DATA ANALYSIS USING GRAPHS

Written by Zikri Yusof

Numerous times during the Introductory Physics Laboratory courses, you will be asked to graph your data as part of your analysis. Almost all of these graphs are straight-line graphs. Here, the major reasons why you are asked to produce these graphs are presented. We will also show a technique on how you can extract information out of straight-line graphs.

There are two major reasons why a graph is necessary when analyzing a set of data. First, it gives you a visual trend on the behavior of your data points. It is certainly easier to observe any pattern emerging from a set of data when it is graphed than by just simply staring at a bunch of numbers.

Secondly, it allows us to test a specific hypothesis or law. For example, the force exerted by a spring when it is extended or compressed from its natural length is described by Hooke's Law as

$$\mathbf{F} = - k\mathbf{x} \qquad (1)$$

where **F** is the force exerted by the spring, **x** is the extension or compression from its natural length, and k is the spring constant. If we consider just the magnitude of this expression, we have

$$F = kx \qquad (2)$$

In a typical experiment to test this law, one uses several masses hanged vertically on the spring, thereby obtaining several values of the extension x corresponding to the various values of the applied force F. Now the question is, what does one do after obtaining this data? With hindsight (this will be explained later), one can plot a graph of F versus x (F as the vertical axis and x as the horizontal axis) as shown in Figure 1.

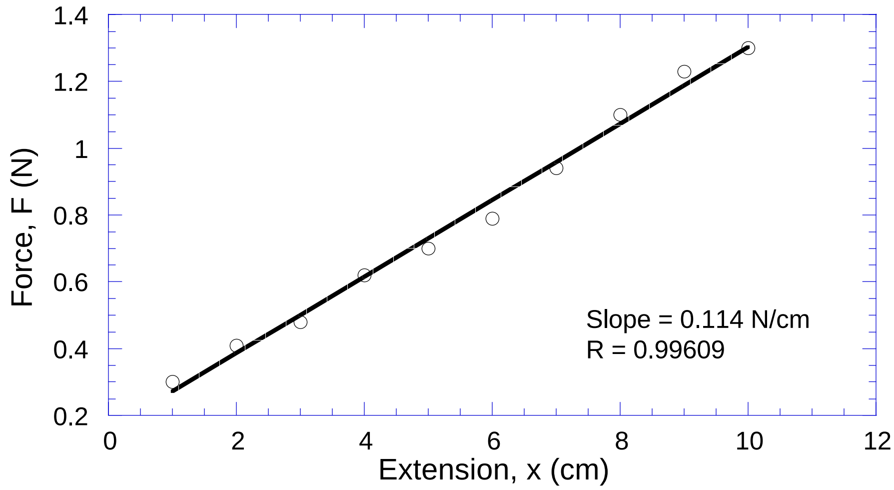

Figure 1: Experimental data of the force on a spring at various extension length (circles). The solid line represents the linear best-fit.

From Equation 2, if Hooke's Law is obeyed by this spring, then the force *F* applied to the spring is direction proportional to the extension of the spring. In Figure 1, simply by looking at the experimental data, one can immediately verify that this is true. Once we have verified that Hooke's Law applies in this case, then the value of the spring constant can be obtained from Equation 2, which is the slope of the best-fit line shown in the figure. Notice that this is just ONE single, smooth straight line, and not a "connect-the-dots" line. This line represents ALL of the data points. For our data shown, the spring constant k = 0.114 N/cm (notice that even slopes may have units).

A common "error" that students do is actually relevant here for discussion. Since we are interested in finding the spring constant k of the spring, why couldn't we just take one value of the extension x, and the corresponding value of F, and just simply use Equation 2 to find k? Why do we need to perform all the other measurements when just once is enough?

There are two ways to answer that question. First is in the technique of performing any experimental measurement. Each measurement that is made has in it errors which are unavoidable in most cases. These errors can be random, systematic, etc. Relying only on one measurement means that the result of our measurement can be highly uncertain. Furthermore, our one measurement may not reflect the true behavior of the system we are investigating. This is similar to making a survey of only one person to find the behavior or pattern for a whole community – it is just not an accurate technique for obtaining information.

The second reason why it is not acceptable is because using just one data point and Equation 2 to obtain *k* ASSUMES that Equation 2 is valid in the first place! The whole spirit of performing the experiment is to *verify* Equation 2, and *then* use it to obtain *k*. What if, instead of Figure 1, your experimental data shows the behavior in Figure 2? One can still obtain a linear fit for all the data points, but it is hard to convince oneself that the straight line shown is a good representation of the data points! It is likely that the spring used for this experiment has been deformed beyond its elastic limit, and thus its behavior no longer follows Hooke's Law. In this case, a straight line fit is not appropriate, and Equation 2 is not obeyed. Hence we cannot use it to obtain the slope and the spring constant *k*. If one perform just one measurement, there is no way one can see that this spring no longer obeys Hooke's Law. Simply plugging experimental numbers into Equation 2 does not confirm its validity.

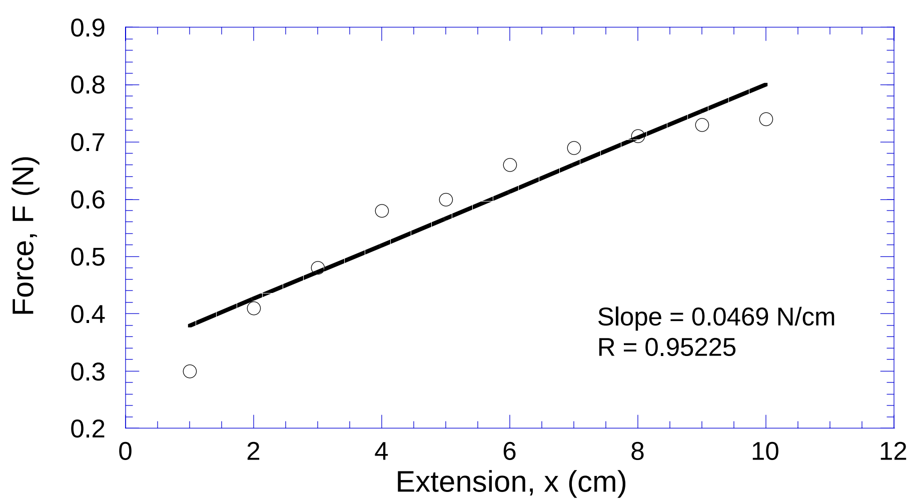

Figure 2: Another set of experimental data, but with a "bad" spring.

Notice that in both Figures 1 and 2, there is an additional number given associated with the quantity *R*. In performing a best-fit routine, one always needs to know how good of a fit the curve is to the data being plotted. Depending on the numerical routine used (typically it is the least-square method), there is a quantity usually known as the "goodness-of-fit". This is sometime known as R, or R^2. This value tells you *quantitatively* how well the best-fit line represents the data point. The closer R is to 1.0, the better the fit. Notice that the value of R in Figure 1 is closer to 1 than the value of R in Figure 2. This confirms our visual inspection earlier in deducing that the data in Figure 2 does not fit into a straight line as well as the data in Figure 1. When performing any kind of fit, be it a straight line or any other curves, it is imperative that one knows quantitatively how well the fit is to the data. One can fit any curve to any set of data. However, it does not mean that the fit is good, or even valid! Knowing the goodness-of-fit factor allows one deduce if this fit is "good enough".

WHAT CAN WE LEARN FROM A LINEAR LINE FIT?

In the previous section, the slope of the best-fit line from an F versus x graph corresponds to the spring constant k. How do we know this? And how can we apply to a more complex situation?

To answer that question, you have to recall your knowledge about straight lines in a simple x-y graphs. From algebra and geometry, the equation for a straight line in cartesian coordinate system is given by

$$y = mx + b \tag{3}$$

where y is the ordinates, x is the abscissa, m is the slope of the line, and b is the y-intercept. If we have a set of values of (x,y), then the values of y are plotted on the vertical axis, while the values of x are plotted on the horizontal axis.

In the Hooke's Law experiment, you are given that for a spring still within its elastic limit, the force F and the extension of the spring follows the simple relationship as

$$F = kx \tag{2}$$

which is the old Equation 2.

Now, if we plot F on the vertical axis and x on the horizontal axis, we can make a *direct comparison* between Equations 2 and 3. We notice that

$$y \to F$$

$$x \to x$$

$$m \to k$$

$$c \to 0$$

Simply by observation, we can deduce that the slope of our F versus x best-fit line should correspond to the spring constant of the spring, if we end up with a good straight-line fit.

This is a rather trivial example, so let us look at a more complex relationship. For the thin lens equation, we are given the expression

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i} \tag{4}$$

where f is the focal length of the lens, o is the distance of the object from the lens, and i is the distance of the focused image from the lens. In this experiment, you measure o and i, and the objective here is to find the focal length f of the lens. So how should we plot the data to be able to obtain f? Let us try a systematic approach to this.

1. Look at the theoretical expression (Equation 4), and decide what you want to plot on the vertical axis, and what you want to plot on the horizontal axis. In some cases, you may need to do a little algebraic manipulation. In this case, let us plot $1/o$ versus $1/i$ (i.e. $1/o$ on the vertical axis, and $1/i$ on the horizontal axis).
2. Once you have decided what to plot, then make a direct comparison with the x and y axes from Equation 3, i.e.

$$y \to 1/o$$

$$x \to 1/i$$

3. Rearrange the theoretical expression to follow Equation 3. In this case, we have

$$\frac{1}{o} = -\frac{1}{i} + \frac{1}{f}. \qquad (5)$$

4. Look at the rearranged equation (Equation 5) and make a direct comparison again to Equation 3. We now have

$y \rightarrow 1/o$

$x \rightarrow 1/i$

$m \rightarrow -1$

$c \rightarrow 1/f$

Notice that the expression corresponding to *m* and *c* are both **<u>constants</u>** for this set of data. The expression for the slope and the *y*-intercept should *not* contain any variable within that data set!

5. Use the comparison above to obtain the relevant information. In this case, plotting 1/*o* versus 1/*i* will result in (if the thin lens equation is obeyed) a straight-line. The best-fit line should have a slope of –1, while the *y*-intercept should correspond to 1/*f*. Thus using the *y*-intercept value from the graph, the value of the focal length of the lens can be obtained.

Sometimes the theoretical expression given can seem to be rather complicated. Do not be afraid of these. If you follow the technique given above, it should be straightforward to derive the necessary expression for the slope and the *y*-intercept of your line. For example, you will encounter the expression

$$V(t) = V_0 e^{-t/\tau} \qquad (6)$$

where the variables that you measure in the experiment are $V(t)$ at various t. Now, no matter what you plot, you will never end up with a straight-line fit, because $V(t)$ has an exponential relationship to t. However, we can perform some algebraic manipulation as suggested in Procedure 1 above. Taking the natural logarithm of Equation 6 and rearranging the expression (you should do this yourself), we obtain

$$\ln[V(t)] = -\frac{t}{\tau} + \ln V_0 \,. \tag{7}$$

If we plot $\ln[V(t)]$ versus t, we will then obtain a straight line graph with expressions for the slope as $-1/t$ and y-intercept as $\ln V_0$. Notice that $-1/t$ and $\ln V_0$ are constants as required earlier.

You will have plenty of opportunity to practice this skill throughout the introductory physics laboratory session, since this procedure is usually required as part of the theory section in your laboratory report.