on a computer screen but not great in a printed document or presentation. If you right-click anywhere within the x-y axis, you will see a list of options for changing the plot appearance. First let's change the plot background from gray to white; just right-click in the plot and select “Format plot area...” to change this. Now right-click again and select “Chart Options” to add major gridlines on the x-axis. I also like to reformat the gridlines to be dashed lines as follows:



In the above I also moved the legend inside the chart (just drag it in there), resized the plot, and eliminated the border around the chart area.

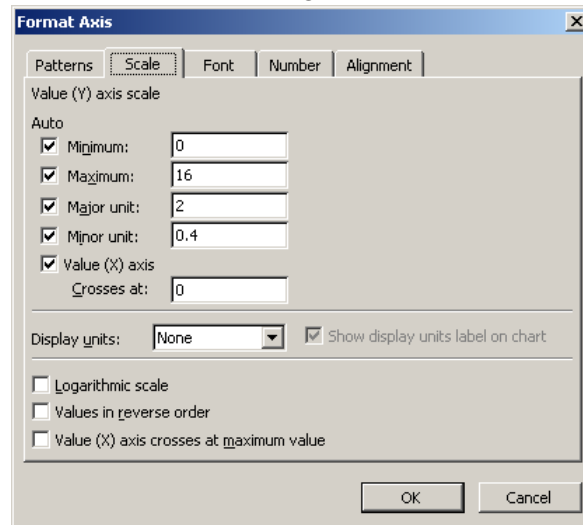
This is starting to look like a nicely formatted plot, but there is still one change I would strongly recommend, and that is increasing the size of the markers and the weight of the connecting line, and the fonts. Do this by double-clicking on one of the data points to launch the appropriate dialog box. You can increase the font size of the axis labels and legend by double-clicking on those labels and making the appropriate change. After doing this I get the following:



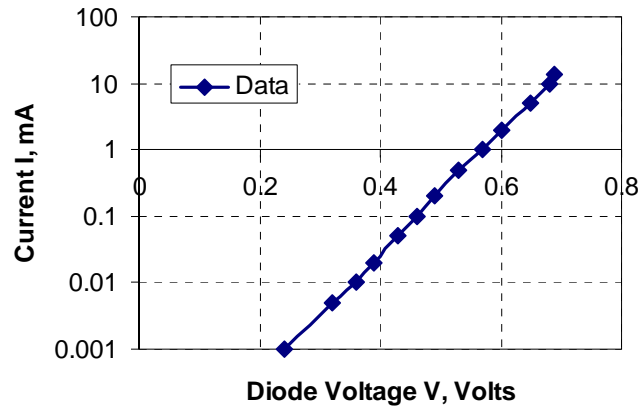
This is a suitably formatted plot for a lab report or power-point presentation.

Log Scales

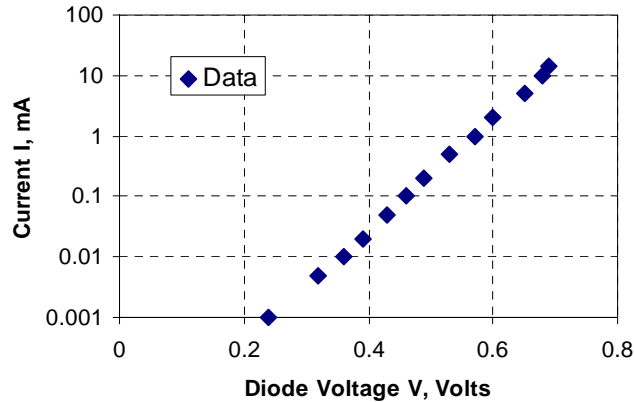
Since many of the diode data points involve small currents, they appear close to zero on the linear scale shown above. A logarithmic scale is sometimes used to effectively expand the scale for small currents. This is done in Excel by double-clicking on the appropriate axis (y-axis in this case) and then selecting the “Series” tab in the “Format Axis” dialog as shown. At the bottom you see a check box labeled “Logarithmic Scale”; click this and press “Okay”



Now your plot should look something like this



The data seems to fall on a straight line, which is what we'd expect for an exponential dependence. A couple of minor formatting issues can be addressed here; first note that the x-axis intersects the y-axis at 1 mA; on a log plot like this we'd prefer to have the x-axis and labels lie along the bottom of the chart. You can do this by going back to the "Format axis" dialog box (by double-clicking on the y-axis) and typing in the appropriate intercept value for the x-axis in the "Value (X) axis crosses at:" box. In this case the appropriate value is 0.001. Note that you could also adjust the limits of the vertical scale if desired, but in this example Excel has automatically chosen appropriate values. So we now have something that looks like this:



Here I have also removed the line connecting the data points in preparation for the next step, which is to add a second "Model" curve to compare the data against our diode model. Just double-click on one of the data points and click the "Line: None" box in the Patterns tab of the "Format Data Series" dialog.

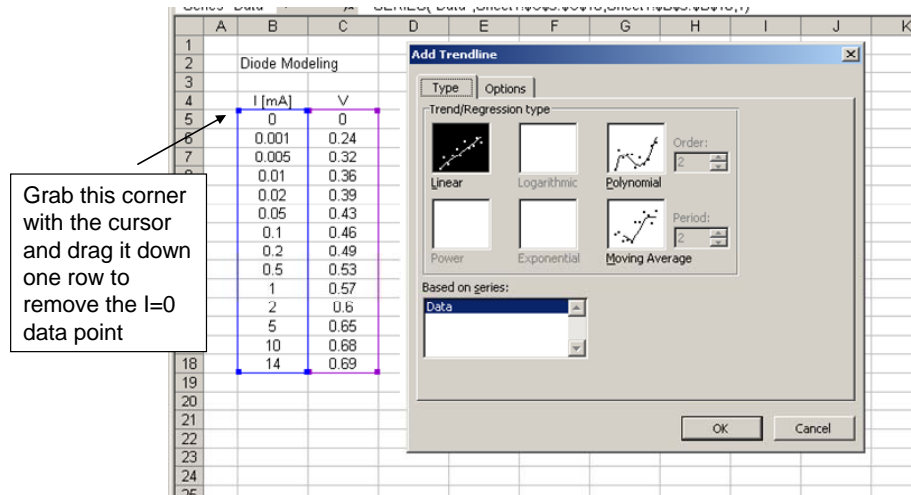
Curve-Fitting, or "Trendlines"

As you know, diodes are usually modeled by a relationship of the form

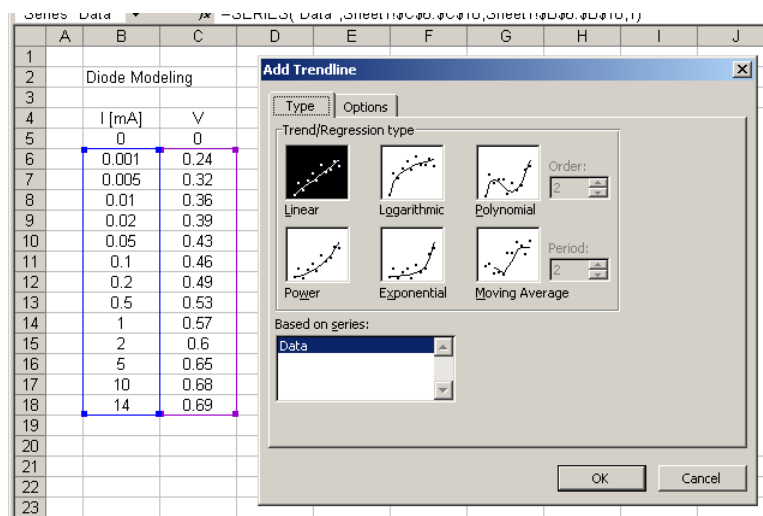
$$I = I_s \left[e^{qV/nkT} - 1 \right] \quad (0.1)$$

where n is the ideality factor, I_s is the reverse saturation current, and $kT/q \approx 26$ mV at room temperature. How can we use the measured data to determine appropriate parameters for the saturation current and ideality factor for this diode? There is an easy way to do this

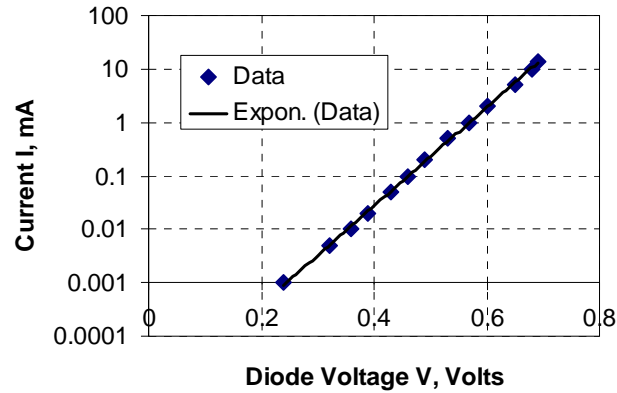
using the “Trendline” feature in Microsoft Excel. Just right-click on one of the data points in the chart and select “Add Trendline...”; you will then see a new dialog box appear as shown below:



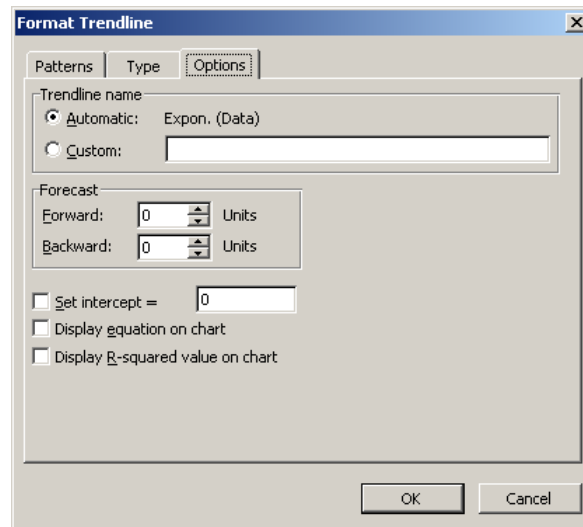
You can see that there are a few different options for curve-fitting the data, but the important one (“Exponential”) is grayed out. Why? The reason is that our data set includes an $I(0)=0$ data point, which you can never theoretically have on a simple exponential model! So the trick here is to remove that data point from the plot; the easiest way to do this is to just grab one of the top corners of the boxes enclosing the data and adjust the range to exclude this point. Do that for both the current and voltage columns, and now you should see all of the available curve-fitting options when you select “Add Trendline...” again:



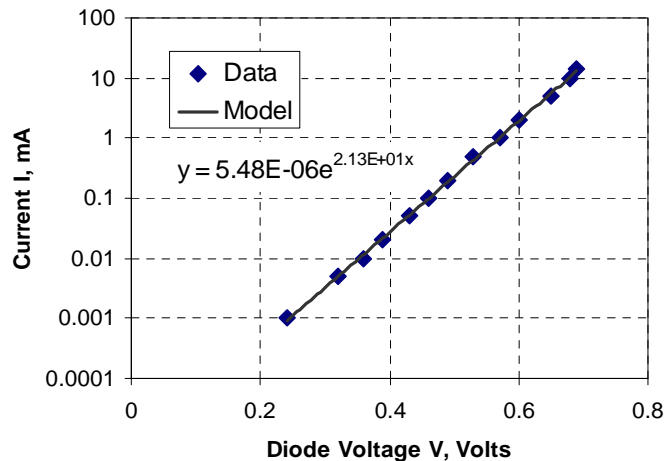
Now select the exponential model and hit “Okay”. Excel will find the best fit of the data to a model of the form $I = Ae^{\alpha V}$ and plot this on the chart as shown below:



A pretty good fit in this case. But we'd really like to know what the curve-fit parameters are, right? Double-click on the trendline and select the "options tab to get the following:



Here we want to select "Display equation on chart". You can also change the name of the trendline to "Model". Now with a few additional but minor formatting changes we have:



The last step is to simply compute the parameters we want from the information Excel has provided. In this case Excel has told us that $I = Ae^{\alpha V}$ is a pretty good fit to the data when $A = 5.48 \times 10^{-6}$ and $\alpha = 21.28$. Note that our vertical scale is in milliamps, so the fitting parameter A also has these units; in other words, $A = 5.5$ nA. Now, going back to our diode model in (0.1), we see that for $I \gg I_s$ the model reduces to $I \approx I_s e^{qV/nkT}$, so we must have

$$I_s \approx A = 5.5 \text{ nA}$$

$$\frac{q}{nkT} = \alpha = 21.28 \quad (0.2)$$

You can use Excel to compute the ideality factor from the last equation. Assuming the data was taken at room temperature I get $n \approx 1.8$, which is comfortably in the range we'd expect for real devices.

Plotting Multiple Data Series

As an example of plotting multiple curves, let's make a plot showing how the diode would behave at other temperatures using the model parameters just provided by our curve-fitting exercise. Note that the saturation current I_s is temperature dependent, approximately given by the relation

$$I_s(T) \propto T^3 e^{-E_g/kT} \quad (0.3)$$

where E_g is the bandgap energy of the semiconductor (1.13 eV for Silicon). If we know the saturation current at some reference temperature T_0 , then we can write

$$I_s(T) = I_s(T_0) \left(\frac{T}{T_0} \right)^3 \exp \left(-\frac{qE_g}{k} \left(\frac{1}{T} - \frac{1}{T_0} \right) \right) \quad (0.4)$$

Below I've made some room in the original spreadsheet so that columns F-G-H will describe the predicted currents at the three different temperatures entered into boxes F9, G9, and H9.

	A	B	C	D	E	F	G	H
1								
2	Diode Modeling Example							
3								
4					TO:	300		
5					Is(T0):	5.48E-06		
6					n:	1.817		
7								
8		Measured at 300K				Temperature, K		
9						290	300	310
10		I [mA]	V		V			
11		0	0		0.00			
12		0.001	0.24		0.05			
13		0.005	0.32					
14		0.01	0.36					
15		0.02	0.39					
16		0.05	0.43					
17		0.1	0.46					
18		0.2	0.49					
19		0.5	0.53					
20		1	0.57					
21		2	0.6					
22		5	0.65					
23		10	0.68					
24		14	0.69					

Annotations in the image:

- Constants found from data-fitting at room temperature (pointing to cells E4-E6)
- Temperatures (in Kelvin) at which we will compute the diode current in columns F-H (pointing to cells F9-H9)
- Enter the first two points and drag down to fill in the rest (pointing to cells E11-E12)

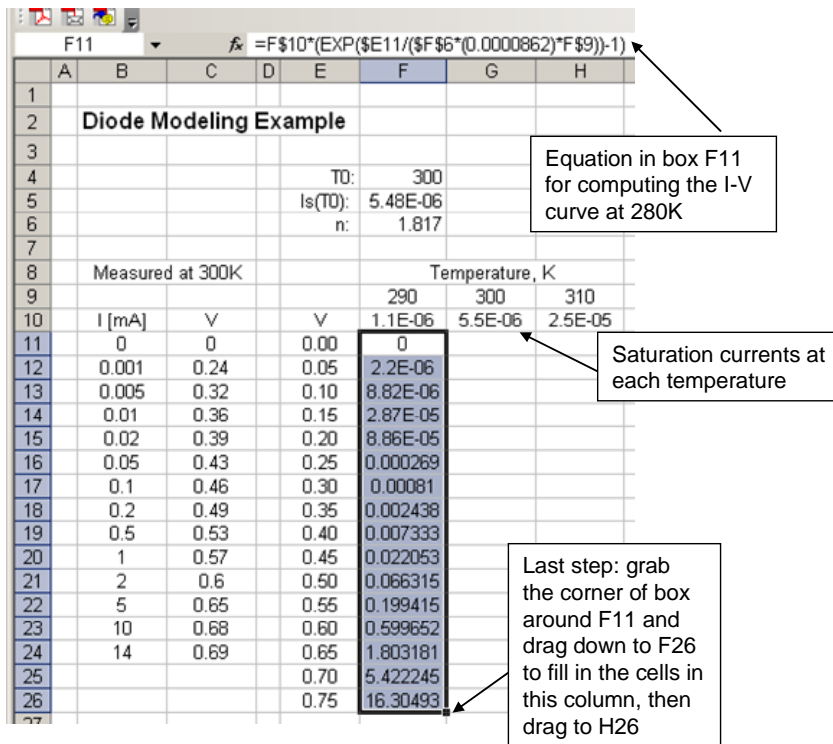
The fitting constants for the diode saturation current and ideality factor and the measurement temperature T_0 are included in F4:F6 as shown. To define the voltages at which we'll compute the currents, enter the first two points (which define the starting value and the increment) and then drag downward to fill in the rest as shown. The range of 0.0 to 0.75 Volts should be fine.

Now we can enter the equation for the diode current in the appropriate boxes. There is a simple way to do this in Excel such that we only need to enter the formula once. Start with the temperature-dependent saturation current. First click on the F10 cell and enter the formula from (0.4) as shown below (formulas are always preceded by an "=" sign). Note that the temperature is referenced as F\$9; this tells Excel that no matter which cell this formula ends up in when we drag it around, we always want to keep the "9" row reference constant. Similarly, the measured diode saturation current at T_0 is referenced as \$F\$5, which tells Excel to keep both the column and row reference constant when the formula moves. Without the dollar sign in front of a row or column reference, the reference will change according to the relative movement of the cell content.

	A	B	C	D	E	F	G	H	I
1									
2		Diode Modeling Example							
3									
4					T0:	300			
5					Is(T0):	5.48E-06			
6					n:	1.817			
7									
8		Measured at 300K					Temperature, K		
9							290	300	310
10		I [mA]	V		V	1.1E-06			
11		0	0		0.00				
12		0.001	0.24		0.05				
13		0.005	0.32		0.10				
14		0.01	0.36		0.15				
15		0.02	0.39		0.20				
16		0.05	0.43		0.25				
17		0.1	0.46		0.30				
18		0.2	0.49		0.35				
19		0.5	0.53		0.40				
20		1	0.57		0.45				
21		2	0.6		0.50				
22		5	0.65		0.55				
23		10	0.68		0.60				
24		14	0.69		0.65				
25					0.70				
26					0.75				

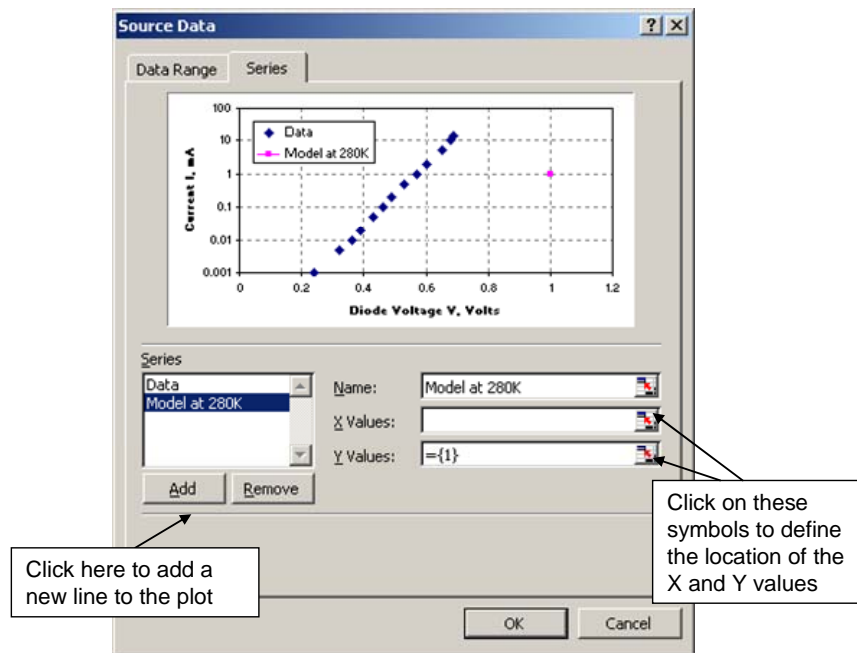
Once the formula is entered you can grab the corner of the box around F10 and drag it to H10 to fill in the saturation current for the other temperatures.

Now enter the equation for the full diode current from (0.1) in box F11 as shown below. Note that this time we reference the voltage as \$E11, because we want the voltage to change as we drag the formula *down*, but we don't want it to change when we drag it *sideways*. Now grab the corner of the box surrounding F11 and drag it down to F26, the last cell in that column; you should now see the following

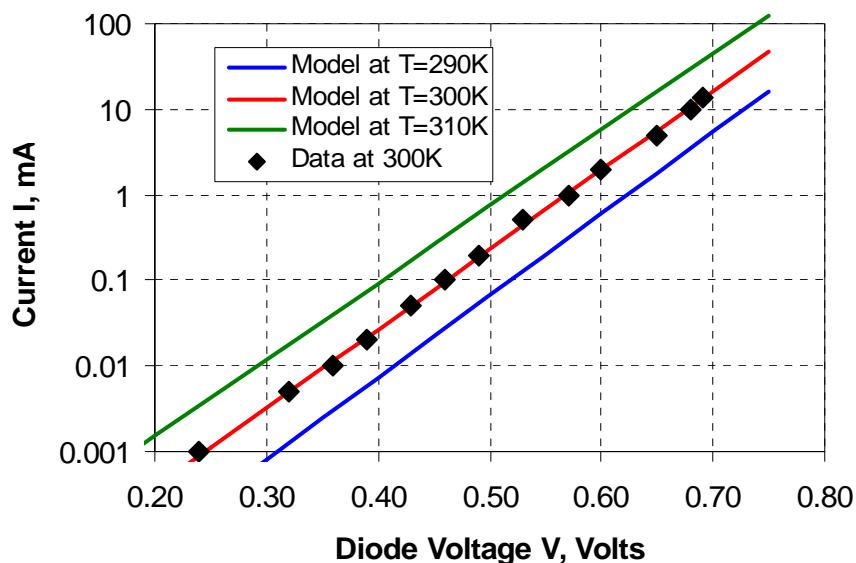


Finally, grab the box and drag to H26 to fill in the other columns. Now, each cell should compute the diode current for the voltage in the particular row, and for the temperature at the top of the particular column that the cell appears in. If you then click in any of these boxes, you can see how the references to the voltage and temperature change as you move from cell to cell.

To plot all three of these curves together we could use the same procedure as before and select the range of data to plot (E11 through H26) and then insert a new chart. However, doing it that way means we'd have to go through all of the formatting stuff again. A somewhat easier alternative is to copy and paste a previously formatted graph (like the one we made for our measured data) and then just change the source data for the plot. To do this, right click somewhere in the new chart and select "Source Data...". Click on the "Series" tab in the dialog box; to add a new data series, then click the "Add" button. For example, after doing this and entering a new name "Model at T=290K" I get the following:



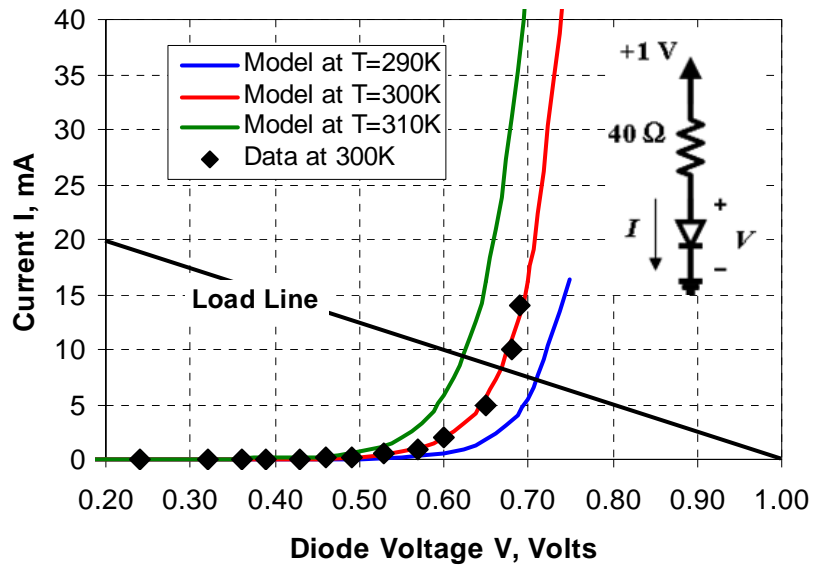
The next step is to define where the X-Y data for this new series is located. The easiest way is to click on the little spreadsheet symbols as shown above, and then use the cursor to select the appropriate cells in the spreadsheet. Note that if you do it this way, you have to first delete the “={1}” entry in the Y-values box, otherwise you will get an error message. Now do this for the other two data sets for T=300K and T=310K. Now all of the data should be on the plot, and we just need to do some formatting to polish it up. Here is what I have on a log scale:



Note that Excel doesn't usually make very intelligent choices of colors and line widths for the plots. For good quality reports and presentations you need to choose bold, dark colors and thicker line widths than the Excel defaults, as I have done here. Also note that the data point

markers (black diamonds) appear *on top* of the red model curve at 300K; to make this happen you need to double click on one of the markers to launch the “Format Data Series” dialog box, and then select the “Series Order” tab to define the order in which the plot lines are drawn.

As a last note, most people do not realize that Excel, like many of the MS Office products, includes similar drawing features that appear in PowerPoint. To access the drawing tools you just select “View → Toolbars → Drawing”. This allows you to add graphics or comments to the plots very easily. For example, here I have reformatted the plot for a linear scale, adjusted the scales a bit, and added a load line by drawing in a straight line through the appropriate intercepts. I also Cut & Pasted the schematic for the load line into the Excel chart from a PowerPoint drawing.



This is a high-quality plot, suitable for lab reports and other publications. The main point here of course is demonstrating how Excel can be used to do the routine data plotting and analysis that we need for ECE 2 and other ECE coursework, but the above also demonstrates some of what we discussed about the merits of using series resistors for diode biasing, *i.e.* that it greatly reduces the variation in bias current over temperature.

There are other nice features of Excel, and the only way to learn them is to experiment with the program. In my experience if there is something you want the program to do, chances are you can find a way to do it if you try hard enough.