

Finding the Best-Fit Function

Fall 2023

Introduction

Today's exercise will give you further practice using KaleidaGraph for curve fitting. You will be presented with three data sets where you won't already know which function best fits. You will determine the best-fitting function and use KaleidaGraph to create a user-defined fit which will allow you to calculate the uncertainty in the parameters involved, as well as the SSR. *Since we are providing you with three large data sets, it is **not necessary** this week to record this data in your journal. Remember to record your data in all future journals!*

Create a Graph of the Data

1. Start KaleidaGraph, and open *DataSet1* from the **PHYS151** directory on the *Teaching Drive (T:)*.
2. Create a scatter plot from the data in this set. Be sure to put your name on the graph and give it a proper title.

“How do I graph my data – which quantity goes on the y-axis, and which on the x-axis?”

When you create a scatter plot, you will typically graph the *dependent* quantity along the y- (vertical) axis, and the *independent* quantity along the x- (horizontal) axis. In a typical experiment where a falling object covers a certain distance in each interval of time, the distance fallen depends on the time interval measured; in this case the distance will be plotted on the y-axis, and time along the x-axis.

There are other experiments we can perform where the time interval depends on the distance travelled; in physics we would still plot this with time along the x-axis. We usually plot graphs with the time along the horizontal axis since the equations derived from our theories are typically arranged this way.

Be sure to look at each data table carefully. The columns may not be presented in the same order!

3. Next you will determine the function that best fits your data. Select one of the following built-in functions from the Curve Fit menu: *Linear*, *Polynomial* (of 2nd order), *Exponential* (base *e*), *Logarithmic* (base 10) or *Power*. These functions are commonly used in scientific applications.

You will initially evaluate the quality of the fit by visual inspection: Does the function fit well across the entire span of your data, or only a portion of it? (Just the ends? The middle?) If more data were available, would the function predict a logical extension of that data? For the purpose of this exercise, you should check all five possible functions to ensure that you have chosen the correct one; don't forget to deselect a rejected fit before choosing another. You should also record any fit that was close, but that you decided to reject.

Compare the SSR value between two functions that fit equally well; the function with the lower SSR will be the correct one (at least for this exercise!) Recall that the SSR will appear on your graph *after* applying a user-defined function (in the next step.) *Remember that we don't use the R-value for our analysis!*

4. With the model of the best-fitting function chosen, you can now create your own “user-defined” fit as follows, so that you can calculate the uncertainties in the parameters.
 - a. Choose **Curve Fit** ⇒ **General** ⇒ **New Fit**. Note that this will keep the original fit on your graph, which will make it easier to ensure that you have created the user-defined fit correctly.
 - b. Click the **Define...** button and replace the contents with your equation. For example, a linear fit would be entered as follows:

$$\mathbf{a*x + b; a = 1; b = 1}$$

The statement “a = 1; b = 1” initializes ‘a’ and ‘b’ at unity, giving KaleidaGraph a place to begin its calculations. *Use 1 for your custom fit; don't use the values of ‘a’ and ‘b’ that were calculated using the built-in function from step 3!*

KaleidaGraph Notes (the **bold** text in the following examples show the notation used to give you the desired function):

Function	Example
Multiplication	b*x → b·x <i>Note that typing bx will give an error!</i>
Raise a value to a power	x^2 → x ² a^b → a ^b
Logarithm (base 10)	log (x) → log ₁₀ (x)
Exponential (base e)	exp (x) → e ^x

- The built-in functions of KaleidaGraph will use *M0*, *M1*, *M2*, etc. to represent variables. You will find it easier to use *a*, *b*, *c*, etc. when creating your custom fit functions.
 - You will always choose a second order fit when you select the built-in *polynomial* function (for work in this course). KaleidaGraph will correctly fit a quadratic and will give you the value of the correct parameters, but it will **always** display an expression representing a 9th order polynomial, the highest order that can fit by KaleidaGraph. *Don't be concerned.*
 - Don't worry if you notice a slight discrepancy between the calculated values of the parameters using the built-in functions compared to your user-defined function. The values should be of the same order if they don't match exactly.
- c. Once you have the function defined, *deselect the fit from step 3 above* so that only the user-defined fit remains on your graph.
5. Print your graph, but don't close it yet; you'll need it to make another plot in a moment.
 6. Record your model equation in your journal; since the data sets you are working with this week have measurements in distance (*d*) and time (*t*), your model equation should be written with *d* and *t*, not *x* and *y*! So, an appropriate *linear* model equation might be $d = a \cdot t + b$, where $a = __ \pm __ \{units\}$, etc. Be sure to state the parameters and their uncertainties with the correct number of significant figures, as well as the calculated SSR. Remember to include the appropriate units for each parameter!
 7. Select the graph just created and use KaleidaGraph to calculate the residuals of the fit (Refer to the *Calculating Residuals* section of "Graphing & Curve Analysis Using KaleidaGraph"). Then create a second graph of the residuals vs. time. Examining a residuals plot can reveal an underlying systematic error. *Be sure that you include the units of residuals on your graph axes!* Add your name and a proper title, and then print this graph as well.
 - *You do not need to print the data and residuals table for today's exercise.*
 8. Now do the same for *DataSet2* and *DataSet3*: Create a graph, determine the best-fit function, calculate the value of the parameters and SSR, and create a residuals graph.

Discussion

- Summarize your results by describing how you determined the function that fits each data set best. You should explain why you rejected another function that may have fit closely.
- Also include the 6 graphs that you printed: a graph of the data and the residuals for each of the three data sets. It should make sense that the title of each graph contains the data set number.
 - *If you read through these directions ahead of time, you would know that the data set number should be in each graph title!* If you have already printed your graphs, simply write the set number on the graph and be sure to thoroughly read the directions in future experiments.