MacCurveFit version 1.0

September 1993

Introduction.

MacCurveFit is a program written for scientists, science students or anyone who wants to fit user defined functions to a set of data points. I originally wrote it because I wanted to fit some rate equations to my chemical kinetics data. The popular commercial graphics packages allow you to fit linear, polynomial and logarithmic curves to data, but most packages don't let you specify your own equation. However MacCurveFit allows you to fit almost any equation you choose. A trivial but annoying example is fitting a straight line through the origin. Many popular programs fit the equation y = ax + b, however they decide what value to use for b and it isn't necessarily zero. MacCurveFit lets you simply fit the equation y = ax. The program also lets you directly manipulate the coefficients or constrain them to chosen values. MacCurveFit is very flexible in this regard and gives you control over which coefficients will be optimized and even which mathematical algorithm will be used.

System Requirements for MacCurveFit 1.0

Curve Fit requires at least the 128K version of ROM which means it will run on the 512K Enhanced or higher machines. It also requires version 6.0 or higher of the System software. The program runs best under system 7.

An Introduction to the Program.

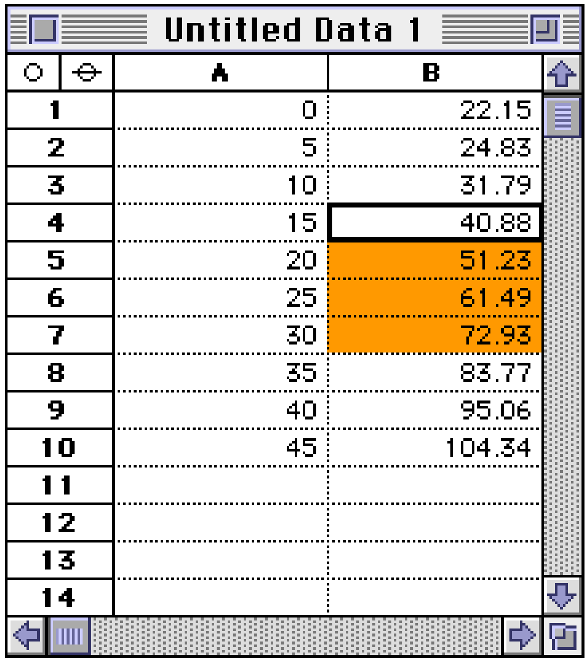
I'll introduce the program bit by bit. The program has three types of display windows: **Data** Windows, **Plot** Windows and **Curve Fitting** Windows. First, I'll describe the Data Window.

The Data Window.

A new untitled data window may be created by selecting **New** from the **File** Menu. The data windows are similar to the windows of spreadsheet programs. The windows have a number of data cells organized into columns and rows. MacCurveFit 1.0 allows you to enter up to 1000 columns and 32767 rows. In principle the real limit is imposed by your computer's available memory as a completely full data window would consume 312 Mb!

Note: If you intend to enter large amounts of data into the data windows then you should increase the **Preferred Size** in the **Get Info** window in the Finder. (Select the MacCurveFit application icon in the Finder and select **Get Info…** from the **File** menu.) The application is distributed with the preferred size set to 512 K, however if you may need to increase this 1 Mb or more. In general you can try running with the smallest possible memory size, the application will notify you if it is running low on memory.

The active cell is indicated by a thick black border and a blinking insertion point. Numbers may be typed into the cell from the keyboard and locked in by pressing the **enter** key. A new cell can be made the active cell by clicking in it. When a new cell is clicked the contents of the previous cell are automatically locked in. Adjacent cells can be selected by using the **arrow** keys (and also with the **tab**, **return**, **shift-tab** and **shift-return** keys).



Untitled Data 1

	A	B
1	0	22.15
2	5	24.83
3	10	31.79
4	15	40.88
5	20	51.23
6	25	61.49
7	30	72.93
8	35	83.77
9	40	95.06
10	45	104.34
11		
12		
13		
14		

A range of cells can be selected by clicking the mouse in one cell and dragging to the final cell. Alternatively a range can be selected by clicking in one cell and then holding down the shift key while clicking in the end cell. A selected range is indicated by the system's highlight colour (selected in the colour control panel). On black and white machines the selected range is inverted.

Editing the Data Window.

Data can be cut, copied and pasted in the usual Macintosh way by choosing the appropriate items from the **Edit** menu. A sequence of digits from the active cell may be copied by selecting just those digits, or an entire range of cells may be copied by selecting the range and then choosing **Copy** or **Cut**. Similarly a range of cells may be pasted into the data window. If the clipboard holds a range of cells and a single cell is selected in the data window, then choosing **Paste** automatically expands the selected range to the right and below before pasting in the data. If however a range of cells is selected before choosing **Paste** then *every cell in the selection is affected and only those cells are affected*. If the selection is larger than the number of cells in the clipboard then the extra cells in the data window are cleared. If the selection is smaller than the cells in the clipboard then not all the data on the clipboard will be pasted. Too be sure to paste all the cells in the clipboard it is best to select a single data cell before choosing **Paste** from the **Edit** menu. These operations are all reversible and can easily be undone by choosing **Undo** from the **Edit** menu.

Here are a couple of tips to remember when copying and pasting. Choosing **Select All** from the **Edit** menu will select the entire data window. Holding down the **option** key while choosing **Select All** will select all the digits in the active cell only. Clicking in a column label will select all cells in that column. Similarly clicking in a row label will select all cells in that row. You can also drag or shift click the labels to select a number of rows or columns. Sometimes you may find that not all of the digits of a number will fit in the data cell. If you hold down the option key while dragging across the digits in the active cell, when you move the mouse outside the cell the hidden digits will scroll into view. Holding down the option key and pressing the left or right arrow keys will also cause the digits to scroll when the insertion point reaches the edge of the cell. This allows you to inspect, delete or copy the obscured data.

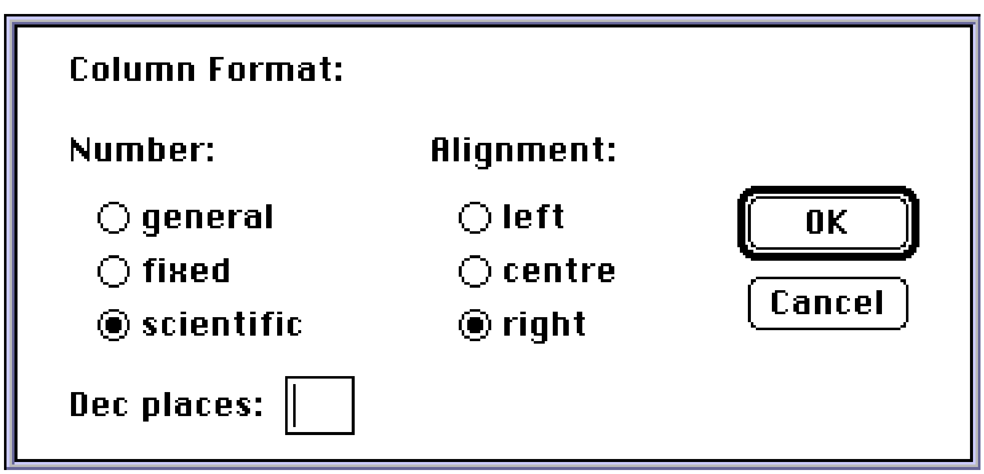
Data may be deleted from the window by using the **delete** key or by choosing **Clear** from the **Edit** menu. If the selection is a single cell then both methods will clear the cell. If the selection is a range of cells you can clear the entire range by choosing **Clear** from the **Edit** menu. To clear only the active cell use the **delete** key.

Importation of Data from a Tab delimited Text File.

Although MacCurveFit 1.0 does not have a direct method of reading text files and extracting data, this can be done via the clipboard. To do this first open the text file with a text editor (such as TeachText, Edit) or a word processor (MS Word, for example) and select the data. Copy this data to the clipboard and then it will be available to MacCurveFit. Open a new data window and select **Paste** from the **Edit** menu; you can now use the data.

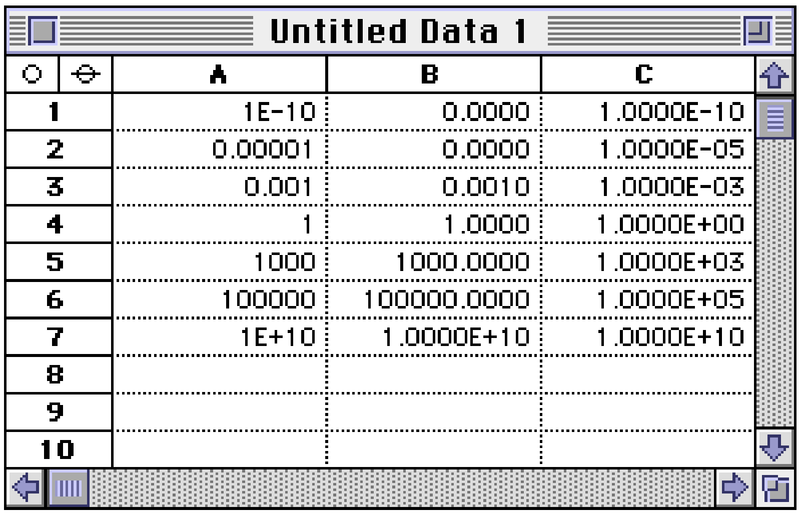
Changing the Format of the Data.

The format of the data in any column or range of columns can be changed by selecting the columns and choosing **Column Format…** from the **Data** menu. The following dialog box will appear.



The **Number** section selects the desired numerical format. **General** format will display numbers with the minimum number of decimal places required to specify the number. The number will be displayed in fixed format preferentially but will default to scientific format for large or small numbers. When **fixed** format is selected an additional dialog item appears which allows you to specify the desired number of decimal places. **Fixed** format will display numbers without exponents to the desired precision. Large numbers will be displayed with exponents. **Scientific** format will always display numbers with exponents to the specified precision entered in **Dec places** field. The figure on the next page shows the format types for three identical data sets. Column A is displayed in general format, column B is displayed in fixed format (4 decimal places) and column C is displayed in scientific format (4 decimal places). The **Alignment** section allows you to select whether the number is left, centre or right justified within the data cell.

The **Edit** menu contains an item titled **Underlying Value**. This is useful in cases where you may have typed in data to 5 decimal places and then formatted the cell to show only 1 decimal place. If you wanted to change or inspect the hidden digits you could choose **Underlying Value** from the **Edit** menu. This temporarily displays the contents of the active cell in general format thus allowing you to see all of the digits. I should point out that this command displays what is locked in memory and not necessarily what is currently displayed in the cell. Hence if you selected a cell already containing a data value and then typed new data over it, you can recover the original value with this command. Note that you can also achieve the same effect by choosing **Undo Typing** from the **Edit** menu. If the replacement data was locked in, you can revert the cell to the original value by choosing **Undo Data Entry** from the **Edit** menu.


Untitled Data 1

	A	B	C
1	1E-10	0.0000	1.0000E-10
2	0.00001	0.0000	1.0000E-05
3	0.001	0.0010	1.0000E-03
4	1	1.0000	1.0000E+00
5	1000	1000.0000	1.0000E+03
6	100000	100000.0000	1.0000E+05
7	1E+10	1.0000E+10	1.0000E+10
8			
9			
10			

By now you will wondering what the two curious icons are in the top left corner of the data window. There may be occasions when you'll want to hide some data points from a plotted graph. You can do this by selecting the data in the data window and then clicking the left icon. The circle icon will change to a dot indicating that the points will be excluded from the graph. The two icons always reflect the status of the current active cell and will change when another cell is selected. The data that is excluded from the plot can be identified at a glance by its different font style. The invisible data is displayed in outline style. When the right hand icon is clicked, the selected data will be disregarded when performing curve-fitting. The icon will change to a circle with no intersecting line and the affected cells will show their data in italic style.

Printing the Data Window.

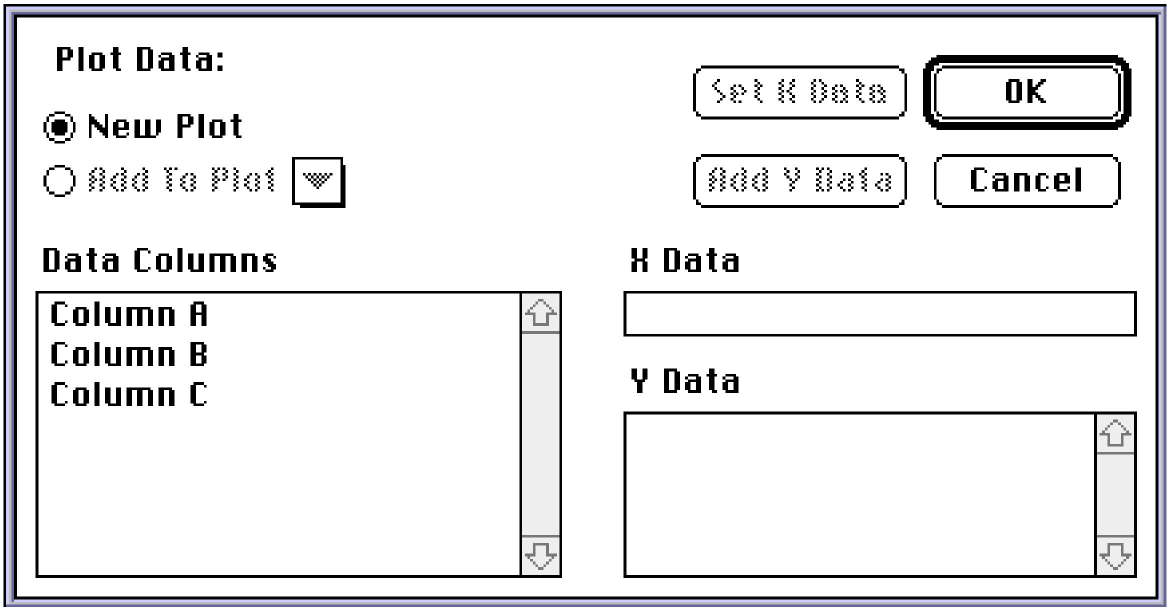
The data window may be printed in the standard Macintosh way. Make sure a printer has been selected in the **Chooser** and select **Page Setup…** from the **File** menu. You can specify any of the standard printing options such as landscape printing or reduction etc. Selecting **Print…** from the **File** menu then puts up the standard printing dialog box and starts the printing process.

Saving and Opening Data Files.

The contents of the data window can be written to a disk file by choosing **Save** or **Save As…** from the **File** menu. The standard Macintosh Save dialog box will appear and you can specify a folder and a file name for the new file. The file can be opened later by choosing **Open…** from the **File** menu. The standard Macintosh open dialog box will appear allowing you to select the file.

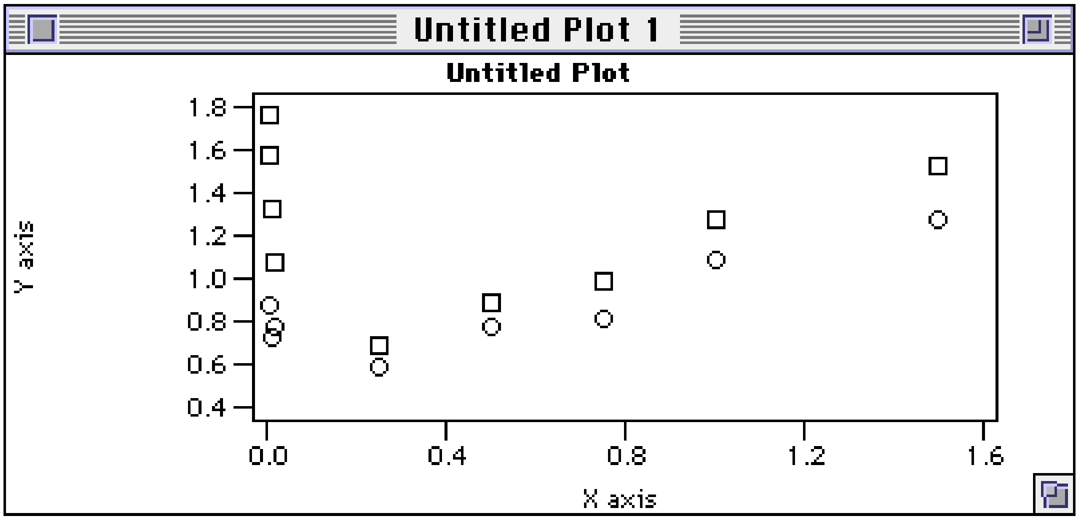
The Plot Window.

When data has been entered into the data window it can be plotted by bringing that window to the front and then choosing **Plot Data…** from the **Data** menu. The following dialog box will appear.



If the **New Plot** radio button is selected then a new plot window will be created. If there are existing plot windows open then the **Add To Plot** radio button will be enabled and the popup menu next to it will contain a list of the plot windows. If you want to add the new data sets to an existing plot window then select the lower radio button and choose the plot window from the popup menu.

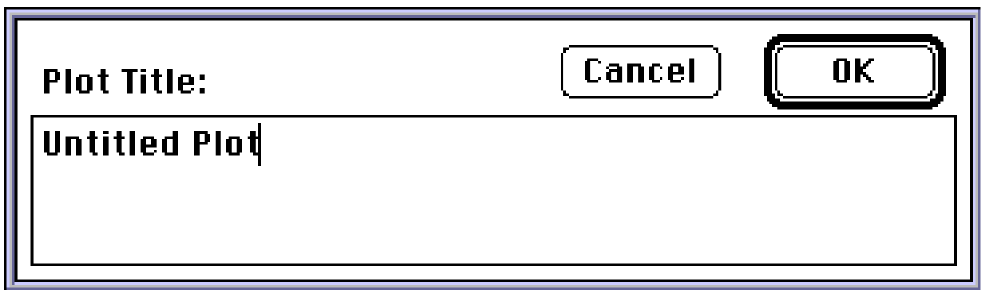
The left list shows the data columns available in the data window. To specify which column should be used as the x data set you select it from the list by clicking it and then click the **Set X Data** button. You can change your selection at any time by selecting another data column from the left list and clicking the **Set X Data** button again. You specify which data columns will be plotted on the Y axis by selecting the data sets in the left list and clicking the **Add Y Data** button. The lists are standard Macintosh lists and behave as such. Multiple data columns can be selected by holding down the shift key and dragging. Data columns can be removed from the right list by selecting the columns and clicking the **Remove Y Data** button. When you have made your selections click the **OK** button and the plot will be constructed. A typical plot is shown in the figure below.



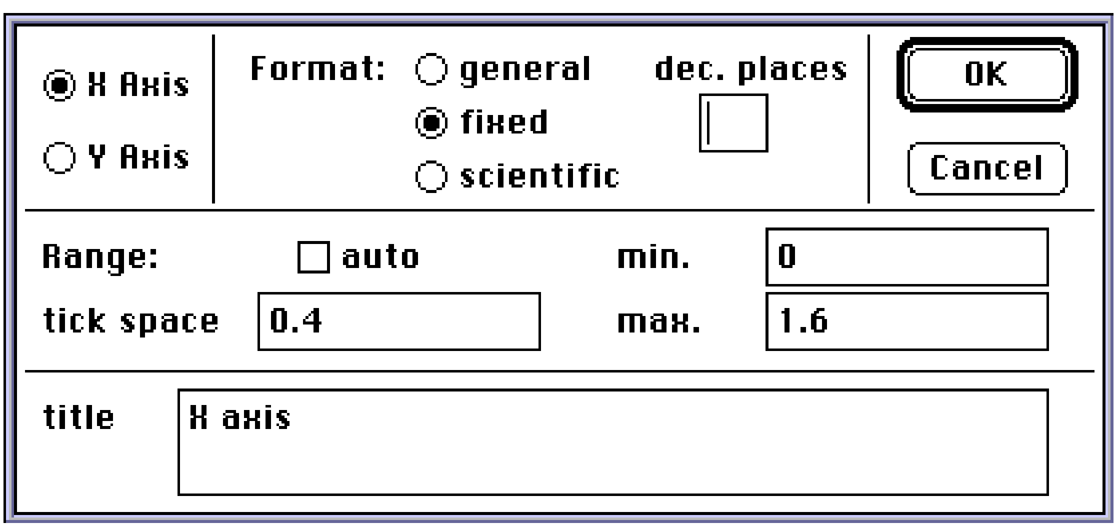
Untitled Plot 1
Untitled Plot
Y axis
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.0
0.4
0.8
1.2
1.6
X axis

Changing the Appearance of the Plot Window.

The appearance of the plot can be altered in a number of ways. The height and width of the plot can be altered by dragging the small icon at the bottom right corner of the window in the same way that any other Macintosh window can be resized. The title of the plot can be edited by choosing **Title…** from the **Plot** menu. The editing is performed in the dialog box that subsequently appears.

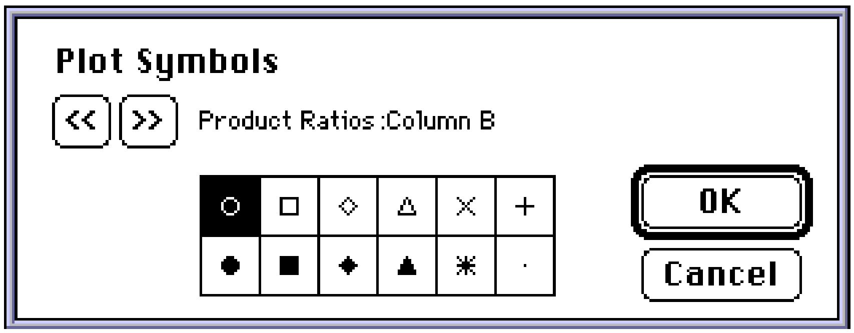


The plot axes can be tailored by choosing **Axes…** from the **Plot** menu. The following dialog box will appear.



The radio buttons in the top left corner allow you to specify whether you are changing the X axis or the Y axis. You can set up the X axis, then click the **Y Axis** button and configure the Y axis and then click on the **OK** button. The centre panel of the top section allows you to specify the numerical format of the axis labels. The panel works just like the **Column Format…** dialog box discussed above in the **Data Window** section. The panel below this lets you specify the axis range. The **min.** and **max.** fields let you set the range itself and the **tick space** field lets you specify the distance between the axis labels. If the **Auto** box is checked then the program will automatically set reasonable values for the range. Finally the title of the axis can be typed into the bottom section of the dialog box. After clicking **OK** the plot will be updated to reflect the changes.

The symbols used to indicate the data points can also be changed. This is done by choosing **Data Symbols…** from the **Plot** menu. A dialog box will appear and will show the available symbols. The arrowed buttons let you choose which data set you want to alter. The name of the data set is composed of the data window name and the column name separated by a colon. Since the plotted data can come from different data windows it is necessary to include the data window's name. Having specified the data set simply click the desired plot symbol and then click **OK**. You may of course change a number of data symbols before dismissing the dialog box.



Updating the Plot Window.

When the **Data Link** item in the **Plot** menu is checked (the default option) the plot window is automatically updated when the plot data is edited in the data windows. This can sometimes cause delays when you want to edit a large number of data cells. The automatic updating can be turn off by unchecking the **Data Link** option in the **Plot** menu. When this option is turned off you can explicitly update the plot window by choosing the **Update Plot** item from the **Plot** menu. The **Data Link** option also governs the automatic updating of the plot when curve fitting is being performed (see later).

Printing the Plot Window.

MacCurveFit's printing methods have been designed to give the highest possible quality on any printer. When the plot is drawn on the screen it is drawn at a resolution of 72 dots per inch (dpi). The Macintosh printing architecture is set up to use this as the default resolution when printing. Although this works well when printing most things it gives poor results when printing curves, they often look jagged and unattractive. Most printers are capable of printing at higher resolution, the ImageWriter manages 144 dpi, the LaserWriter 300 dpi and the StyleWriter 360 dpi. MacCurveFit interrogates the chosen printer just prior to printing and finds out what resolutions the printer supports. It then re-images the plot to take advantage of the printer's best resolution, this way you will always get smooth curves even on non-postscript devices.

To print just select the **Page Setup…** and **Print…** items from the **File** menu, the program will set up the printer for optimum results automatically.

Copying the Plot to the Clipboard.

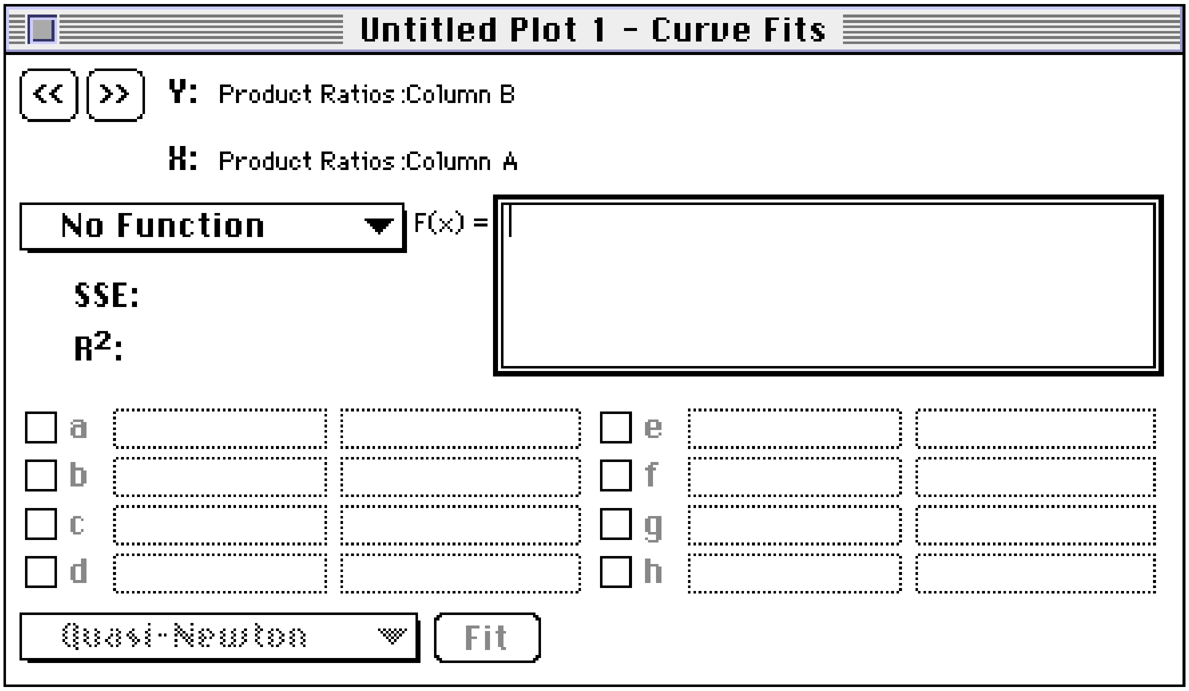
The plot can be copied to the clipboard by choosing **Paste** from the **Edit** menu. The graphic information sent to the clipboard is imaged at 72 dpi so that other applications can deal with it in the standard way. However in an effort to achieve good results when printed by the recipient program, the graphic information contains some postscript code which will produce high quality results on postscript devices such as the LaserWriter.

Saving and Opening Plot Files.

The plot window can be saved by choosing **Save** or **Save As…** from the **File** menu. The resulting plot file contains information about the plot, the curve fits (see the Curve Fit Window section) and information on how to locate the dependent data files and function files. When you want to continue with work you saved earlier, open the plot file first and all of the windows you were working with will open automatically. To avoid clutter the curve fit window doesn't automatically open but it is reconstructed so that you can simply open it later by choosing **Curve Fit…** from the **Fit** menu. The Macintosh acquired many new features when Apple released System 7. If you are using System 7 then MacCurveFit will be able to find the dependent data and function files even if they have been moved to another folder or renamed.

The Curve Fitting Window.

The **Curve Fit** window is the heart of MacCurveFit. It allows any function to be plotted in the plot window and, of course, allows the function to be fitted to the data in the plot. The curve fitting window will open when you choose **Curve Fit…** from the **Fit** menu. The curve fit window is pictured below.



The arrowed buttons in the top left corner allow you to select the current data set. The name of the data set is composed of the data window name and the column name. The popup menu below the arrowed buttons allows you to choose a function to be fitted. This menu is divided into three sections. The top section shows the items **No Function**, **Linear**, **Linear (0,0)** and **Polynomial**. **No Function** is the default menu choice and indicates that you do not wish to fit a function to the data. Choosing **Linear** will load the function “a*x + b” into the function text box to the right of the function popup menu. If you wish to fit a straight line through the origin the next item **Linear (0,0)** will place “a*x” into function text box. Choosing **Polynomial** will bring up a submenu allowing you to choose the desired polynomial order. The polynomial function will then be loaded into the function text box.

The bottom section of the popup functions menu contains commands for maintaining function files. A new function can be created by choosing **New Func** from the menu. You can also create a new function by simply typing directly the function text box when either no function is selected or a built-in function such as a polynomial is selected. After creating a function they can be saved to a file by choosing either **Save Func** or **Save Func As…** from the function popup menu. A previously saved function can be opened by choosing **Open Func…** from the menu. A standard Macintosh dialog box will open allowing you to select the file. The function's text will be loaded into the function text box and the name of the function will be placed in the central section of the function popup menu. A function file can be closed when no longer needed by selecting **Close Func** from the popup menu. This will remove the function's name from the popup menu and set the current function as **No Function**.

Organizing Windows.

Because large numbers of windows may be open in MacCurveFit, I have included a hierarchical **Windows** menu. You can bring any window to the front by selecting its name from the **Windows** menu. The menu has submenus for the **Data**, **Plot** and **Fits** windows. You can also open a window displaying the contents of the clipboard. Windows that are hidden are shown in italics.

There is one point to remember about curve fit windows. The curve fit window and the plot window can be regarded as two parts of the same document. Closing a curve fit window does not close the document. In this manner curve fit windows may be hidden to reduce clutter. The hidden windows can be opened from the **windows** menu or by bringing the plot window to the front and choosing **Curve Fit…** from the **Fit** menu. However if the *plot* window is closed then the plot file will be closed along with the curve fit window.

Function Syntax.

Functions entered in the function text box must conform to a set of syntax rules. The symbols for the allowed arithmetic operators are:

+	addition
-	subtraction, unary minus
*	multiplication
/	division
^	exponentiation

The recognized operands are **x**, **a**, **b**, **c**, **d**, **e**, **f**, **g**, **h**, **pi**, **π** and numeric constants. **X** is the function's argument and **a**, **b**, **c**, **d**, **e**, **f**, **g** and **h** are variable coefficients. Constants can also be used, e.g.

$$f(x) = 2.3*x + \pi$$

In the above function, **2.3** is a numeric constant. The numeric constants can also be entered in scientific format, i.e. **2.3e+0**. Pi can either be typed as '**π**' (option-p) or as '**pi**'.

The function you define can also include other standard mathematical functions. The functions that are recognized are:

sin()	sine
cos()	cosine
tan()	tangent
asin()	arcsine
acos()	arccosine
atan()	arctangent

ln()	natural logarithm (base e)
log()	common logarithm (base 10)
exp()	natural exponent (e^x)
sinh()	hyperbolic sine
cosh()	hyperbolic cosine
tanh()	hyperbolic tangent
asinh()	inverse hyperbolic sine
acosh()	inverse hyperbolic cosine
atanh()	inverse hyperbolic tangent
sqrt()	square root

Comments may be appended to the end of a function by using a semicolon to mark the function's end.

f(x) = a*x + b*x/exp(1-x*x); a comment may be appended here

Tip. In the above example, x^2 is typed as '**x*x**' rather than '**x^2**'. For reasons I won't go into, the multiplication operation is much faster than the exponentiation operation. Using 'x*x' rather than 'x^2" will significantly speed up plotting and curve fitting.

MacCurveFit is not case sensitive when it reads the user defined function, hence a combination of upper and lower case characters can be used. The function can occupy more than one line of text in which case the function text box will wrap the text. Return characters can be typed to end a line prematurely or to leave a blank line. All white space characters (i.e. tab, return, space etc.) between operands and operators are ignored. If the function text requires more lines than are available in the text box the text can be made to scroll by dragging the mouse or by using the arrow keys. After the function has been typed, it is locked in by pressing the **enter** key.

When the function is locked in it is internally converted to machine code so that it can be evaluated extremely efficiently. Those who are interested in the technical details are referred to my article in MacTutor (now known as MacTech) magazine 1992, *8(3)*, 24.

The Functions Folder.

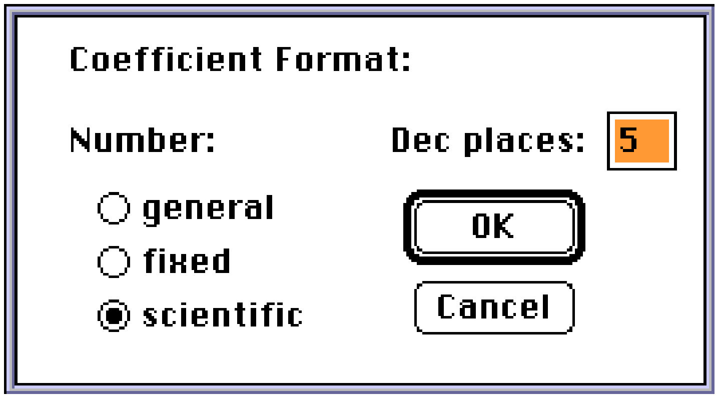
If there are functions that you use often you can store them in a folder named "Functions" in the *same folder that contains the MacCurveFit application*. When MacCurveFit starts up it looks for the Functions folder and loads the names of all the enclosed function files into the function popup menu. This way they can be easily selected without you having to locate the files from a file dialog box. If you wish you may keep the Functions folder elsewhere, provided that you create a Finder alias for the Functions folder. The alias should be called "Functions" and should reside in the same folder that contains the MacCurveFit application. If using an alias the actual folder containing the functions may have any name, only the name of the alias is important. In this way you can maintain several function folders and control which one is opened at startup by using the appropriate alias file.

The Function Coefficients.

The curve fit window displays the function's coefficients in the cells below the function text box. The cells are labelled by the coefficient a, b, c etc. The cell immediately to the right of the label contains the coefficients value and the cell to the right of that displays the uncertainty in the coefficient. The coefficients can be typed directly into the cells like typing into a cell in the data window. When all of the coefficient cells have values then the function will be plotted in the plot window.

Changing the Numerical Format of the Fit Coefficients.

The numerical format of the coefficients can changed by selecting **Coeff Format** from the **Fit** menu. The following dialog box appears.



The dialog box behaves in the same manner as the data window's **Column Format** dialog box.

Assessing the Quality of the Least Squares Fit.

The area of the curve fit window just below the function popup menu displays two parameters which give you a quantitative indication of the fit. The first indicator **SSE** is the sum of squares error and is defined as

$$SSE = \sum(\psi_i - \phi(\xi_i, \alpha, \beta, \chi, \delta, \varepsilon, \phi, \gamma, \eta))^2$$

x_i and y_i are the i th data pair in the data window, f is the function in the text box and a, b, … h are the coefficients. For a perfect least squares fit SSE would equal zero, the larger the value of SSE the poorer the fit. The value $\mathbf{R^2}$ is the correlation coefficient and is calculated as

$$R^2 = 1 - \frac{\nu\sum(\psi_i - \phi(\xi_i, \alpha, \beta, \chi, \delta, \varepsilon, \phi, \gamma, \eta))^2}{\nu\sum\psi_i^2 - (\sum\psi_i)^2} = 1 - \frac{\nu\Sigma\Sigma E}{\nu\sum\psi_i^2 - (\sum\psi_i)^2}$$

where n is the number of data points. A perfect fit has a correlation coefficient of 1 and the lower the value the poorer the fit.

The coefficients can be manually adjusted and the quality of the curve fit can be viewed graphically in the plot window as well as quantitatively by inspecting the sum of squares error and the correlation coefficient.

Performing Curve Fitting.

The object of least squares curve fitting is to minimize the sum of squares error. There are 4 methods by which MacCurveFit can minimize the SSE. There are a number of built-in functions in the program, these are the functions visible in the top section of the function popup menu e.g. linear, polynomial etc. When one of these is chosen all 4 fitting methods will be available and will be listed in the popup menu at the bottom of the curve fit window. The **Special** algorithm is the default method and is capable of fitting the function in a very efficient manner. Consider, as a simple example, the function

$$y = \alpha\xi$$

It can be shown that the least squares solution is

$$a = \frac{\sum \xi_i \psi_i}{\sum \xi_i^2}$$

Hence the special algorithm can calculate the best fit value of the coefficient, a, directly. In general if the **Special** algorithm is enabled you should use this as the preferred fitting method All you need do is click the **Fit** button and everything runs automatically.

However if you have entered a function yourself then the program won't be equipped with a special method for that function. In these circumstances there will be three general purpose algorithms you can choose from to minimize the SSE. The choices are the **Steepest Descent**, **Quasi-Newton**, and the **Newton** algorithm. The default is the quasi-Newton which will give the best results in most cases.

Steepest Descent Curve Fitting.

All of the reiterative curve fitting algorithms rely on the user to supply a starting set of coefficient values to be optimized. The success of the curve fitting depends on how close the starting set of values are to the optimal values. The Steepest Descent method is a simple method and can be quite slow when many coefficients are being optimized. However this method can perform well even when the starting set of coefficients are a long way off the optimal values.

The Steepest Descent method can be visualised as follows. Consider a contour map of a collection of mountains and valleys. Any point on the map can be characterized as having two position coordinates, longitude and latitude, and a third value indicating height. If we want to get to the lowest point on the map we need to walk down the mountains and into the valleys. The contour map indicates which way to go. This corresponds to a situation where we are optimizing two coefficients to fit a function to a set of data points. The two coefficients, a and b, are coordinates on a contour map, the contours indicate the sum of squares error (SSE) at that point. To get the best fit we want to minimize the sum of squares error, i.e., find the lowest point on the map. The Steepest Descent method is a two step reiterative process. The first step is to take a small step in the 'a' direction to test the steepness in this direction and then take a small step in the 'b' direction. This enables the calculation of which direction provides the steepest descent. The second step is to conduct a line search, i.e., to proceed along this direction until the lowest point is found. Since this new point is lower than the starting point, the sum of squares has been decreased and a better fit has been found. This new point is then used as a starting point for the next iteration.

Newton Method.

The Newton method is a rapidly converging algorithm and is the method of choice for systems that are already close to the best fit. Unfortunately, this method does not perform at all well when you are a long way from the best fit, under these conditions the method may not converge at all.

The Steepest Descent method doesn't converge rapidly because it is essentially short sighted. It works out the best direction to go based on the gradient at the current point. The steepest direction found in this manner doesn't necessarily continue to be steep as you move along that way. Furthermore, the method has to find out the step length by trial and error. However, the Newton method calculates not only the gradient vector (steepness) but also the Hessian matrix (curvature) at the current point. By assuming that the contour surface is quadratic, it can calculate not only which direction to go but how far it should go. Unfortunately, the basic assumption breaks down when you are a long way from the minimum sum of squares and the algorithm doesn't converge. One further point to mention is that the algorithm doesn't actually look for a minimum in the sum of squares. It proceeds in a direction towards a point where the gradient vector is zero, i.e., it may converge to a maximum sum of squares thus giving a worse fit.

Quasi-Newton Method.

The quasi-Newton method used by this program is the Davidon-Fletcher-Powell algorithm. It is a compromise between the Steepest Descent method and the Newton method. It has the stability of the former and the rapid convergence of the latter method, for this reason it is the default algorithm.

The method involves two steps with the first being the calculation of a good direction to go. This direction is not necessarily the direction of steepest descent. The second step is to conduct a line search and hence find the lowest point along this direction.

For a good introduction to Nonlinear Optimization, read "Practical Methods of Optimization. Volume 1, Unconstrained Optimization" by R. Fletcher (Wiley) 1980-81.

Calculation of the Uncertainties in the Coefficients.

After a curve fit has completed, the uncertainties in the fir coefficients will be estimated automatically and displayed in the cells next to the coefficients. However you may have arrived at a fit by adjusting the coefficients manually or perhaps you may have lost the uncertainties by changing a coefficient and then restoring it. In these cases you may want to know the uncertainties without actually running the fitting algorithms. You can request MacCurveFit to estimate the uncertainties at any time by choosing **Calc Coeff Errors** from the **Fit** menu.

The coefficient uncertainties are estimated from the variance-covariance matrix and the values displayed by MacCurveFit are the square roots of the diagonal elements. The variance-covariance matrix is calculated from the Jacobian matrix, this is described in Fletcher's book referred to above.

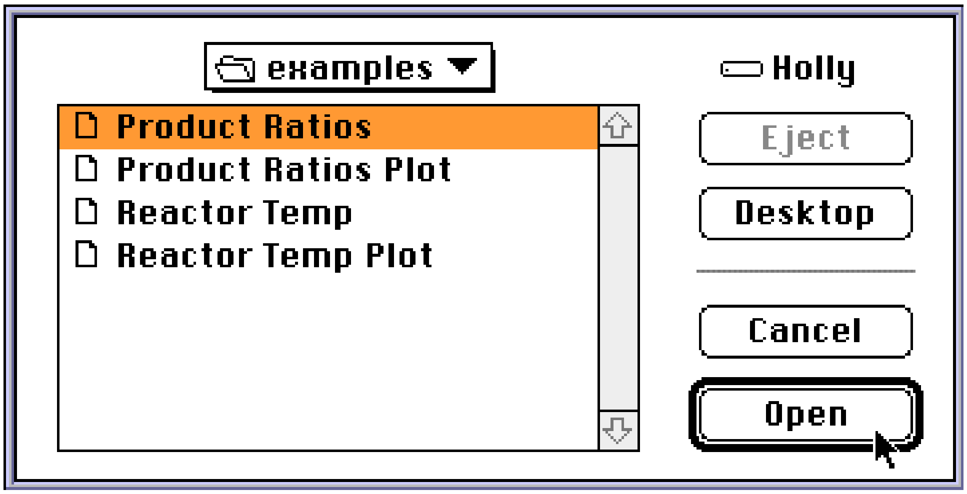
MacCurveFit and the Cooperative Multitasking Environment.

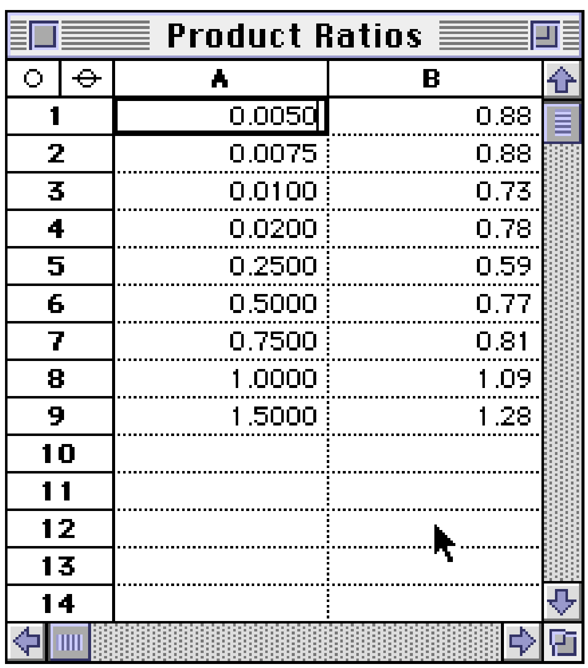
Cooperative multitasking is Apple's jargon for the way the Macintosh allows more than one program to run at once. Apple intends that applications are "good citizens" are do not hog the microprocessor. MacCurveFit is a model citizen and never hogs the microprocessor for more than 0.5 seconds when curve fitting. This enables you to continue work with other applications while MacCurveFit crunches along in the background. The downside of being considerate is that MacCurveFit's performance suffers. For those who insist on the highest possible performance, future versions will be configurable so that you will be able to seize the whole machine if you feel the need.

Tutorial 1.

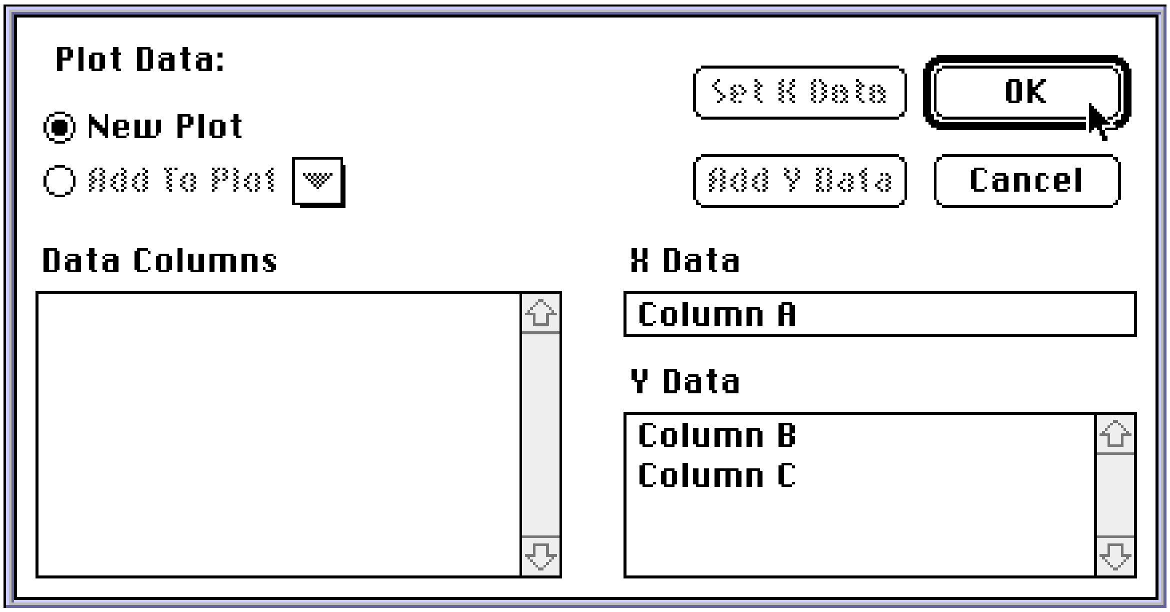
This tutorial will illustrate how this pieces fit together and show you how to do some curve fitting. The problem chosen here is the problem that prompted me to write MacCurveFit. It involves a kinetic investigation of a particular free radical chemical reaction. If you are not interested in chemistry I'll spare you the details.

Launch MacCurveFit by double clicking its icon in the Finder. When it has started up select **Open…** from the **File** menu and choose "Product Ratios" from the dialog box. Click on the **Open** button and a new data window will open displaying the data from the file.

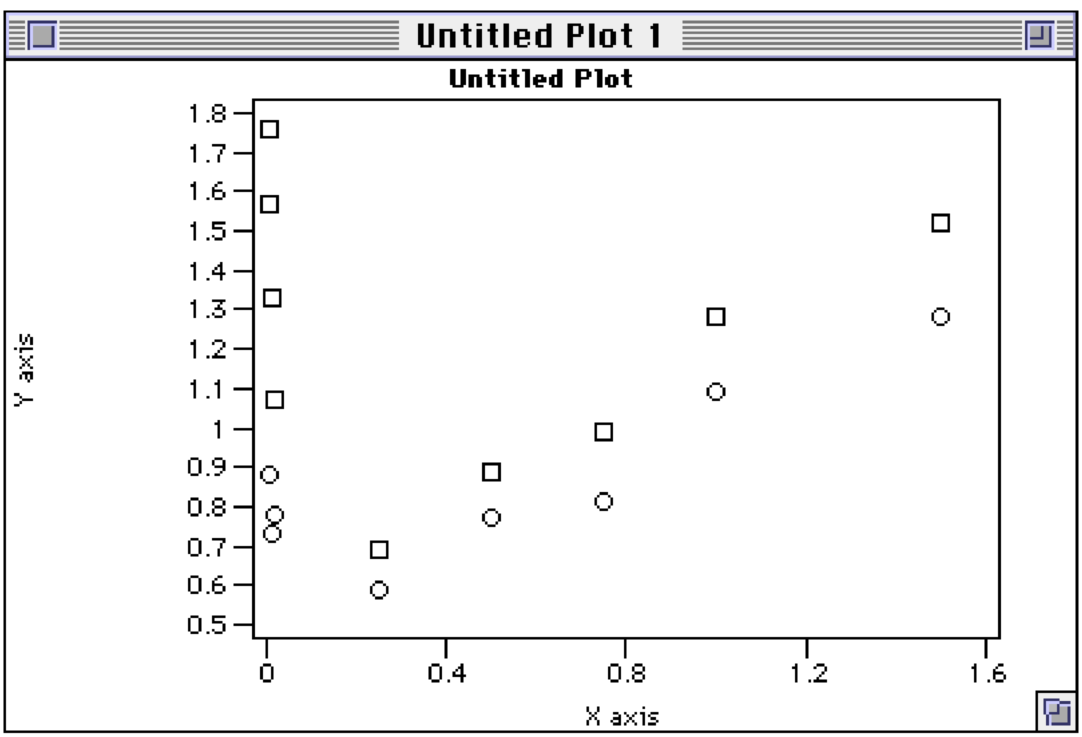




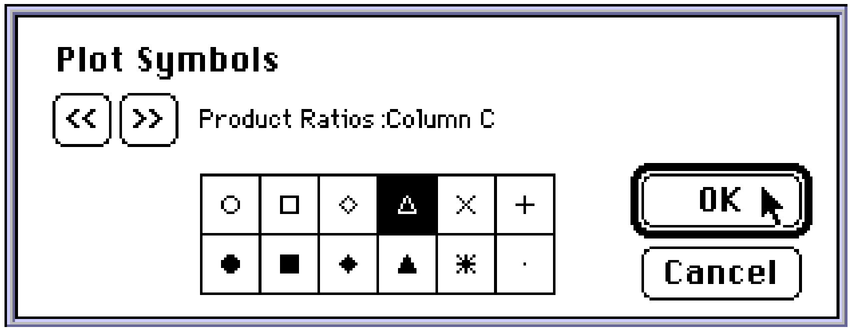
The data window has three columns. The first columns contains concentration of tri-n-butylstannane in a series of chemical reactions and the second and third columns list two ratios of the reaction products. I'll just refer to the products as **5**, **6**, **7** and **8**; and the ratios are ([**5**]+[**6**])/[**7**] and ([**5**]+[**6**])/[**8**]. The next step is to plot the ratios in columns B and C against the stannane concentrations in column A. To do this choose **Plot Data…** from the **Data** menu. A dialog box will appear and you should click on column A in the left list then click the **Set X Data** button on the right hand side. Next hold down the option key while you drag across the columns B and C. Both columns will be highlighted and you can then click the **Add Y Data** button. The dialog box should be as shown below.



Click on the **OK** button and a new plot window will appear.



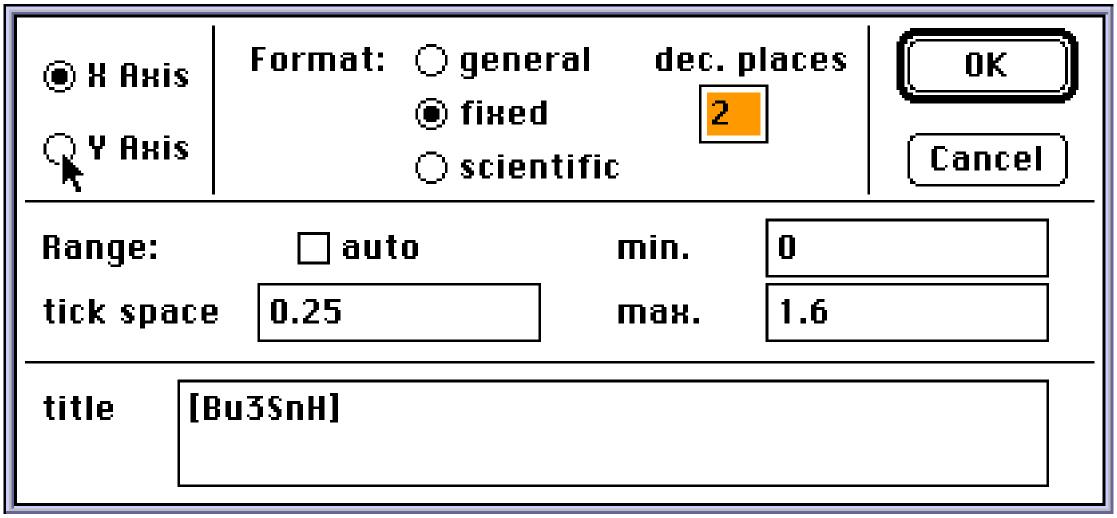
The data from column B will be displayed as circles and the data from column C will be represented by squares. Let's change the plot so that the ratios in column C are represented as triangles. Choose **Data Symbols…** from the **Plot** menu. Click the right arrow button to change the selected column to "Product Ratios:Column C". Then click the triangle symbol followed by the **OK** button.

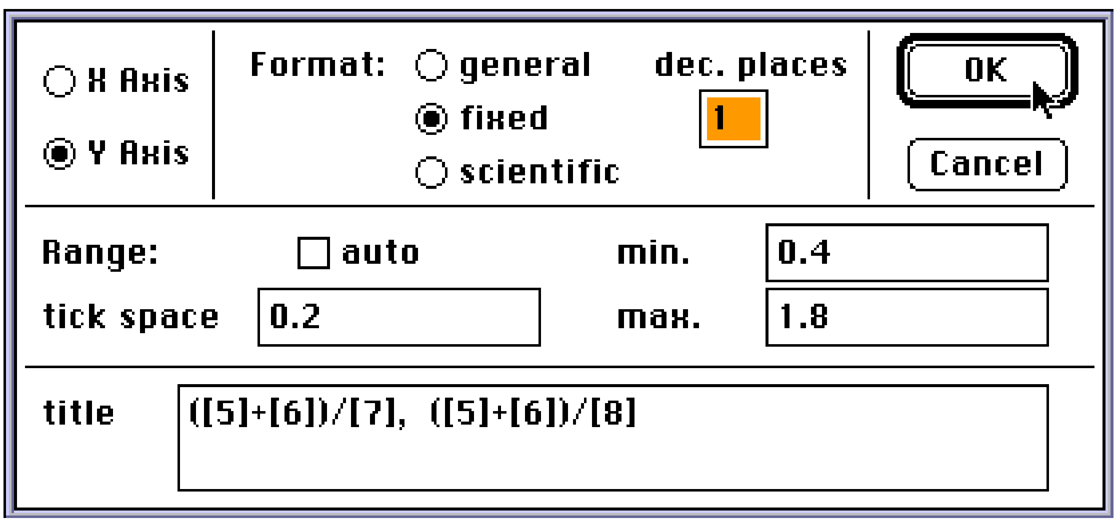


The plot will be redrawn with triangles marking the data from column C. Next change the title of the plot by choosing **Title…** from the **Plot** menu. Type "Dependence of Product Ratios on Stannane Concentration" into the dialog box and click **OK**.

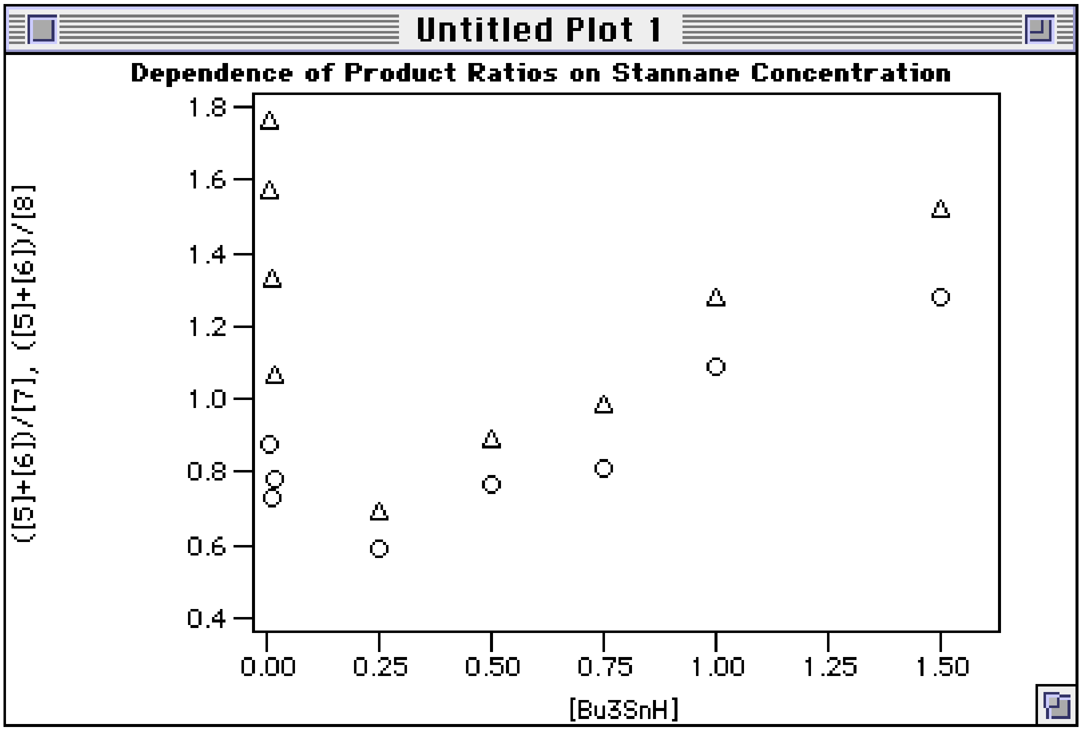
Most changes are easy to undo in MacCurveFit. To see this choose **Undo Plot Title** from the **Edit** menu and then **Redo Plot Title** from the same menu.

Finally let's get the axes looking the way we want (I'll assume for the moment that we agree on style). Bring up the axes dialog box by choosing **Axes…** from the **Plot** menu. Make the necessary changes so the dialog box matches the figure on the top of the next page. Then click the **Y Axis** radio button and set things to match the figure at the bottom of the next page.


X Axis
Y Axis
Format:
general
fixed
scientific
dec. places
2
OK
Cancel
Range:
auto
min.
0
tick space
0.25
max.
1.6
title
[Bu3SnH]


X Axis
Y Axis
Format:
general
fixed
scientific
dec. places
1
OK
Cancel
Range:
auto
min.
0.4
tick space
0.2
max.
1.8
title
([5]+[6])/[7], ([5]+[6])/[8]

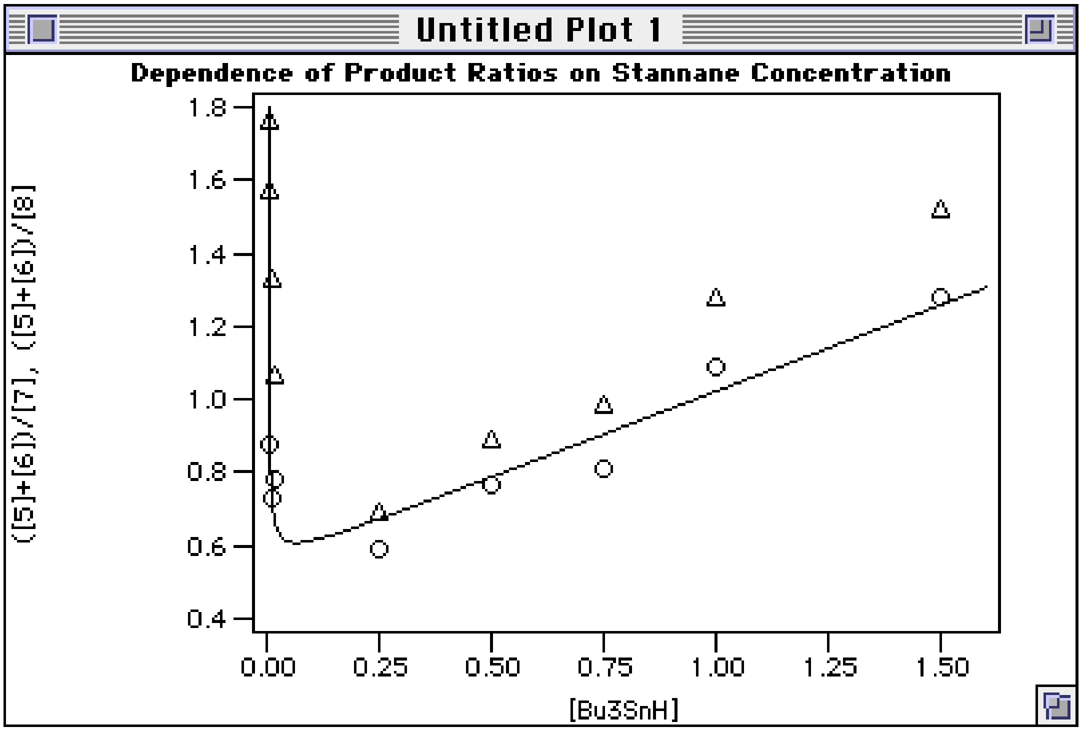
The plot will now look like the one shown below.

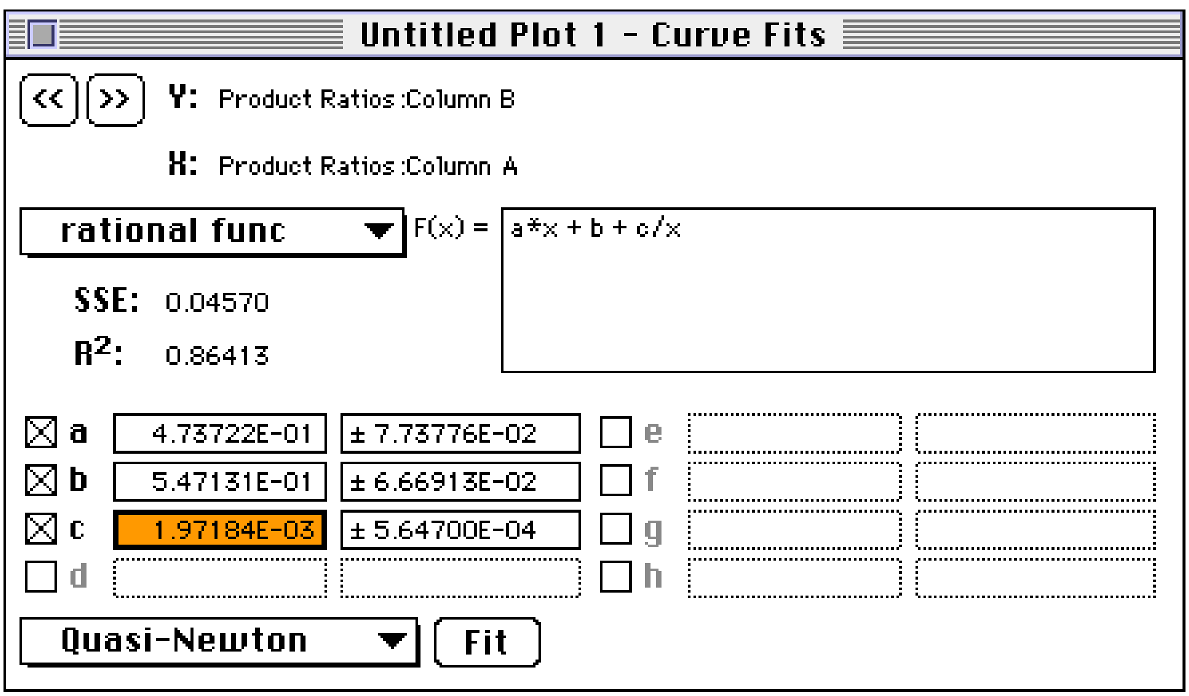


At last it's time to do some curve fitting. We'll fit the function $y = ax + b + c/x$ to each data column. Summon the curve fit window by selecting **Curve Fit…** from the **Fit** menu. Choose the function "rational func" from the function popup menu. If it is not in the menu quit the program and check that the supplied "Functions" folder is in the same folder as the MacCurveFit application.

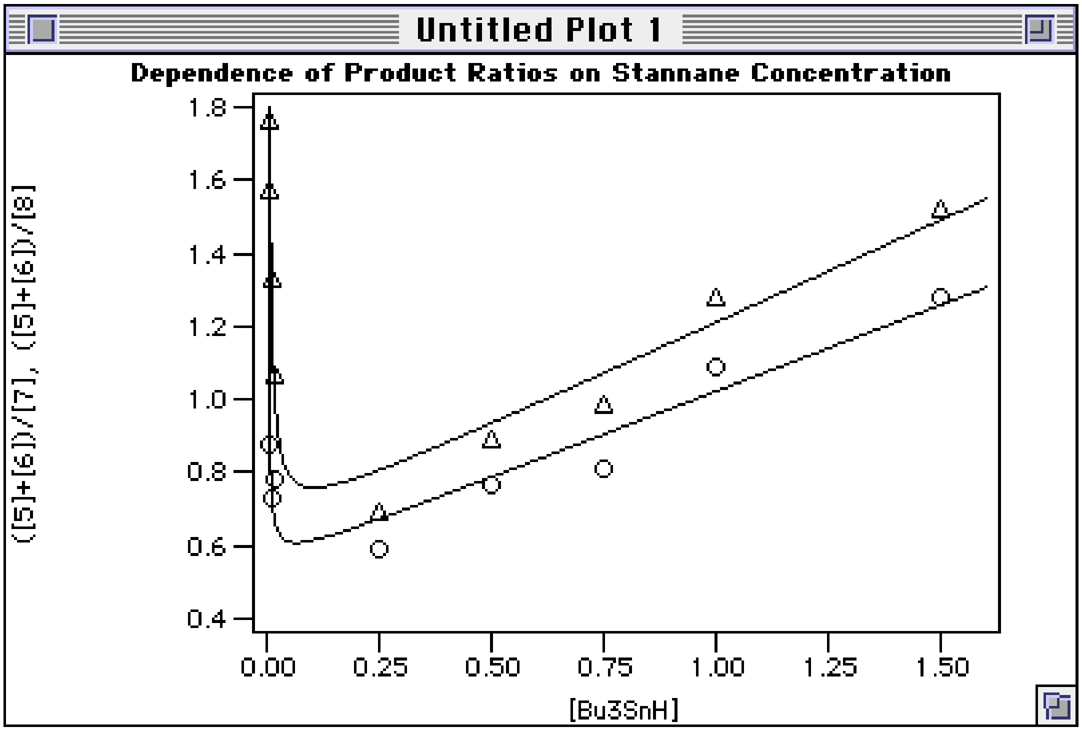
Next give the fitting algorithm a starting point to work from, type 1 into the cells for a, b and c. Then click on the **Fit** button. You will see the SSE, the R^2 and the coefficients updated every iteration. The algorithm quickly arrives at a best fit as shown on the next page.

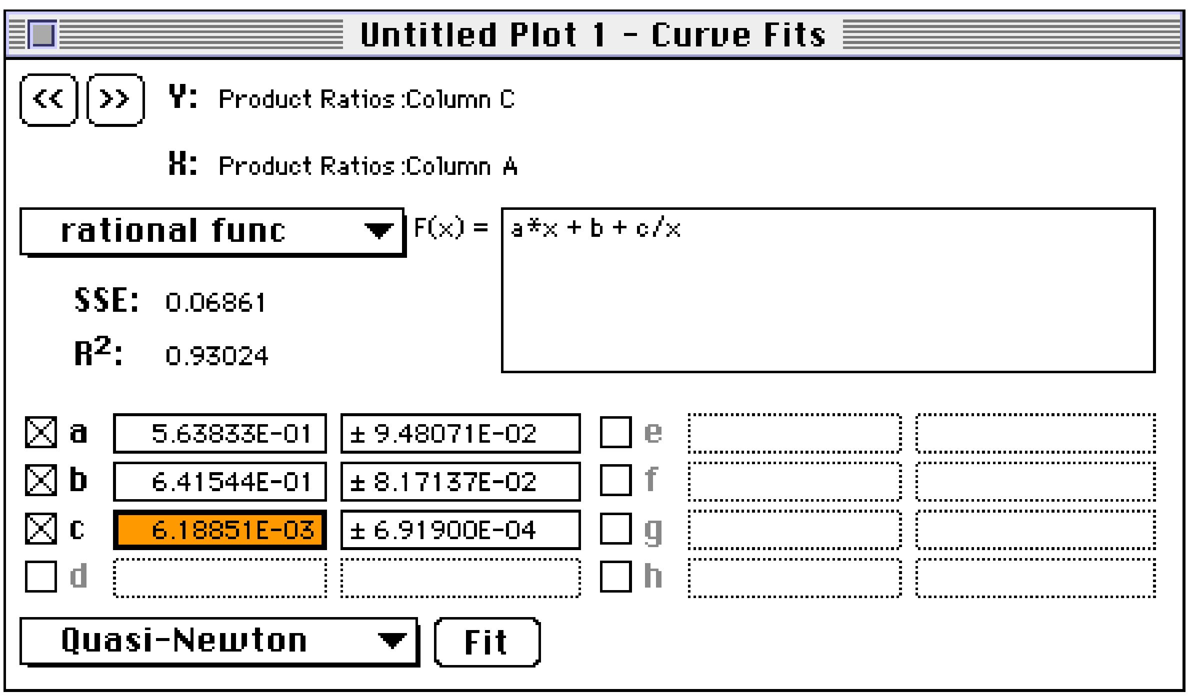
Earlier I mentioned that MacCurveFit's printing code always strives to give the best results possible. If you print this document you may be puzzled by the poor quality. To show you how the whole window looks as opposed to how the plot itself looks, I have simply presented a screen dump (command-shift-3) in this manual. If you print the plot directly from the application at this point you'll be more satisfied by the printing quality.





To fit the function through the second set of data click the right arrow button in the curve fit window to select the data in column C. Repeat the earlier proceedings, namely choose the rational function and enter 1 into the cells for a, b and c. Click the **Fit** button. You should now have a plot that looks like the one on the next page.





Finally save the plot by selecting **Save** from the **File** menu while either the plot or curve fit window is the front window.

Now back to the chemistry, from the values of a, b and c, it was possible to calculate the kinetic rate constants of certain steps in the free radical reaction I was investigating.

Tutorial 2.

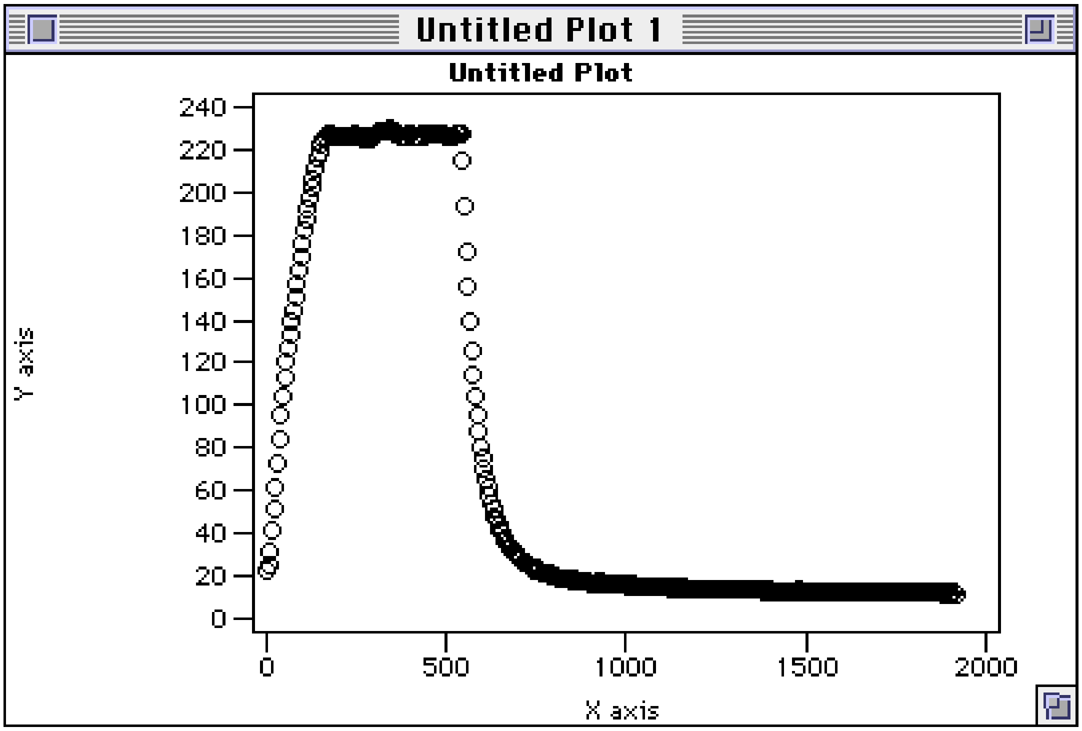
This tutorial demonstrates the use of those funny icons in the top left corner of the data window. Start up MacCurveFit and open the file “Reactor Temp” (in the examples folder). Since the basics have been covered in Tutorial 1 you’ll forgive me for not going through the tedium again.

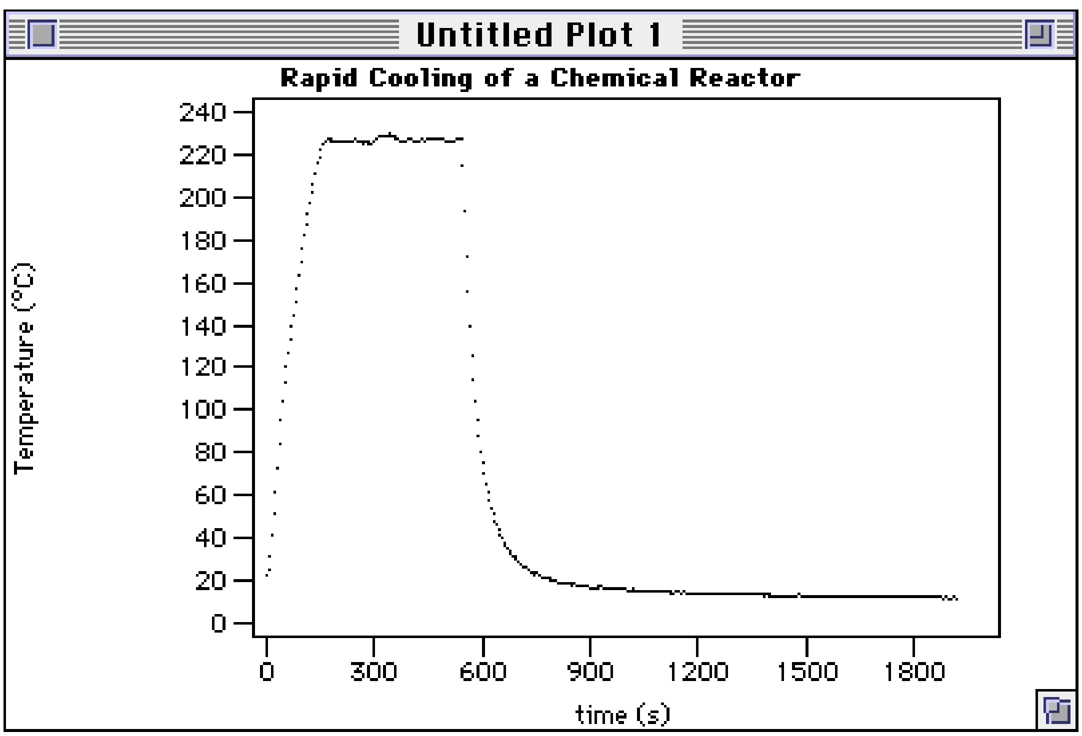
Reactor Temp

	A	B
1	0	22.15
2	5	24.83
3	10	31.79
4	15	40.88
5	20	51.23
6	25	61.49
7	30	72.93
8	35	83.77
9	40	95.06
10	45	104.34
11	50	112.81
12	55	120.32
13	60	127.35
14	65	133.30

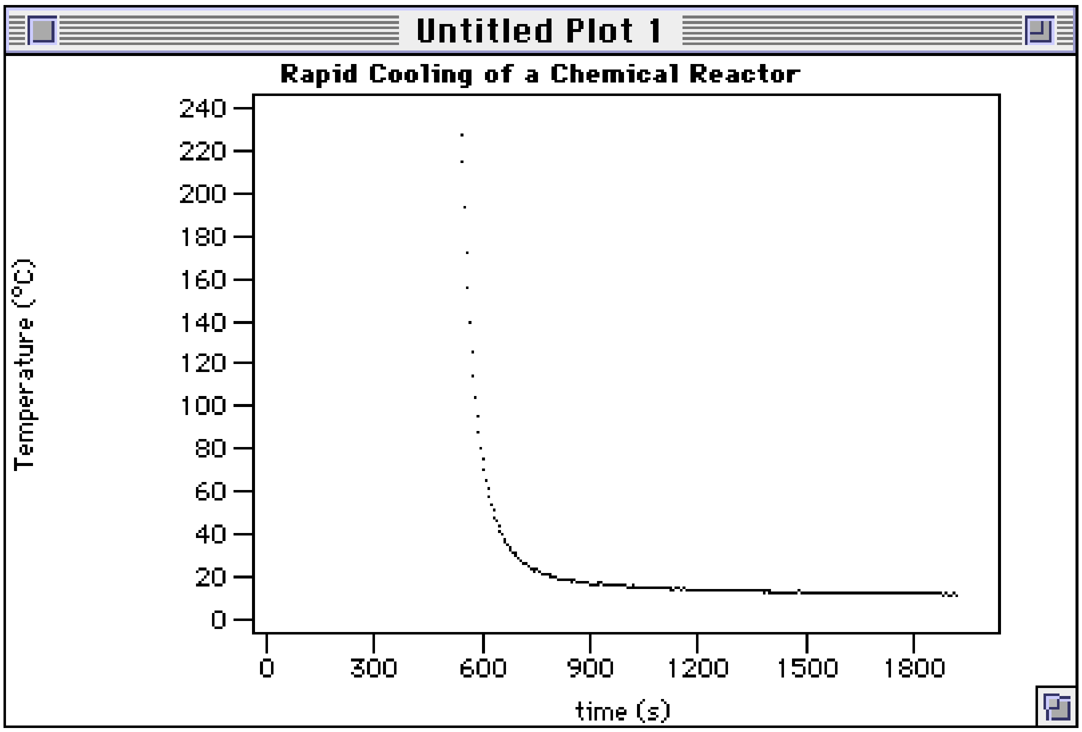
The window contains temperature data from a microwave chemical reactor that has been rapidly heated, held at a constant temperature for a while and finally cooled using a patented cooling device. Column A lists the time in seconds since the reactor was switched on and column B is the temperature in degrees Celcius.

Plot the temperature against the time and you’ll get a plot like the one on the top of the next page. Because of the large number of temperature measurements the plot looks quite unattractive. You can remedy this by changing the plot symbols to single pixels. You can also label the axes and give the plot a title. Your spruced up plot should look like the one at the bottom of the next page.


Untitled Plot 1
Untitled Plot
Y axis
X axis
0
20
40
60
80
100
120
140
160
180
200
220
240
0
500
1000
1500
2000


Untitled Plot 1
Rapid Cooling of a Chemical Reactor
Temperature (°C)
time (s)
0
20
40
60
80
100
120
140
160
180
200
220
240
0
300
600
900
1200
1500
1800

Next select the data window to bring it to the front. Make the cell B1 the active cell then scroll the window to show the cell B108. Hold down the shift key and click cell B108, you should now have a range of cells selected. Click the left small icon in the top left corner of the data window and observe the plot. All of the points before 540 seconds disappear from view as shown below.



Click the left icon again and the points will return. Let's fit an exponential decay to the cooling part of the curve. While the range B1:B108 is still selected click on the right icon. The plot will look no different however the points will not be considered when curve fitting.

Open the curve fit window and select the function "exp decay" from the function popup menu. Set the value of coefficient a to 100 and b to 0.01. The function will be plotted as shown on the top of the next page. Since our temperature curve doesn't decay to zero edit the function text to read

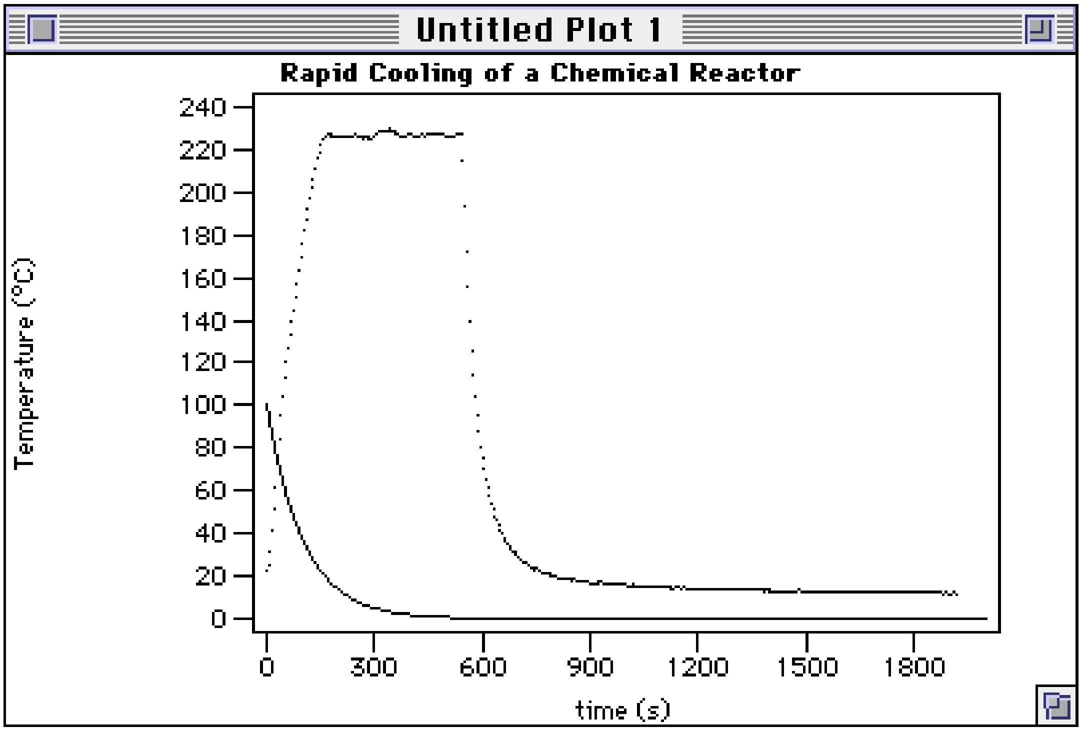
$$f(x) = a*exp(-b*x) + c$$

Press the enter key and then set the value of coefficient c to 20. Also, since the start of the cooling is at 540 seconds, edit the function again to read:

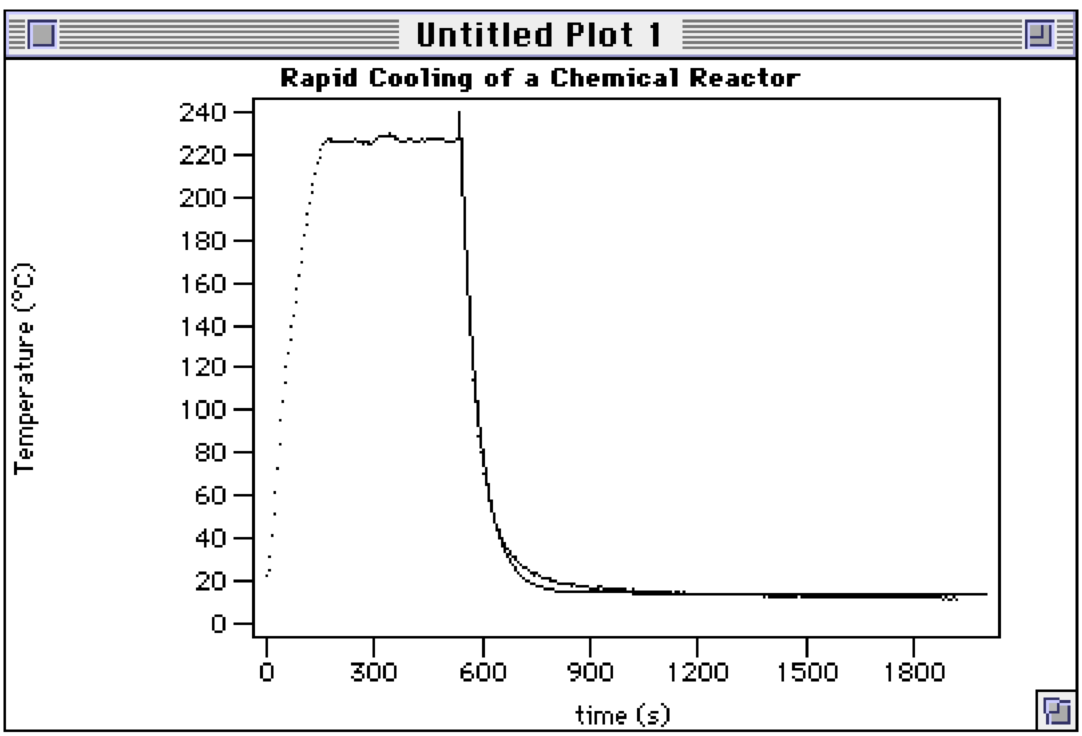
$$f(x) = a*exp(-b*(x-540)) + c$$

and then press the enter key. The more astute reader will notice that the two functions above are equivalent. However the latter gives better performance in curve fitting. Whenever the optimum coefficients are different by several orders of magnitude the algorithms in MacCurveFit may terminate prematurely. (Try curve fitting with the first function). The problem will be rectified in version 1.1. However as a general tip try and construct your functions so that this situation doesn't arise.

Next click the **Fit** button. The resulting plot is shown on the bottom of the next page.



Untitled Plot 1
Rapid Cooling of a Chemical Reactor
Temperature (°C)
240
220
200
180
160
140
120
100
80
60
40
20
0
0
300
600
900
1200
1500
1800
time (s)



Untitled Plot 1
Rapid Cooling of a Chemical Reactor
Temperature (°C)
240
220
200
180
160
140
120
100
80
60
40
20
0
0
300
600
900
1200
1500
1800
time (s)

The optimum values of the coefficients will be:

a = 209.7 ± 1.5

b = 0.01942 ± 0.00021

c = 14.43 ± 0.17

The sum of squares error was reduced to 1824.3 and the correlation coefficient (R^2) was.0.99215.

Well that covers the general usage of MacCurveFit. You can now start to do some curve fitting of the own, which I'm sure you'll find far more interesting.

Future Directions.

There are a number of things that I want to put into version 1.1. I promised someone (a year and a half ago) that I would implement error bars in the plots. This is top of the list at the moment. Other good suggestions I have received include a routine that reads a column of x values and outputs a column of predicted y values. I also want to allow formulae to be associated with a column so that you can type e.g. x values in one column and have logx values calculated automatically in another column. Another suggestion was to copy the fitting equation and coefficients to the clipboard in a format ready to paste into a word processor or spreadsheet program.

I mentioned earlier that MacCurveFit's number crunching performance is compromised so that it can behave correctly under multifinder. I want to have an option to allow number crunching of large jobs to be carried out fast without yielding any time to other programs. I have also developed an FPU version of my equation compiler as part of a contract for a commercial company. The company granted me permission to use the FPU equation compiler in my own programs. This means I can now bring out a high performance version of MacCurveFit if anyone is interested. I am planning to exclude the FPU version of MacCurveFit from the public domain but will happily provide it free of charge (plus postage and handling) to anyone who has paid their shareware fee for the regular version of MacCurveFit.

I am now planning to take a little interlude from MacCurveFit as version 1.0 has commanded most of my spare time for the last two years. I am looking forward to trying out an interesting idea that may eventuate in another MacTech (MacTutor) article. I also want to play with my new (Symantec) C++ compiler.

Finally, thankyou for your encouragement. The email messages and postcards helped me to keep going and finish version 1.0. And Yes! I received a postcard of York Minster.