Table of Contents

Nonlinear Curve Fitting: Fit Plot 1
- ***Fit Function is a Sum of Fit Curves*** 1
- ***Coping Fit Curves between Fit Plots*** 2
- ***Setting Initial Values of Parameters*** 2
- ***Parameter Locking*** 2
- ***Fit Intervals*** 2
- ***Baseline Fitting and Extraction*** 3
- ***'Data-Baseline' Table Column*** 3
- ***Fit One Curve*** 3
- ***Joining the Parameters of Fit Curves*** 4
- ***Fitting Algorithm*** 4
- ***Weighting of y data*** 5
- ***Standard Deviation of Fit Parameters*** 6
- ***Formulas*** 6

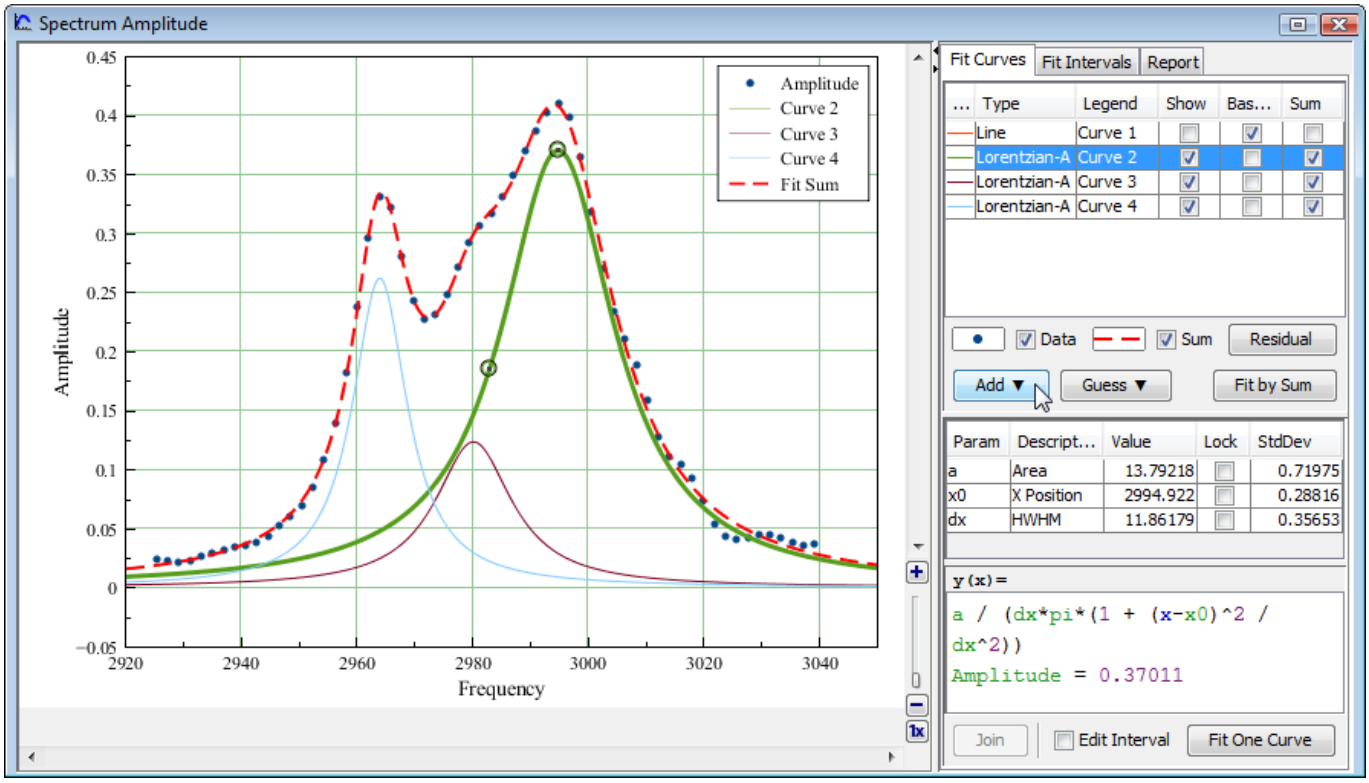
Nonlinear Curve Fitting: Fit Plot

Nonlinear least squares data fitting can be performed using Fit Plot. To create Fit Plot select x and y columns in table, then select `Table → Create Fit Plot` in main menu, or use context menu in table, or use Create Fit Plot button in the toolbar.

'Nonlinear' means here that analytical fitting function depends nonlinearly on varying parameters (fit parameters). Linear fitting is a quite simple method, which is based on solving the system of linear equations. Unlike linear fitting, nonlinear fitting is performed by iterative algorithm which needs the user to set the initial values of fit parameters.

To fit the data you must implement these steps:

1. Create Fit Plot, specify y weighting in Plot properties, if any
2. Specify fit function by adding Fit Curves
3. Specify initial values of fit parameters (drag curves or enter accurate values)
4. Specify used x data interval
5. Run fitting



Fit Function is a Sum of Fit Curves

MagicPlot considers fit function as a **sum** of Fit Curves. Ordinarily in peaks fitting each Fit Curve corresponds to one peak on the experimental data. There are a number of predefined Fit Curves (Line, Parabola, Gauss, Lorentz, etc.) You can specify custom Fit Curve too. The baseline fitting components also may be added to the fitting sum.

Fit Plot window contains the list of Fit Curves. Each Fit Curve in the list has 3 check boxes: Show,

`Baseline`, `Sum`:

- `Show`: Specifies whether to show this Fit Curve on plot. Active only if Baseline check box is not set
- `Baseline`: Toggles the subtracting of this Fit Curve from experimental data
- `Sum`: Specifies whether to use this Fit Curve in sum fit function

Below the Fit Curves list is parameters table which shows parameters names, values, descriptions of the selected Fit Curve.

Coping Fit Curves between Fit Plots

You can copy and paste Fit Curves in curves table as usual. Use context menu in curves table or press `Ctrl+C`, `Ctrl+V` on PC and `Cmd C`, `Cmd V` on Mac (curves table must be focused, not the plot itself). MagicPlot copies Fit Curves parameters values, baseline/sum/show checkbox values and style.

Setting Initial Values of Parameters

Nonlinear fitting assumes that certain initial values of parameters are set before fitting. This procedure is just easy if you use predefined Fit Curves: you can drag curves on plot.

Initial parameters values for each Fit Curve can also be set in parameters table.

Parameter Locking

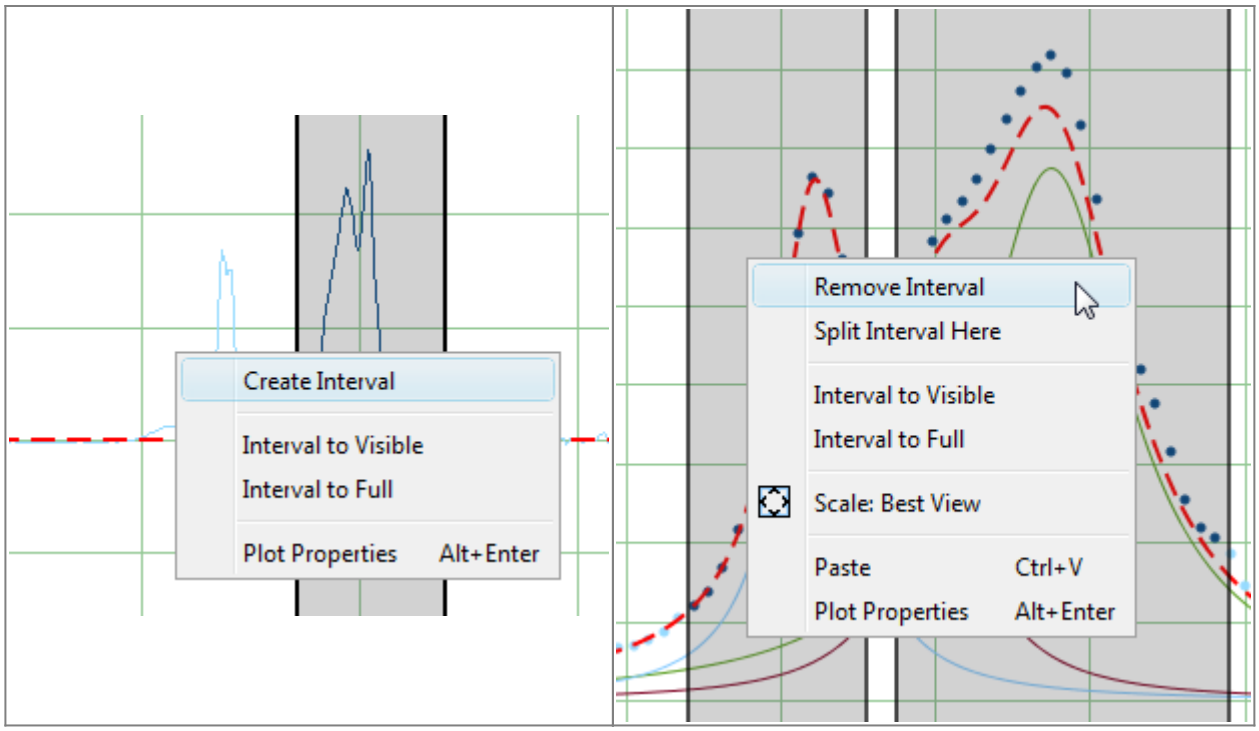
You can lock parameter(s) to prevent varying this parameter during fit and to prevent its changing due to setting initial values by mouse dragging (for built-in functions). Use checkbox on `Lock` column in parameters list.

Fit Intervals

You can set the x intervals of the data. The data points outside this intervals are not used to compute the minimizing residual sum of squares (see below). You can use this feature if some data points (especially in the beginning and the end) are inaccurate, e.g. noisy.

Select `Fit Interval` tab to set intervals visually or edit accurate borders values:

- Double click on interval to split it
- Drag the interval border to move it. If intervals intersect, intervals will be merged
- Use context menu on the plot to create, delete and split intervals



Baseline Fitting and Extraction

Fit Interval is also usable when baseline fitting. Before baseline fitting you can specify the interval which does not contain any signal points and contains baseline only. Set `Baseline` check boxes at baseline Fit Curves after baseline fitting to subtract baseline from data. Then set interval to full and fit the data.

Note that if you execute one of data processing algorithms (integration, FFT, etc.) on Fit Plot, then the difference between the data and baseline curves (which you do see on the plot) will be processed. You can use this behaviour to exclude baseline from data before integrating, see Integration for more information.

'Data-Baseline' Table Column

The 'Data-Baseline' column is appended to the Table with initial (x, y) fit data when you create Fit Plot. The 'Data-Baseline' column contains the difference between initial y data and baseline approximation (the sum of Fit Curves for which `Baseline` checkbox is selected).

It is 'Data-Baseline' column that is actually plotted on Fit Plot.

Use 'Data-Baseline' column in Table if you want to process the data without baseline. This column is also used as initial data if you use `Processing` menu when Fit Plot is active.

Fit One Curve

You can also use MagicPlot to fit the data with single selected Fit Curve by pressing `Fit One Curve` button. In this case a specific data interval for each Fit Curve is used and the main fitting data interval

(set in `Fit Interval` tab) is ignored. Select `Set Interval` check box in the bottom of the Fit Plot panel to set specific fit intervals for each Fit Curve.

Joining the Parameters of Fit Curves

In some cases you may want to fit the data with two Gauss or Lorentz peaks with the same width but different positions and amplitudes, for example. You can do this in two ways: by specifying custom Fit Curve with your equation or by *joining* the 'width' parameters of two peaks.

To join parameters of two or more Fit Curves select one of desired Fit Curves, select desired parameter in parameters table and press `Join` button. Add parameters which will be joint in the opened dialog window. Joined parameters are treated as one fit parameter.

Joined parameters are showed with blue color (instead of black) in parameters table.

Fitting Algorithm

MagicPlot uses iterative Levenberg–Marquardt nonlinear least squares curve fitting algorithm which is widely used in most software.

Fit procedure iteratively varies the parameters β_k of fit function $f(x, \beta_1, ..., \beta_p)$ to minimize the residual sum of squares (RSS, χ^2):

$$\chi^2 = \sum_{i=1}^{N} w_i \left(y_i - f\left(x_i, \beta_1, ..., \beta_p \right) \right)^2 \rightarrow \min$$

here:

- x_i and y_i are the data points,
- N is total number of points,
- $f(x, \beta_1,...,\beta_p)$ is the fit function which depends on value of x and fit parameters β_k,
- p is the number of fit parameters β_k,
- w_i are normalized y data weighting coefficients for each point y_i:

$$\sum_{i=1}^{N} w_i = 1$$

The calculation of the new guess of parameters on each fit iteration is based on the fit function partial derivatives for current values of fit parameters and for each x value:

$$\frac{\partial f}{\partial \beta_m}\left(x_i, \beta_1, ..., \beta_p \right)$$

To start minimization, you have to provide an initial guess for the parameters.

Fit Procedure Stop Criteria

After each iteration except the first MagicPlot evaluates *deviation decrement D*:

$$D = \left| \frac{\chi^2_{curr.\,iter.}}{\chi^2_{prev.\,iter.}} - 1 \right|$$

Deviation decrement shows how the residual sum of squares (RSS) on current iteration relatively differs from that on the previous iteration.

The iterative fit procedure stops on the one of two conditions:

- If the deviation decrement *D* is less than minimum allowable deviation decrement, which is 10^{-9} by default

or (and)

- If the number of iterations exceeds maximum number of iterations, which is 100 by default

You can change the minimum allowable deviation decrement and maximum number of iterations in Fitting tab of MagicPlot Preferences.

Fit progress window

MagicPlot indicates fit process with a special window. Fitting curves are periodically updated on plot while fitting so you can see how fit converges.

MagicPlot shows current iteration number and deviation decrement by two progress bars while fit is performed. The fit process stops when one of this progress bars reaches the end.

Two buttons are located on fit progress window:

- `Break Iters`: Breaks iterations after current iteration. Use this button if you suppose that further iterations will not change the result.
- `Undo Fit`: Breaks iterations and reverts the fit parameters to before fit values. Use this button if you see that fit process converges to wrong result; change initial values of parameters and run fit again.

Weighting of y data

MagicPlot can use weighting of *y* values based on y errors s_i:

- If standard *y* errors are **not** specified: all w_i=1
- If standard *y* errors s_i are specified:

$$w_i = C\frac{1}{s_i^2}$$

here *C* is normalizing coefficient (to make the sum of w_i be equal to one):

$$C = \sum_{i=1}^{N} s_i^2$$

In `Fit Plot Properties` dialog (`Plot Data` tab) you can set one of the following method to evaluate standard y errors s_i:

- Get y errors from table column(s),
- Percent of data for every point,
- Fixed value or Standard deviation — do not use in weighting because in this case the error values are the same for all data points.

Standard Deviation of Fit Parameters

The standard deviations (*std. dev.*) of fit parameters β_k are evaluated after fit using the following formula:

$$s_k = \sqrt{\frac{\chi^2}{N-p}\left[\alpha^{-1}\right]_{k,k}}$$

here α is the matrix of partial derivatives of fit function by parameters β_m and β_n which is used for fitting:

$$\alpha_{m,n} = \sum_{i=1}^{N} w_i \frac{\partial f}{\partial \beta_m}\left(x_i, \beta_1, \ldots, \beta_p\right)\frac{\partial f}{\partial \beta_n}\left(x_i, \beta_1, \ldots, \beta_p\right)$$

Formulas

In the table below you can find the formulas which MagicPlot uses to calculate fit parameters and values in `Fit Report` tab.

Because of some confusion in the names of the parameters in different sources (books and software), we also give many different names of same parameter in *note* column.

Parameter Name	Symbol	Formula	Note
Original Data and Fit Model Properties			
Number of used data points	N	—	This is the number of data points inside specified Fit Interval.
Fit parameters	$\beta_1,\ldots,\beta_p$	—	For peak-like functions (Gauss, Lorentz) these parameters are amplitude, position and half width at half maximum.
Number of fit function parameters β	p	—	This is the total number of parameters of all fit curves which are summed to fit.
Degrees of freedom	ν	$N-p$	
Estimated mean of data	$\overline{y}$	$\frac{1}{N}\sum_{i=1}^{N} y_i$	
Estimated variance of data	s^2	$\frac{1}{N-1}\sum_{i=1}^{N}\left(y_i-\overline{y}\right)^2$	Not used by fit algorithm, only for comparison.
Data total sum of squares, TSS	TSS	$\sum_{i=1} w_i\left(y_i-\overline{y}\right)^2$	*TSS is also called sum of squares about the mean and acronym SST is also used.*
Fit Result			

Parameter Name	Symbol	Formula	Note
Original Data and Fit Model Properties			
Residual sum of squares, RSS	χ^2	$\sum_{i=1}^{N} w_i \left(y_i - f\left(x_i, \beta_1, \ldots, \beta_p \right) \right)^2$	This value is minimized during the fit to find the optimal fit function parameters. *RSS is also called the sum of squared residuals (SSR), the error sum of squares (ESS), the sum of squares due to error (SSE).*
Reduced χ^2	$\chi^2_{red.}$	$\frac{\chi^2}{\nu} = \frac{\chi^2}{N-p}$	The advantage of the reduced chi-squared is that it already normalizes for the number of data points and model (fit function) complexity. *Reduced χ^2 is also called mean square error (MSE) or the residual mean square.*
Standard deviation of the model	s	$\sqrt{\left(\chi^2_{red.} \right)}$	*Standard deviation is also called root mean square of the error (Root MSE)*
Coefficient of determination	R^2	$1 - \frac{\chi^2}{TSS}$	R^2 will be equal to one if fit is perfect, and to zero otherwise. This is a biased estimate of the population R^2, and will never decrease if additional fit parameters (fit curves) are added, even if they are irrelevant.
Adjusted R^2	$\overline{R}^2$	$1 - \frac{N-1}{N-p-1}\left(1 - R^2 \right)$	Adjusted R^2 (or *degrees of freedom adjusted R-square*) is a slightly modified version of R^2, designed to penalize for the excess number of fit parameters (fit curves) which do not add to the explanatory power of the regression. This statistic is always smaller than R^2, can decrease as you add new fit curves, and even be negative for poorly fitting models

From:
https://magicplot.com/wiki/ - **MagicPlot Manual**

Permanent link:
https://magicplot.com/wiki/fitting?rev=1288036750

Last update: **Sun Nov 8 12:20:32 2015**