

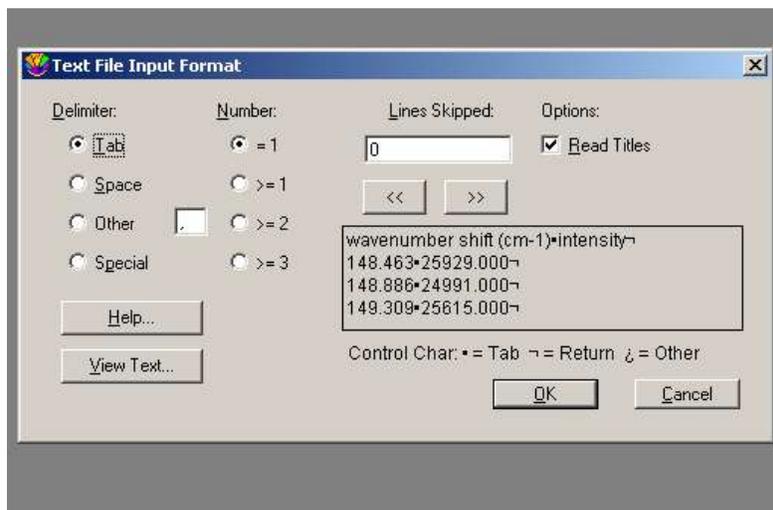
KaleidaGraph tutorial

Kaliedagraph is a useful plotting and data analysis tool. It can produce publication quality graphs and fit data with arbitrary curves. Memory permitting, it can handle data sets of 1000 columns by 1,000,000 rows. The following tutorial is designed to demonstrate most of the common features and is based on Kaleidagraph3.6 on a Windows 2000 machine. Extensive documentation is in the Kaleidagraph manual, it's 300 pages. There is also a built in tutorial from the menu Help>tutorial.

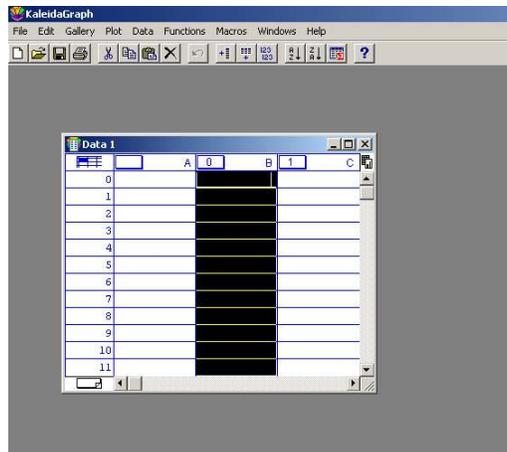
Plotting data

Most data is generated somewhere and then imported into Kaleidiagraph. From the menu File>Open. Under type of file change it from any file, this means any Kaleidagraph file, to all files (*.*) A dialog box appears.

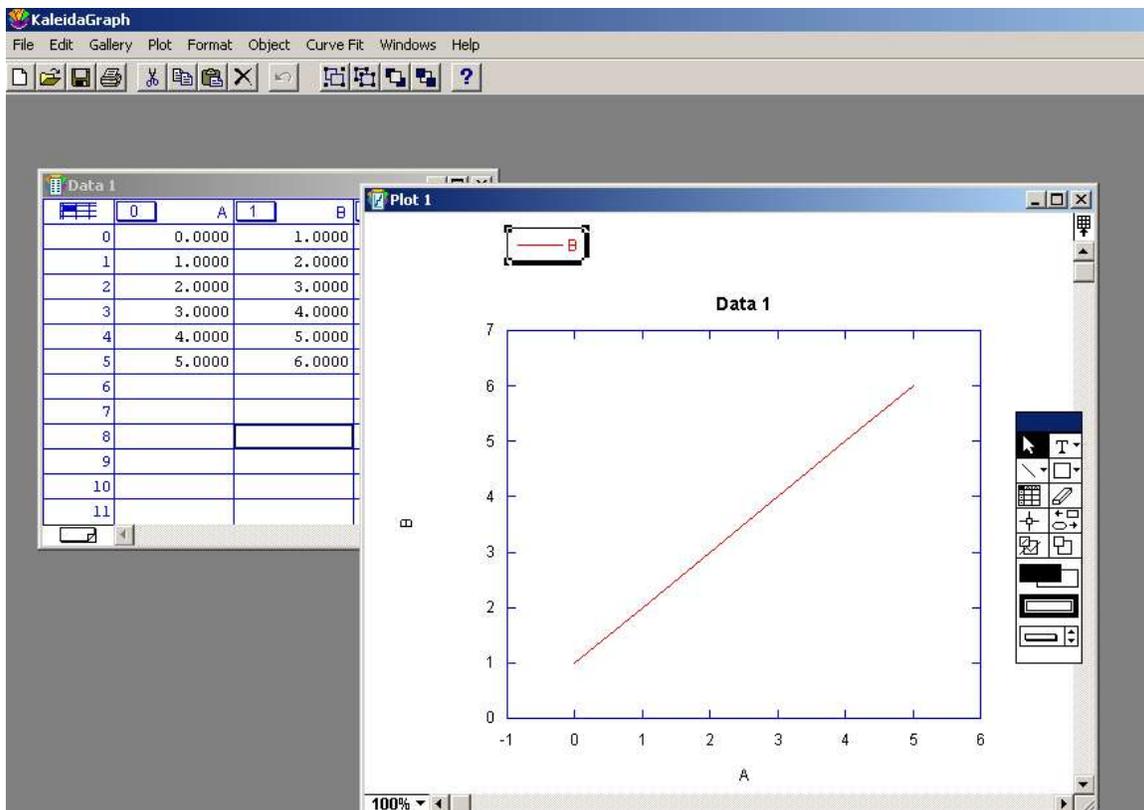
The character which separates one data column from another is called the delimiter, it is often a tab character. If the file has header information that is not data, you can skip these lines by increasing the Lines Skipped value. The read titles check box labels the data columns. When it is all good, click O.K. If you have a Microsoft Excel spreadsheet, it can be opened, use File>Import>Excel instead of File>Open. The Special delimiter is used if you have data rows in the form of text and numbers “Time 1 temp 71”, see manual section 4.3.4 for details.



For this tutorial we will instead manually type in a small data set. Start a new data set File>New. The rows are indexed numerically (0,1,...), and the columns are indexed alphabetically (A,B...Z,AA,AB....). Also notice the columns have a relative index, in the figure below column “B” is column “0” and column “c” is column “1”

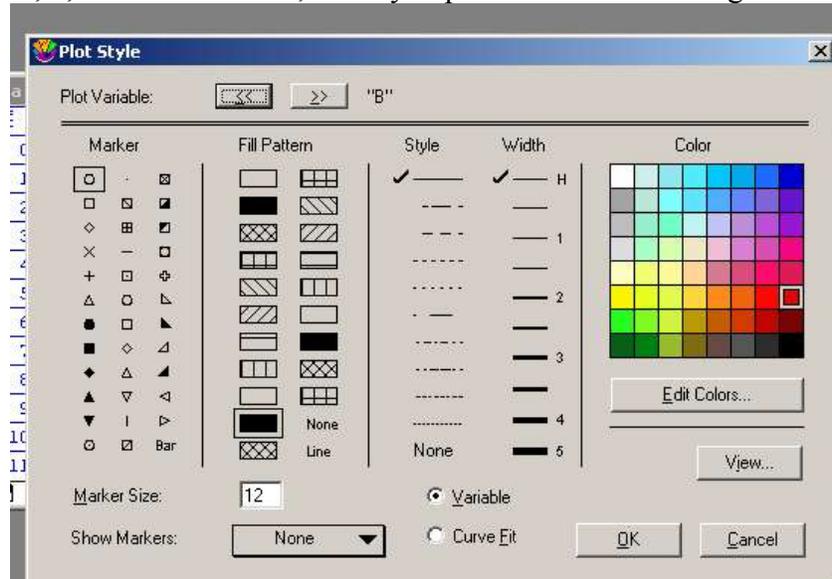


For our simple data set, in column A type the numbers 0,1,2,3,4,5 and in column B type the numbers 6,5,4,3,2,1. We can now plot this data. Go to the menu item Gallery>Linear>line. A new pop-up window will appear, chose column A for X and column B for Y. You will generate a plot similar to the following picture. Notice that if you click on the graph or the data window, the menu items change.



Click on the plot, then go to menu item Plot>Plot Style. Here you can control many aspects of the plot. Choosing a different color is straightforward. The item at the bottom says show markers, you can chose “none, all, every Nth”. If you chose to show markers,

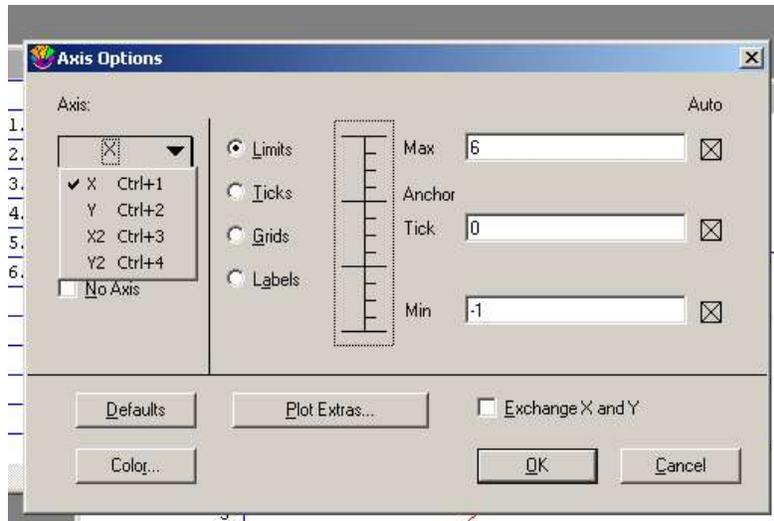
then you can choose one of many different markers. “Style” and “width” refer to the line, choosing none will have markers only. Although you might not notice the difference for H (hairline) .5, 1, 1.5 on the screen, when you print there will be a big difference.



With the plot is a new tool palletted, shown in the right of the figure. The arrow selects different elements on the graph. The “T” is for text, you can place a box of text anywhere on the plot. The diagonal line is to create lines, click on it, and then hold the right mouse button down for about one second. Another popup appears and you can change the line to an arrow. The box is to draw squares or rectangles, if you hold down the right button another popup appears and you can select circles or polygons. The next one is to create a table of text. The eraser tool is to delete labels or objects. The crossbar and square is the identify tool. It displays the X and Y values when clicked inside the plot area. If you press the “Alt” key on the keyboard a text label is produced when you release the mouse button. The box and arrow is the align tool and it is used to align objects. The two boxes with lines is the data selection tool. This is used to define a region and “mask data”. Only data inside the defined region is used for plotting and curve fitting. If you press the “Alt” key this has the opposite effect, data inside are excluded from the plot. (To unmask select the data columns and goto Functions>Unmask.) The two boxes is the zoom tool, this zooms in on the selected region without masking data. To return to the original size, double click this.

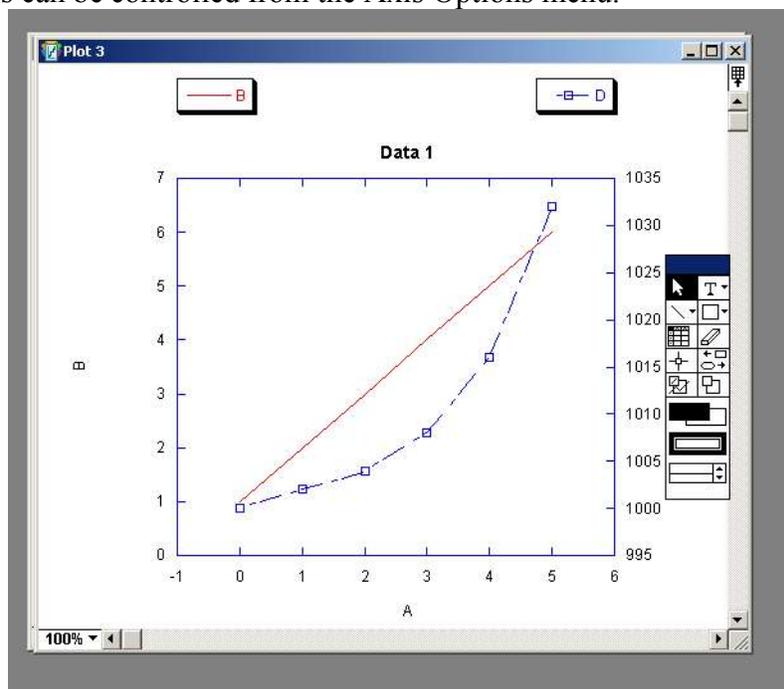
The axis color, range of the plot can be selected from Plot>Axis options. To access the different axis, click on the down arrow at the left of the “X” to bring up “Y,X2,Y2”. You can change the limits, grid, tick marks independently for each axis. Under color you can change the axis colors and background color.

If you double click on the “A” labeling the X axis, you can change the text. Under style you can add subscripts or superscripts. Greek letters are available from style and alt. Font, this allows both Greek and roman letters in the same text line.



You can plot multiple data on the same graph. Click on the data window, and for column C type in “6,5,4,3,2,1”. Make a new plot, Gallery>linear>line. Chose column A for x and columns B and C for Y. The second set of data can be independantly controlled from the Plot Style menu. Click on the double arrow “>>” to move the display from “B” to “C”. Again the color, style, and markers can be chosen for the second column of data.

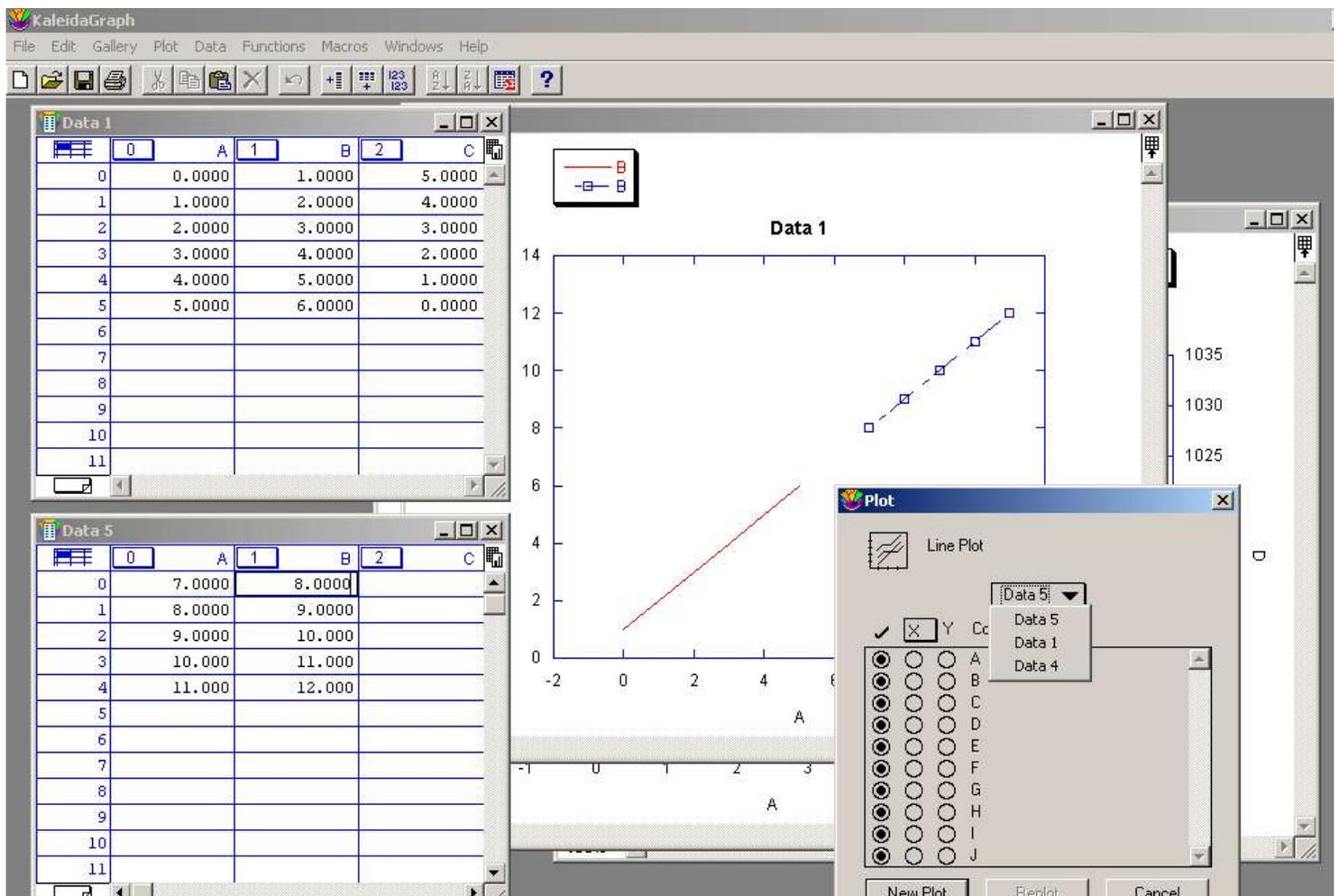
Some times you want to plot two data sets which share the same X axis, but have very different Y scales. In the data window in Column D type the numbers “1000,1002,1004,1008,1016,1032”, use the scroll bar at the bottom of the data window to move to column D. If you try to plot both of these on the same Y scale you will only see two horizontal lines. A better way is to use a double Y plot. Gallery>Linear>Double Y, chose column A for X axis, column B for Y1 axis and column D for Y2 axis. The range of each Y axis can be controlled from the Axis Options menu.



If you change the data and want to update the corresponding plot, goto Plot>Update plot.

If your data has a break (1,2,3,4,12,13,14,15) and you would like to show this break, Plot>Axis Options>Plot Extras and click on data break on the left.

Sometimes you have data in two files and would like to plot them on the same axis. Make a new data set File>new. In column A type the numbers “7,8,9,10,11” and in column B type “8,9,10,11,12”. Start a new plot Gallery>Linear>line, chose columnA for X and column B for Y. Before you hit “Plot” click on the down arrow for the data window, similar to choosing other axis from the Axis Options menu. Chose the second data window and chose column A for X and column B for Y. Now hit plot. The X axis will be the range in the combined A columns and the Y axis will be from each column B.



Layouts

Often you want to have one figure with several plots, maybe an inset. One approach is to copy one plot and paste it into the second. Unfortunately what is pasted is a static image, if you change the plot you have to copy and paste again. The layout feature allows you to arrange plots of relative size and position. Note, when you save a layout you are saving the relative size and position information and not the plots themselves.

From the Windows>Show layout>KGl layout a window will appear. From menu Layout>Arrange layout. Select number of rows and columns for plots, 2x1 for an example. You can then select a position and from the Layout>Select plot put a plot in that position. Each position is independently scalable, so you can resize one and move it to an appropriate position for an inset. Then print the layout for the final figure.

Data manipulation

Often you want to do some numeric manipulation of your data, for example normalize all values relative to the largest. This can be done with the formula entry tool. Another common task is to create a large series of numbers, use Function>create series.

Open a new data window File>New. Highlight column A by clicking on it once (if you double click you can change the name of the column). Go to Functions>Create Series. For initial value enter -5, increment .1, multiplier 1, final value 5 and click the box next to final value. Click O.K. You have just created 100 numbers. Let's calculate the square of each of these numbers. First click on column A, the relative column number is "0". Next go to Windows>Formula Entry. A window appears with buttons F1...F8, these are different functions you can enter. The syntax is a little strange and uses the relative column numbers. To calculate the square type " $c1=c0*c0$ " and then click "run". This says "for each row, put into relative column 1 the value of column 0 multiplied by column 0". To put a normalize the column of squares, first click on Column B to make it relative column 0, then type in the formula entry " $c1=c0/cmax(c0)$ " and hit run. This says "for each row in relative column 1 place the value of the row divided by the maximum value in column 0". The same procedure to fill column C can be achieved by first clicking on column A and typing in the formula " $c2=c1/cmax(c1)$ ".

You can also perform conditional calculations. Say we want to calculate the square root of the column of numbers in column A. Kaleidagraph isn't smart enough to know the square root of $-1=i$ so it will give an error for the negative numbers. We want an expression similar to "if the number is less than zero then multiply by -1 and take the square root, otherwise just take the square root." A pseudocode implementation would be "if ($x<0$) $\sqrt{x*(-1)}$ else \sqrt{x} ". The Kaleidagraph syntax for the formula entry is " $c3=(c1<0) ? \sqrt{-c1} : \sqrt{c1}$ ". Manual Chapter 6.9 for more advanced details.

Assume you have a two column data file where the data range is $0, X1, \dots, Xn, 0, X1, \dots, Xn$, that is the same measurement has been performed twice. If you want to average the two readings first click on the data column to make it relative column 0. In the formula entry

enter “c1=(cell(index(),0) + cell(index()+npts(c0)/2,0) /2)” This produces an average value in column 1, there are half as many data points as in column 0.

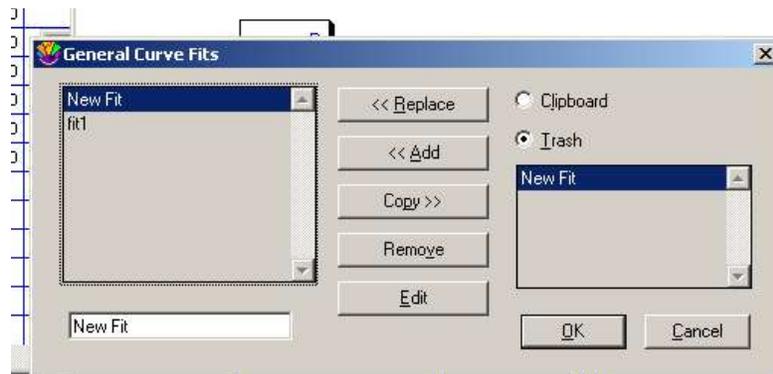
Curve fitting

Kaleidagraph has a powerful routine for fitting a set of data to an arbitrary curve. This is an implementation of the Levenburg-Marquardt algorithm for nonlinear curves.

Open a new data window File>New. Create a series from -5 to 5 with .1 steps, just like in the above section Data manipulation. Enter the following formula in the formula entry window “c1 = 1+5*sin(8*c0)+2*exp(-c0*c0/3)”, make sure the box next to “rad” is checked (top left of formula entry) so the angles are expect to be in radians. Run the formula, and plot the result. You should see a Gaussian envelope with many wiggles.

Select the plot and chose the menu item Curve Fit>Linear. Click on the box for column B and then click O.K. A line through the data should appear along with an equation “y=1.6079+0.008531x r=0.031251”. If the equation is not displayed on the graph, go to Plot>Display equation. “R” is a statistical measure where 0 is no fit and 1 is perfect. If you then go back to Curve Fit>Linear there is a new box titled “View”. Here you can have the fitted values appended to to a data column at the end of your data. Obviously a linear fit to this data is very poor.

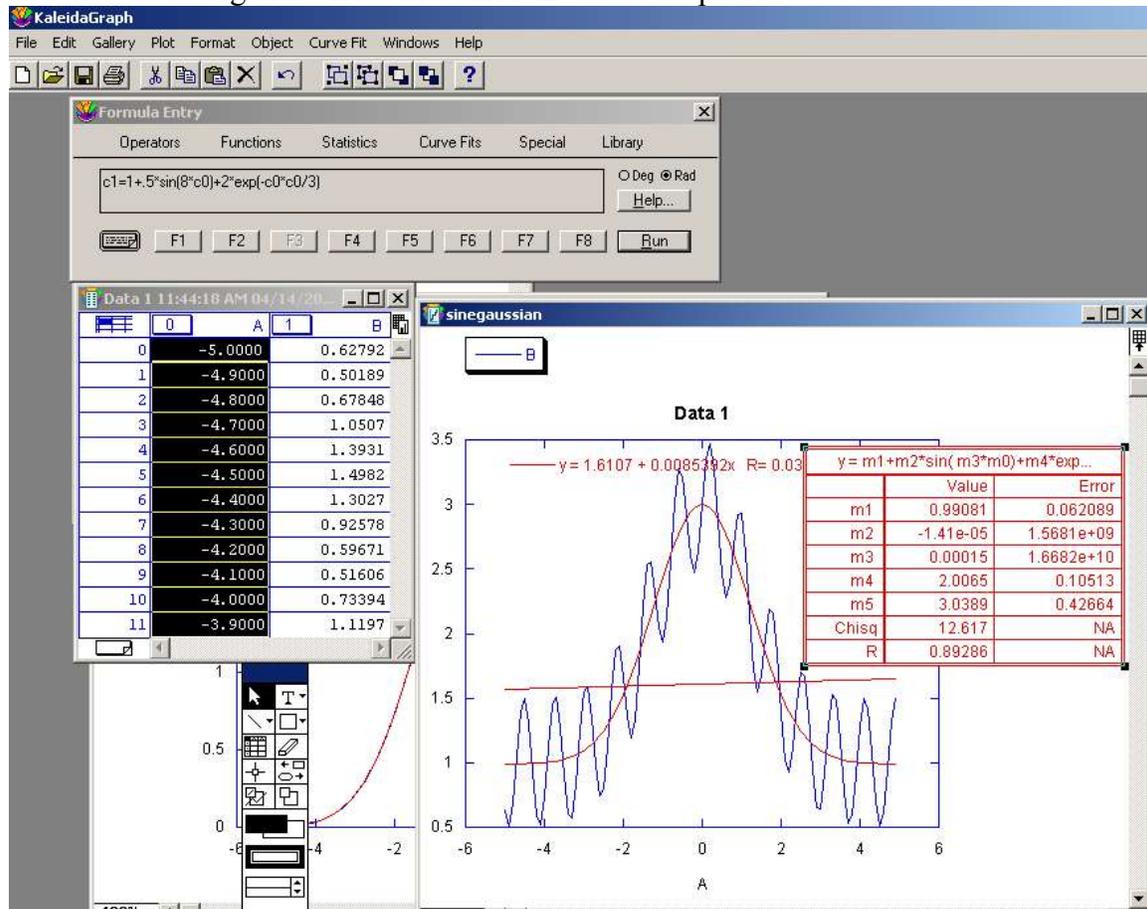
We can enter a different equation to try and fit the data. Go to Curve Fit>General>Edit General, a new window appears. Click on New fit in the left box and then add. In the right box the “new fit” will appear. On the right box click on New fit.



In the lower left box, change the name of “New fit” to something like “sine gaussian” Then click edit. Our curve is of the form “y=A+B*sin(C*x)+D*exp(-x^2/E)”. There two tricky things about making general curve fit work properly. The first is that in the equation the independent variable is called M0 and all parameters are M1..M5. Under the edit window, type m1+m2*sin(m3*m0)+m4*exp(-m0*m0/m5). The second trick is to remember to give an initial guess for each parameter. In our example we will try 1 for each. The full text in the equation box should be “m1+m2*sin(m3*m0)+m4*exp (-m0*m0/m5);m1=1;m2=1;m3=1;m4=1;m5=1”. The algorithm looks for new values of

the parameters by looking at the derivative of the function. If you click specify partials you can enter explicitly the partial derivative for each DF/DM_i , otherwise it will do it numerically. For our example $df/dm_1=1$, $df/dm_2=\sin(m_3*m_0)$, $df/dm_3=m_2*\cos(m_3*m_0)*m_0$, $df/dm_4=\exp(-m_0*m_0/m_5)$ and $df/dm_5=m_4*m_0*m_0*\exp(-m_0*m_0/m_5)/m_5/m_5$.

Now go to Curve Fit>General>Sine Gaussian, select column B and O.K. The fit will be approximately correct. The error is an indication of how well the parameter is fit, the relative error of (error/value) should be small. Notice the large relative error for M2, M3 and that M2 is negative. Let's limit the M2 value to a positive value.



Go to Curve Fit>General>Edit General and select sine gaussian, you can also get to the same window from Curve Fit>General>Sine gaussian>define. We will limit values for m2 to be between 0 and 10. Change the text for m2 to “(m2=limit(0,10,m2))” so the formula should read:

”m1+(m2=limit(0,10,m2))*sin(m3*m0)+m4*exp(-m0*m0/m5);m1=1; m2=1; m3=1; m4=1; m5=1”

For the limit to work, you must specify the partial derivatives and click the check box. Now when you try to fit with this curve you will receive an error message ”Singular coefficient matrix” This means that the fitting routine got stuck at a local minimum somewhere in the parameter space and couldn't fit the curve. This usually means that the

initial guesses were off. Try again, but change the initial m_3 to 5. Another singular matrix error. It turns out that the range which this will converge is for a starting value for m_3 between 7.45 and 8.55 with all other initial guesses at 1.

Kaleidagraph plots in other places

Sometimes you might want to use a KaleidaGraph plot somewhere else. If you want to post a graph on a web site you need an image. From File>Export select JPG, PNG, or TIFF. JPG is a lossy compression, these files are small but the graph axis and text tend to come out blurry. PNG is a lossless compression, the files sizes a a little larger but everything is clear. TIFF is an uncompressed image and is the largest file size.

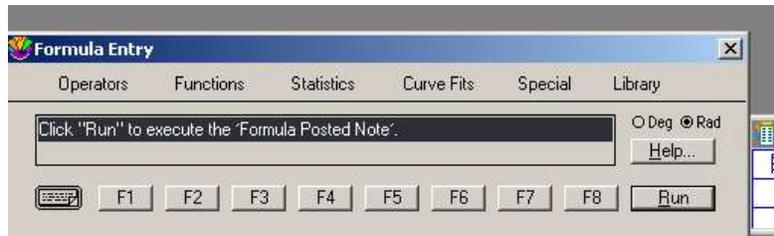
Encapsulated PostScript is possible using the printer driver. Many journals accept EPS files and LaTeX can directly incorporate these. The first step is to install a PostScript printer driver. In Windows, Start>Printers>Add printer. Chose a HP or Lexmark printer model with PS in the name. Set it to print to file. Once this is complete, in KaleidaGraph, or any application, print the graph to the PostScript printer. A popup window will appear where you can select a printer, select the postscript one. In the tab titled "Layout" click on the advanced button in the lower right. Under Document options>Postscript options> Postscript Output option, change this to "EPS". In Kaleidagraph printing Layouts with the position near the top of the page seems to work better than printing graphs directly. The postscript file can then be viewed with a postscript viewer, ghostscript on the unix machines. IF there is too much white space, you can edit the postscript file with a text editor. One of the first few lines is "##BoundingBox LX LY UX UY" where LX and LY are the coordinates of the lower left corner and UX and UY are the coordinates of the upper right corner of the data. Change these and save the file.

You can embed a Kaleidagraph plot into another application, Microsoft's Word or Powerpoint for example. The best way to do this is to select the plot in Kaleidagraph and from the menu Edit>Copy Graph. A popup window will appear. Make sure that the box next to "copy OLE data" is checked. In the other application go to Edit>Paste Special. This places a linked object in the other application. When you click on the graph in the other application, Kaleidagraph starts and you can make changes to the graph. From KaleidaGraph select File>Update the graph in the other application is changed. If you select File>Update and Return the graph is changed and Kaleidagraph is closed if it wan't already open.

Automation

Once you have a running experiment, you often do the same set of manipulations to the data. As an example, assume you have data from a Raman scattering experiment and you always convert the wavelength to wavenumber, normalize the data, and plot the result. You can write a script that performs all of this with a single click. The place to enter the script is in the formula entry window. In the lower left corner next to the F1 button is a

picture that is supposed to look like the top page of a post-it note. Click on it, and you enter a simple text editor window. Enter the commands there. You can save and load other scripts. To Run the script first exit the text editing window then click run.



The following script assumes that the data file is in two columns and has no header information. The script allows you to choose a file to open, then calculates the wavenumber shift from the first column, places it in the third column and titles the third column “wavenumber.” It then takes the second column, normalizes it and puts the result in the fourth column. It then plots data using column 3 for X and column 4 for Y, the column indexing starts with 0.

```
#DATAFILE
file = "*.*)"
read_titles = false
#END
```

```
#FORMULA
c2= 1e7*(1/488-1/c0);
name("wavenumber",c2);
c3=c1/cmax(c1);
name("normalized data",c3);
#END
```

```
#SCRIPT
plot_type = line
begin_group
x = 2
y = 3
end_group
#END
```

More information about scripts is in Chapter 6.2.10.

Another common task is to create many graphs that have the same form. The limits are the same, the marker is the same, and you've added a stick drawing in each that you'd like to include in all of them. Window>Plot script. In the template hold the right mouse button down and then choose which the graph which you'd like the others to be like. Then click add, choose the data window and X and Y columns. An entry appears named “Group

1” When finished adding groups, click on run. If you have a data file with one x and many y's that you'd like, after the third script and new option will appear, “repeat”. Clicking on this generates a group for all the remaining data columns.