4.1. Validating the Model–I: How Do We Know the Model Is OK?

There are two issues that arise when we speak of the validity or correctness of a model. The more obvious one is whether or not the model can predict the measured or observed behavior of whatever object or device is being modeled. Thus, if we are modeling the period of the oscillations of a pendulum, we could reasonably expect that changes in the pendulum length would produce oscillations at correspondingly different periods or frequencies. As we see from eq. (2.2), if we double the length *l* of a pendulum, we would expect its period to increase by about 41% . Similarly, were we doing pendulum experiments on the moon, we would expect to see an increase in the period of about 145%. These predictions of the pendulum's behavior are confirmed by the available experimental data, and so the model is validated. Alternatively, given empirical data without an underlying theory, we could construct a model to explain the empirical data—although it is also quite likely that the (new) model or theory would be further tested by making predictions about experiments as yet undone or measurements as yet untaken.

(We note parenthetically that the measurement [and containment] of experimental error is a complex subject that is closely linked to the field or discipline in which the experiment is intellectually housed. However, there are some fundamental ideas about error and about statistics that apply generally, and we will introduce them in Sections 4.2–4.3.)

The less obvious question about model validity is concerned with the inherent consistency and validity of themodel. If we hark back to themodeling meta-principles outlined in Section 1.2, we see issues and questions that pertain directly to model validation. For example, have we identified the right governing principles? Have we used the right equations? And, is the model consistent with its principles and assumptions? The first two of these questions are about ensuring that we apply the proper principles and formulations when we try to find what we are seeking. Again, when modeling the pendulum, our basic principles are Newton's law of motion, and our assumptions will depend on whether we are anticipating small angles of oscillation or large. As we will see in Lecture 5, a linear equation of motion suffices in the former case, while a complete nonlinear formulation is needed for the latter (large oscillations).

4.1.1. Checking Dimensions and Units

There are several checks or tests we can bring into play while we build models and approximate the mathematics. The first is the application of the principle of dimensional homogeneity (cf. Lecture 2), which requires that each term in an equation has the same net dimensions. For example, the stiffness or spring constant of a cantilever beam, *k*, can be written in terms of the beam's length, *L*, second moment of its cross-sectional area (commonly but erroneously called the "moment of inertia"), *I* , and modulus, *E*, as:

$$k = \frac{3EI}{L^3}. \qquad (4.1)$$

The physical dimensions of the parameters in eq. (4.1) are **F/L** for the spring constant, **L** for the beam length, $\mathbf{L^4}$ for *I* , and $\mathbf{F/L^2}$ for the modulus. Thus, we can apply the principle of dimensional homogeneity to ensure that eq. (4.1) has the correct dimensions and is dimensionally consistent:

$$[k] = \left(\frac{\mathrm{F}}{\mathrm{L}}\right) = \left[\frac{3EI}{L^3}\right] = \frac{1 \times (\mathrm{F}/\mathrm{L}^2) \times \mathrm{L}^4}{\mathrm{L}^3} = \left(\frac{\mathrm{F}}{\mathrm{L}}\right). \qquad (4.2)$$

If the dimensions of all the terms in an equation or model are not known, as is sometimes the case, then the principle of dimensional homogeneity can be applied to properly determine the dimensions of the unknown quantity. In the case of the cantilever beam, if we didn't know the

dimensions of I , we would solve eq. (4.1) for I and then apply the principle of dimensional homogeneity again:

$$[I] = \left[\frac{kL^3}{3E}\right] = \frac{(\mathrm{F/L})\mathrm{L}^3}{\mathrm{F/L}^2} = \mathrm{L}^4. \qquad (4.3)$$

We can also take the principle of dimensional homogeneity one step further and use it as a guiding principle for checking the specific *units* used in a numerical calculation. If we measured the properties of a particular cantilever beam, say a standard (12 in) steel ruler to be used in a classroom project, we would find

$$E = 2.05 \times 10^2 \mathrm{GPa},$$
$$I = 6.78 \times 10^{-5} \mathrm{cm}^4, \qquad (4.4)$$
$$L = 2.81 \times 10^{-1} \mathrm{m}.$$

If we substitute these values into eq. (4.1), we see immediately that we have a mismatch of units:

$$k = \frac{3(2.05 \times 10^2 \mathrm{GPa})(6.78 \times 10^{-5} \mathrm{cm}^4)}{(2.81 \times 10^{-1} \mathrm{m})^3}. \qquad (4.5)$$

The units' mismatch is easily rectified if we use proper unit conversions, that is,

$$k = \frac{3\left[2.05 \times 10^2 \times 10^9 \ \mathrm{Pa}\left(\frac{\mathrm{N/m}^2}{\mathrm{Pa}}\right)\right]\left[6.78 \times 10^{-5} \mathrm{cm}^4 \left(\frac{\mathrm{m}}{10^2 \mathrm{cm}}\right)^4\right]}{(2.81 \times 10^{-1} \mathrm{m})^3}. \qquad (4.6)$$

or

$$k = \frac{3[2.05 \times 10^{11} \ \mathrm{N/m}^2][6.78 \times 10^{-13} \mathrm{m}^4]}{(2.81 \times 10^{-1} \mathrm{m})^3} = 1.88 \times 10^1 \ \mathrm{N/m}. \qquad (4.7)$$

Two final notes here. First, it is generally a better strategy towrite all of the data to be used in the same systemof units at the beginning of a calculation as this reduces the chance for error. Thus, here we could have converted the units immediately after the measurements were taken. Second, note that we have used scientific notation in both writing the measurements and performing the arithmetic. Thus, there can be no doubt about the number of significant figures in the answer (4.7).

4.1.2. Checking Qualitative and Limit Behavior

Model validation is integral to the modeling process. Models are validated by having their predictions confirmed experimentally, or statistically, or by some other quantitative means. In both our physical and mathematical reasoning we must make explicit our assumptions and their limits, and we must ensure that our mathematics does indeed reflect the physics we are modeling. In addition to looking at numbers, the mathematical behavior should "feel right" in qualitative terms. We did just such qualitative analysis at the beginning of this section when we described the expected behavior of the pendulum as a function of its length, l . Similarly, as also indicated by eq. (2.2), it feels intuitively right that pendulums will swing faster and have shorter periods in stronger gravitational fields. Thus, when we are constructing mathematical models, and especially when we are making mathematical approximations, we need to take care that we are admitting mathematical behaviors that are qualitatively appropriate.

Still another example of such reasoning is available from our just completed dimensional check of the stiffness of a beam. Here we rewrite eq. (4.1) in a form that explicitly identifies the physical meaning of each parameter that appears in the equation:

$$(k = \text{beam stiffness}) \propto \frac{(E = \text{material stiffness})(I = \text{cross} - \text{sectional 2nd moment})}{(L = \text{beam length})^3}. \qquad (4.8)$$

Equation (4.8) can be viewed through the eyes of a structural engineer talking about the meaning of its mathematical version, eq. (4.1). It supports the engineer's intuitions as follows. It stands to reason that the beam's stiffness is proportional to the material stiffness, that is, it increases or decreases as does *E*. The beam's stiffness is also proportional to the second moment of the beam's cross-section, *I* . It also is intuitively pleasing that the stiffness is inversely dependent on the length, so that the beam's stiffness increases as *L* becomes very small and decreases as *L* becomes very large. Finally, if we look at the limiting cases of each parameter decreasing to zero or becoming indefinitely large, we would see that each of the trends exhibited by eq. (4.8) is consistent with the reasoning just outlined, as well as with our practical experience of beams in the real world.

Reasoning about the way that variables appear in equations is of second nature in mathematical modeling, and we will have many opportunities to invoke such reasoning in the discussions of applications that follow. One simple example is afforded by the fundamental frequency of free vibration of a cantilever beam, ω, of mass density, ρ, and cross-sectional area, *A*, with a mass, *m*, at its tip. That frequency is, approximately,

$$\omega \cong \sqrt{\frac{3EI/L^3}{\rho AL(1 + m/\rho AL)}}. \qquad (4.9)$$

Does eq. (4.9) exhibit the right qualitative and limit behavior? It does. It reduces to a well-known result for a cantilever beam when the tip mass, *m*, vanishes, and eq. (4.9) correctly describes the frequency of a mass-less beam with a tip at its end when that tip mass gets so large that it dominates the beam mass.

It may seem that much of what has been said in this section is *common sense*. It is, as long as it is commonly applied! To invert a popular saying, "If we expect our model to be a duck, then it should look like a duck, walk like a duck, and quack like a duck."

4.2. Validating the Model–II: How Large Are the Errors?

Building mathematical models means using numbers derived from experimental or empirical data, or from analytical or computer-based calculations. Errors are thus always present, whether due to data reading or data manipulation. Since error is always present, we turn now to a discussion of error and statistics—the way we deal with error.

4.2.1. Error

Error is defined as the difference between a measured (or calculated) value and its true or exact value. Error is *always* present. How much error is present depends on how skillfully the data is read or manipulated. Therefore, error analysis should be a part of every modeling process.

There are two types of error. *Systematic error* occurs whenever an observed or calculated value deviates from the true value in a consistent way. Systematic error occurs in experimentswhen instruments are improperly calibrated because their output varies during use. Thus, instruments must be properly calibrated before an experiment is run and before data is measured and recorded. Improper calibration affects both analog and digital data recorders, although analog displays are also subject to other kinds of systematic error, such as a bent needle on ameter face such as that shown in Figure 3.3.(Lecture 3) Systematic error also affects calculations, although this is more controllable as it is likely due to using incorrect values of "known" variables or to improper control of the number of significant figures retained during the calculation process.

Random errors are, not surprisingly, due to chance. They arise largely in experimental work because unpredictable things happen and because not everything in an experimental set-up is known with complete certainty: Connections can be loose or break altogether, dirt may get into a sensitive moving part, or the amount of friction present in a moving part may not be controllable. The resulting random error varies in both magnitude and sign. The laws of statistics help us to describe and account for the distribution of such random errors. Indeed, it has been said that randomness is a mathematical model for variability that cannot be explained in a deterministic way.

The *absolute error* is defined as the difference between the true or expected value, X_e , and the measured value, X_m, that is, as $X_e - X_m$. The true value, X_e , may be known or it may have an expected value based on a calculation or some other data source. The relative error is the absolute error divided by the measured value, that is, $(X_e - X_m)/X_m$.

The statistic found most useful is the percentage error, which is the percentage-based relative error:

$$\%\text{error} = (100)\frac{(X_e - X_m)}{X_m}. \qquad (4.10)$$

For example, suppose that an ammeter has a systematic error of +2A (amperes) because of either a bent needle (analog) or improper calibration (digital or analog).When the display reads 100 A the percentage error is

$$\%\text{error} = (100)\frac{(102 - 100)}{100} = 2\%$$

while if the same ammeter reads 20 A the percentage error is

$$\%\text{error} = (100)\frac{(22 - 20)}{20} = 10\%$$

The percentage error is much larger in this instance, providing another example of how scale affects results!

Similarly, errors are introduced when series expansions are truncated (cf. Lecture 3). For example, for $\theta = \pi/12(15^\circ)$, the percentage error incurred by replacing sin x with x is:

$$\%\text{error} = (100)\frac{(\sin \pi/12 - \pi/12)}{\pi/12} = -1.14\%$$

Note that *errors* and *mistakes* are not the same thing. Errors are defined as the difference between a true or expected value and a measured (or calculated) value. Further, as we discussed above, some error is unavoidable. On the other hand, *mistakes* are blunders made by the person doing the experiment (or analysis or calculation). Blunders are made by reading or recording erroneous data, using instruments inappropriately (e.g., improperly calibrated instruments, inadequately sensitive meters), using he wrong formulas, using inconsistent or wrong units, and so on. These kinds of mistakes can—and obviously should—be avoided.

4.2.2. Accuracy and Precision

Since we have to contend with systematic and random errors, as well as with the hopefully rare mistake, it is important that we be able to estimate the effects of these errors and mistakes. *Accuracy* is defined as a representation of how close a measured or calculated value is to an established or true value. In experimental work, accuracy is usually expressed as a percentage of the maximum scale value. Thus, voltages read on a 100V scale with an accuracy of 5% are accurate to within ±5V.

Precision is defined in terms of the ability to reproduce a set of data with a specified accuracy. The more precise a set of readings or calculations, the closer the individual readings or calculations are *to each other*. Thus, suppose we measured an input voltage that is known to be 50V with the voltmeter having an accuracy of 5%. Five individual readings are taken and recorded as, respectively, 54, 53, 55, 53, and 55V. These clearly fall within the meter's accuracy bounds of ±5 V. Since the average or mean reading of the five readings is 54V, and since the maximum deviation from this mean of any one of the measurements is 1 V, the precision of the five measurements is determined to be ±1% (remember that the meter has a 100V scale). As we illustrate in Figure 4.1, our little virtual experiment has produced precise but relatively inaccurate readings.

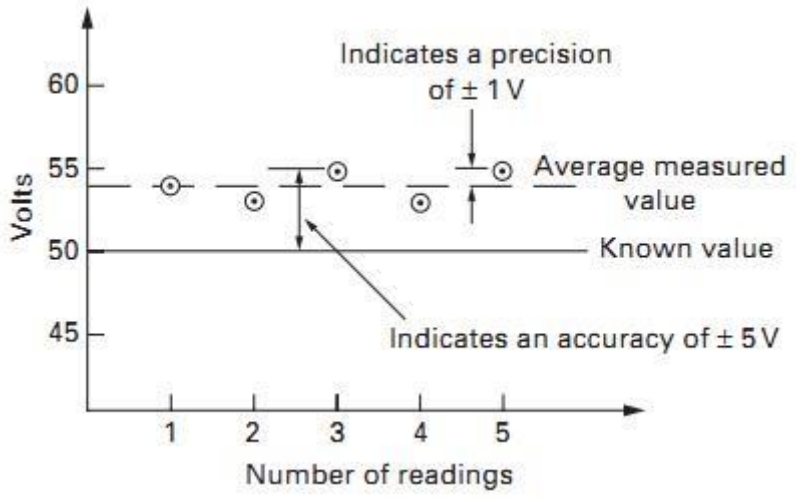

Figure 4.1. Some (made-up) experimental data that illustrates: *accuracy*, the closeness of the measured value to an established value, and *precisio*n the ability to reproduce a set of measurements within a specified accuracy. These data reflect measurements that are rather precise, yet relatively inaccurate.

It is worth noting that the accuracy of a measuring device is controlled by its *sensitivity* because it is the *sensitivity* that identifies the minimum amount of change that the device can detect and indicate. Suppose we wanted to measure very small voltages, say less than 1 millivolt (mV). Our trusty voltmeter allows us to choose one of three measurement ranges: 0–50V, 0–2.5V, or 0–5mV. With either of the first two ranges we will see no reading at all. However, with the third scale, 0–5mV, there will be a noticeable measurement that can be recorded. Thus, moving from either of the first two scales to the third produces a more sensitive voltmeter, and so our readings will be more accurate. Hence, we see how scale influences sensitivity and, therefore, accuracy.

4.3. Fitting Curves to Data

Graphical presentations of calculations and experimental results are the most convenient—and often the most informative—presentation of data available.We can spot trends, identify discontinuities, and generally get an intuitive "feel" for what the data "says" when we look at plots or curves. Given this very human proclivity, how do we draw curves for a given collection of points? That is, since plotted data points rarely align themselves perfectly on a known or identifiable curve, howdowe fit a curve through them? Still further, how do we generate the "best fit" of a curve through the data?

The short answer to these questions is in a familiar spirit: It depends on what you want. If the accuracy of the curve is not too important, and if we're only looking for a rough, qualitative idea of how one variable depends on another, then we can draw the curve "by eye." That is, we draw a smooth curve that seems to go through the plotted data points with an eye to perhaps "distributing" the data in roughly equal amounts above and below the curve drawn, as we have done in Figure 4.2.

Often, greater accuracy is desirable, as when we want to *interpolate* to obtain values between measured values, or even more so when we want to *extrapolate* to estimate values beyond the range of the measured values. Extrapolation can easily magnify errors in the estimated values, so that greater accuracy is quite important. Further, extrapolation ismost accurate when the curve drawn is a straight line.

The *method of least squares* is the most commonly used approach to obtaining a best straight line through a series of points. It assumes that all of the *scatter*, the variation of the data from the drawn curve, derives from error in measuring one of the variables. That variable is chosen as the ordinate for the axes on which the straight line will be plotted. Then the best-fit straight line is the one that has the minimum errors in the ordinate.

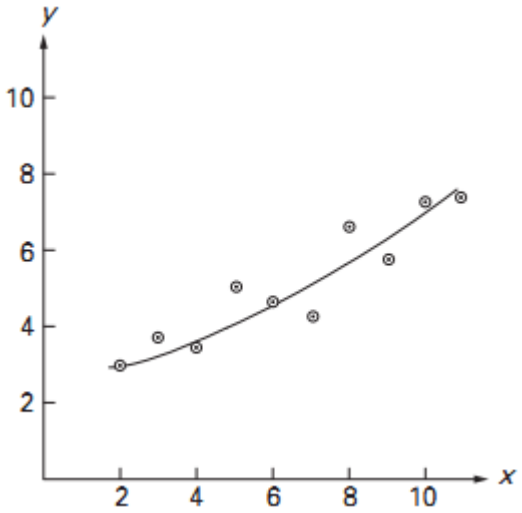

Figure 4.2. A best-fit curve that is drawn by hand using visual estimation (i.e., "drawn by eye").

We are thus looking for an equation of the usual form

$$y = mx + b, \qquad (4.11)$$

where b is the y-intercept with [b]=[y], and m is the slope with [m]= [y/x]. We first define the error in each reading as the difference in the ordinate between the measured value, y_i , and the straight line's ordinate, $(mx_i + b)$, for all values of the abscissa, x_i :

$$E_{y_i} = y_i - (mx + b). \qquad (4.12)$$

We define a measure S of the total error as the sum of the square of the errors at every point on the abscissa, x_i , where values of the ordinate, y_i , are given, that is, as

$$S = \sum_{i=1}^{n} \left(E_{y_i}\right)^2 = \sum_{i=1}^{n} [y_i - (mx + b)]^2. \qquad (4.13)$$

The minimum of the measure of the total error is then found by differentiating S with respect to m and b and so determining the values of m and b needed to plot eq. (4.11):

$$\frac{\partial S}{\partial m} = 2\sum_{i=1}^{n} [(y_i - mx_i - b)(-x_i)] = -2\sum_{i=1}^{n} x_i y_i + 2m\sum_{i=1}^{n} x_i^2 + 2b\sum_{i=1}^{n} x_i = 0, \qquad (4.14)$$

$$\frac{\partial S}{\partial b} = 2\sum_{i=1}^{n}[(y_i - mx_i - b)(-1)] = -2\sum_{i=1}^{n} y_i + 2m\sum_{i=1}^{n} x_i + 2bn = 0. \qquad (4.15)$$

Equations (4.14) and (4.15) are a pair of linear algebraic equation that can be solved to yield the following values of *m* and *b*:

$$m = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}. \qquad (4.16)$$

and

$$b = \frac{(\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i) - (\sum_{i=1}^{n} x_i y_i)(\sum_{i=1}^{n} x_i)}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}. \qquad (4.17)$$

Note that eqs. (4.16) and (4.17) have different physical dimensions that depend on the particular physical problem being modeled.

Consider now the data displayed in the first two columns of Table 4.1, which are the result of another, virtual experiment. We will now determine the best straight line that can be drawn through the data. First, we calculate the products shown in the third and fourth columns of Table 4.3. Then we sum all four columns to find the data in the last row of the table, which are then substituted into eqs. (4.16) and (4.17) to find $m = 0.85$ and $b = 1.26$. The best straight-line fit through the data of Table 4.1 is, then,

$$y = 0.85x + 1.26, \qquad (4.18)$$

Equation (4.18) is plotted in Figure 4.3, together with the data from Table 4.3, and we see that the straight line seems to fit the data pretty well. Can we characterize the *quality* of that fit, that is, just how well does eq. (4.18) fit the given data? The quality of fit is expressed in terms of R^2, called "*R* squared," which describes how well a curve *regresses* toward the

Table 4.1. A table of data from a virtual experiment used to calculate the best-fit strainght line approximation shown in Figure 4.3.

i	x_i	y_i	$x_i y_i$	x_i^2
1	0	1.0	0	0
2	1.0	2.1	2.1	1.0
3	2.0	2.8	5.6	4.0
4	3.0	3.6	10.8	9.0
5	4.0	5.0	20.0	16.0
6	5.0	5.5	27.5	25.0
7	6.0	8.	48.0	36.0
8	7.0	6.4	44.8	49.0
9	8.0	7.4	59.2	64.0
	$\sum_{i=1}^{9} x_i = 36.0$	$\sum_{i=1}^{9} y_i = 41.8$	$\sum_{i=1}^{n} x_i y_i = 218.0$	$\sum_{i=1}^{n} x_i^2 = 204.0$

data from which it was derived. R^2 is a number between 0, which indicates no fit at all, and 1, which describes a perfect fit. (There are many mathematical and statistical computational packages that include the formulas needed to calculate R^2.)

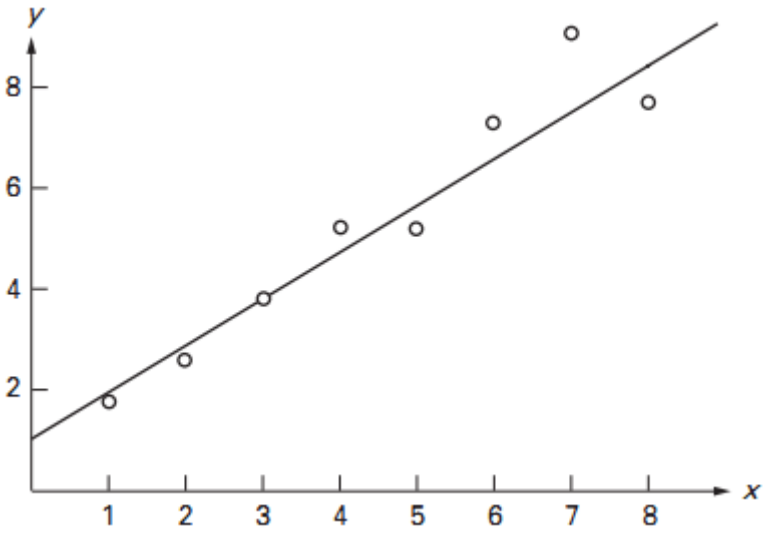

Figure 4.3. A best-fit straight line for the data in Table 4.1 produced by least squares. It is analytically represented as y=0.85x+1.26.