Logger *Pro* Modeling, Fitting, and Linearization

Excerpt From: *Physics with Video Analysis (Appendix C)*
Authors: Priscilla Laws, Robert Teese, Maxine Willis and Patrick Cooney

When physicists compare theory with experiment, they usually consider a *physical model* of the situation. The Bohr model or quark model may be the first that come to mind, but in fact nearly every application of physics in introductory physics courses involves a simple model that is analogous to the real-world phenomenon. A moving object, for example, regardless of whether it is an atom, an elephant or a solar system, is often represented in the first approximation as a point mass with various forces applied to it. A real spring has mass, it can move in three dimensions and it is not ideal, but we model it as a straight-line, massless object that obeys Hooke's Law. A real capacitor may have internal resistance, frequency dependence and polarity, but we treat it as a pure capacitance. A cow rolling down a hill might be modeled as a sphere on an inclined plane as shown in Figure 1.

Figure 1: A spherical cow

Before we can compare a model to data, we have to turn the physical model into a *mathematical model* by applying the laws of physics. Newton's Second Law applied to a falling object, modeled as a point mass in a uniform gravitational field, leads to a simple differential equation. Its solution is the *analytic mathematical function*

$$y = \tfrac{1}{2} a_y (t - t_0)^2 + v_{y0}(t - t_0) + y_0 \quad (1)$$

where a_y is the acceleration of gravity, v_{y0} is the initial velocity, y_0 is the initial height and t_0 is the initial time. Applying Kirchoff's Laws to an LC circuit, modeled as a pure capacitance connected across a pure inductance, leads to a differential equation whose solution is the analytic mathematical function

$$i(t) = i_0 \cos(\omega(t - t_0) + \phi) \quad (2)$$

where i_0 is the current at $t = t_0$, ϕ is the phase angle and $\omega = \sqrt{1/LC}$ is the angular frequency.
An analytic equation contains variables and coefficients. In Equation 1, for example, the measured quantities, position and time (*y* and *t*), are variables whose functional form in the equation determines the basic shape of its graph. The numbers a_y, v_{y0}, y_0 and t_0 are the coefficients. They determine the size, orientation and position of the graph. Once the coefficients are determined, the model may be compared to the experimental data. This can be done visually by plotting the equation and the data on the same graph. In Logger *Pro* the Root Mean Square Error (RSME), a measure of the goodness of fit, can also be calculated.

There are several methods for determining the coefficients. The method that we consider the most versatile and educationally useful is to exploit the physical significance of each coefficient and determine its value from the data or other information about the situation. This method is called Analytic Mathematical Modeling. The Demon Drop activity has a good example of how it is carried out. In this activity, students use the Logger *Pro* video analysis tools to find height vs. time data for a falling cage of people in an amusement park ride. After thinking about the physical model, they call up the ***Model*** feature in the Logger *Pro* ***Analyze*** menu and choose the most appropriate general equation. This should be

$$y = \mathrm{A}t^2 + \mathrm{B}t + \mathrm{C} \tag{3}$$

They use the initial position y_0 to set C and the estimated initial velocity v_{y0} to set B. The theoretical acceleration $a_y = -9.8\ \mathrm{m/s^2}$ that can be calculated from Newton's gravitational constant and the radius of the earth is used to set the coefficient A = 0.5 a_y. They compare a graph of height vs. time from the model to an overlay graph of height vs. time from the experimental data. The agreement is good, and by making small changes in the estimates of y_0 and v_{y0} they can improve it.

A quick, easy substitute method of determining the coefficients is ***Curve Fitting***. With the advent of powerful computers in the classroom, curve fitting has become a widely used technique even though it has strong pedagogical disadvantages. In curve fitting, a standard least-squares algorithm is used to find the coefficients automatically. The main disadvantage of this method is that students do not have to think about the physical significance of the coefficients, so the method does not help them learn what the coefficients mean. Also, they can be deceived about the quality of the model. This is illustrated in the example of the next section.

Another substitute method is linearization. This method was very useful before computers were ubiquitous, or even today whenever people plot graphs on paper. The idea is to find a way to plot functions of the variables so the data points lie along a straight line, then determine the line's slope and offset by curve fitting. For example, the impedance of an RC circuit is

$$Z = \sqrt{R^2 + 1/\omega^2\mathrm{C}^2} \tag{4}$$

where Z and ω are the variables. The data can be linearized by plotting Z^2 versus $1/\omega^2$, since

$$Z^2 = \frac{1}{\mathrm{C}^2}\frac{1}{\omega^2} + \mathrm{R}^2 = m\frac{1}{\omega^2} + b\,. \tag{5}$$

where m and b are the slope and offset. The coefficients C and R are determined from m and b. The use of log paper or log-log paper is another example of linearization.

However, not all interesting functions can be linearized. If B is not zero, Equation 3 cannot be put in the form

$$g(y) = mf(t) + b \tag{6}$$

where $g(y)$ and $f(t)$ are functions that do not depend on the coefficients. Thus linearization cannot be used for free-fall motion unless $v_{y0} = 0$. It also cannot be used for trigonometric functions.

The three methods mentioned above, analytic mathematical modeling, curve fitting and linearization, are methods of comparing a physical model to experimental data by determining the coefficients in an analytic mathematical function. Another way of comparing a physical model to data is to skip the step of finding an analytical mathematical function (especially if it is not possible to find one). Instead, the differential equation of the model is solved numerically and its numerical solution is compared directly to the data. This method is called dynamical modeling. It can be carried out by transferring experimental data from Logger *Pro* into a spreadsheet or other software that can be set up for numerical integration and graphing.

An Example of Modeling and Fitting

The data shown in the graphs below come from the video analysis of a falling stack of ten coffee filters. Since the nested filters have not reached terminal velocity, it is reasonable to ask if the motion can be described by the kinematic equation that works for small, dense falling objects.

Analytic Mathematical Modeling: The investigators are told to see if the physical model of a point mass in a uniform gravitational field near the earth describes the motion using the modeling method. They use Equation 3, choosing A = –4.9 m/s^2 because it is half the acceleration of gravity, B = 0.00 m/s because the initial velocity is zero, and C = 2.06 m because the filters are released that high above the x-axis of the video. After making an overlay graph of the model on the data (Figure 2a), it is clear that this model does not fit the data well. Small changes in the coefficients do not improve the agreement significantly.

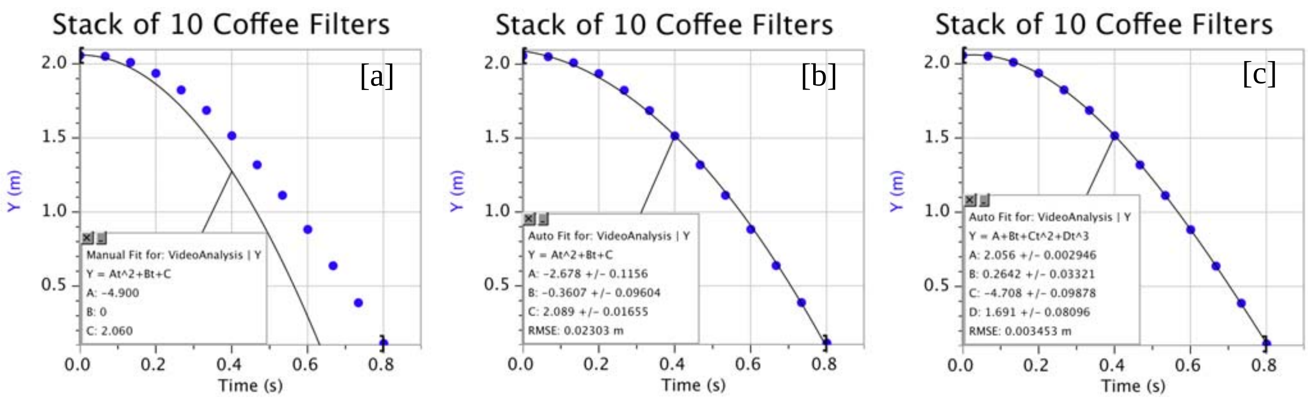

Figure 2: Three graphs of the same falling coffee filter data with different models or fits. [a] analytic mathematical modeling using Eq. 3. [b] a curve fit to Eq. 3. [c] a curve fit to a cubic polynomial.

Curve Fitting: The investigators are told to use curve fitting to see if the quadratic function for objects falling near the earth fits the data. They select the quadratic function under Curve Fit in Logger *Pro*. The result is in Figure 2b. It is a fairly good fit, with RSME = 0.03 m. Since the investigators followed the instructions and got a good fit, it may be hard to convince them that it is not a good model. The value of coefficient A gives an acceleration of 5.4 m/s^2, which is not the acceleration of gravity. The non-zero value of B implies that the coffee filters had a significant initial velocity in this model, even though they were actually dropped from rest. Moreover, this model does not predict that the filters will ever reach terminal velocity. The pitfall for investigators is that if they place too much value in getting a good fit, they are likely to overlook the reasonableness of the fitted coefficients or of the function itself.

Another pitfall occurs when the investigators are given the more general instruction to “find a simple function that fits the data well.” Without the guidance of a physical model, they will probably try the next higher order polynomial, a cubic function. The result is in Figure 2c. It is a much better fit than in Figure 2a, having RSME = 0.004 m. The fit line goes beautifully through all the dots representing data points. However, it is a terrible model – it has a local minimum because it is a cubic function, and it predicts that after another second the filters will begin falling upward! With an over-reliance on the mathematical goodness-of-fit, investigators may easily overlook the importance of understanding how the analytic function and its coefficients must relate to a physical model.

The previous example shows some pitfalls with using curve fitting. We feel that Analytic Mathematical Modeling is the most versatile and educational method for analyzing data that can be described by an analytic function. The reasons for this assertion are outlined in the following tables where we describe how Logger *Pro* software can be used for mathematical data analysis.

A practice we especially want to discourage is telling students to "find a function that fits the data." Until students have a robust sense of how and why functions must be connected to physical models, they are likely to choose functions that the instructor considers unsuitable. This practice also reinforces the idea that physics is just a collection of formulas for special cases rather than a general framework for analyzing nature.

Using Logger *Pro* for Mathematical Data Analysis

In the table below we outline specific procedures for three modeling methods, along with some Pros and Cons of each method. More detailed steps can be found using the Logger *Pro* Help Menu.

Analytic Mathematical Modeling	
Steps	1. Enter or record the data in Logger *Pro.* 2. Insert a graph of the two variables being analyzed (with ***Connect Points*** turned off under ***Graph Options***). 3. In the ***Analyze*** menu choose ***Model*** and then the analytic function you think will match. 4. Think about the physical meaning of each coefficient. Estimate the magnitudes of the coefficients based on what you know about the experiment. If a coefficient cannot be estimated, set it equal to 1. If the model graph does not match the data, either improve the estimates of the coefficients or choose a different model equation. 5. Close the window with the small graph you matched. This displays a larger overlay graph. If needed, make more changes in the coefficients.
Pros	1. Efficient modeling requires users to understand or learn the role each coefficient plays in providing information about a physical situation. 2. Modeling (or manual curve fitting) is the best way to learn how each coefficient in an equation affects the graph of the analytic function. 3. A skilled modeler can match sinusoidal functions efficiently in some cases that require repeated attempts at curve fitting. 4. Envelopes of functions can be seen visually and modeled. For example the local maxima of a damped oscillation can be modeled visually but not fit.
Cons	1. Modeling takes longer than automatic fitting, especially when there are coefficients that cannot be estimated. It is tedious in projects where many similar graphs of data must be analyzed graphically. 2. Modeling may not be as beneficial to users who already know how each coefficient in a particular function affects its graph. 3. Although modeling can give users uncertainty estimates, finding a minimum uncertainty is more tedious than it is for fitting.

Curve Fitting	
Steps	1. Enter or record data in Logger *Pro.* 2. Insert a graph of the two variables being analyzed (with ***Connect Points*** turned off under ***Graph Options***). Select the graph. 3. In the ***Analyze*** menu choose ***Curve Fit***, then the analytic function you think will match. 4. The coefficients will be calculated for you using a least squares analysis along with the Root Mean Square Error (RSME). 5. If the match is not good, try picking or entering another function (or repeat the fitting if you are matching a sine, cosine or tangent function). 6. Close the window with the small graph you matched. This displays a larger overlay graph. Make more changes on it, if needed.
Pros	1. Fitting usually takes less time than modeling, 2. It provides standard error analysis
Cons	1. Fitting does not help users focus on what shapes various analytic functions have. 2. Fitting does not help users learn how each coefficient in an equation affects the graph of the analytic function. 3. Sometimes when fitting many cycles of a sinusoidal function, fitting must be repeated several times to obtain a fit. 4. Envelopes of functions that can be seen visually cannot be fit without creating a new data set with local maxima or minima. For example the local maxima of a damped oscillation can be modeled but not fit without a lot of extra work.

Linearization	
Steps	1. If theory or the shape of a graph of y vs. x suggests that the equation linking these variables can be put in the form $g(y) = Af(x)+B$, then linearization is possible. For example, if the equation is $y = Ax^p$ where p is a non-zero number, then $g(y) = y$ and $f(x) = x^p$. 2. Enter or record the x, y data in Logger *Pro*. 3. Use ***New Calculated Column*** in Logger *Pro* to create new variables $g(y)$ and $f(x)$. 4. Plot $g(y)$ versus $f(x)$. If the graph is not linear try finding other functions $g(y)$ and/or $f(x)$. For example, if $f(x) = x^p$, try another value of p. 5. If the linearized data appear to lie along a line, then choose ***Linear Fit*** in the ***Analyze*** menu The coefficients A and B will be calculated for you using a Linear Regression analysis along with the Root Mean Square Error (RSME).
Pros	1. Many older scientists and teachers are comfortable with it. 2. If a function can be linearized and if no computer is available, then manual calculations and graphing can be performed more easily than for other methods. In particular, log or log-log paper might be useful. 3. The uncertainty in the linearized parameter is easy to obtain using a basic least squares analysis. However, it may be difficult to interpret.
Cons	1. A graph of linearized data is a straight line so the nature of the relationship between the original variables cannot be visualized. 2. Putting an equation in the form $g(y) = Af(x)+B$ may require a lot of cleverness. 3. Data that matches many interesting functional relationships cannot be easily linearized. Trigonometric functions, quadratic functions where the coefficient $B \neq 0$, and various other functions cannot be analyzed using linearization.