Error Analysis and Graphs

Review material taken from the introduction to Experiment 1 in the PHYS 200 Lab Guide

The goal of physics is to understand the world around us, and discover the laws of nature. Therefore, any theory (no matter how fancy or attractive it might be) is actually useless, unless it is supported by experimental evidence. Doing an experiment usually involves making quantitative measurements. However, in order for the measured values (or data) to be meaningful it is very important to understand the limitations of the instruments used and also recognize the possible sources of error.

Significant Figures and Measurement Uncertainty

Assume that you wanted to find the area and volume of your physics textbook. To do this, of course, you needed its dimensions (length, width and thickness). Using a ruler, you first measured the length of the textbook. Since the smallest division on the ruler is the millimeter you could only give an upper limit (26.3 cm) and a lower limit (26.1 cm) for the length. Due to the limitations of the instrument (the ruler in this case) you could then only say that the correct length (L) of the textbook, most probably, lies between these two values. This is written as ($L = 26.2 \pm 0.1$ cm), where the first number is called the measured value of L and the second number is the called uncertainty (or error) in the measurement. Similarly, you measured the width (W) and thickness (T) of the textbook to be ($W = 20.6 \pm 0.1$ cm) and ($T = 3.9 \pm 0.1$ cm) respectively. In these measurements, the first digit after the decimal point is uncertain, making it meaningless (or insignificant) to include any digits beyond that. The digits in a measured value, up to and including the first uncertain digit, are called significant figures. So, there are 3 significant figures in L and W and 2 significant figures in T. Note that we cannot write the length as $L = 26.20$ cm, since it will then include two uncertain digits.

In counting the number of significant figures, in a measurement, we start from the first non-zero digit and end by the first uncertain digit. If the uncertainty of a measurement is known, then it becomes easy to specify the uncertain digit and count the number of significant figures. However, if the uncertainty of a measured value is not known, then we have to be careful. If there is no decimal point in the number, then the last non-zero digit is considered to be the first uncertain digit. If the number contains a decimal point then the last digit after the decimal

Table 1.1

Measured Value	Scientific Notation	Number of Significant Figures
9.80	9.80	3
120	1.2×10^{2}	2
120 ± 2	$(1.20 \pm 0.02)\times10^{2}$	3
120.0	1.200×10^{2}	4
1.00560	1.00560	6
0.00560	5.60×10^{-3}	3
375×10^{-9}	3.75×10^{-7}	3

(whether or not it is a zero) is assumed uncertain. This ambiguity in the number of significant figures is avoided by using the scientific notation, in which only one digit is kept to the left of the decimal point, while the remaining digits, up to an including the uncertain one, are moved to the right. Of course, we have to multiply by an appropriate exponent. See Table 1.1 for examples.

Limit Errors

Knowledge of measurement errors and how they combine is very important in understanding how much we can trust an experimental result. There are two major types of measuring error, random and systematic. For example, the error in the exercise above where you had trouble reading the ruler to better than 1 mm, is a random error. If, on the other hand, you were using a ruler marked in tenths of an inch and you thought it was in mm, that would give a systematic error of scale.

Going back to your exercise, the surface area (*A*) of the textbook is calculated using the equation ($A = L \times W$). If you plug the measured values of length and width into the calculator you will get $A = 539.72\ \text{cm}^2$. However, are all these digits significant? How many significant figures are there in *A*, and what is its uncertainty? As a general rule for combining numbers *by multiplication or division* the final result should not have more significant figures than the original value with the least number of significant figures. For *addition and subtraction*, however, the final result should not have more decimal places than the original value with the least number of decimal places. In this case since both *L* and *W* have three significant figures, *A* should be rounded off to 3 significant figures, and should be written as $A = 540\ \text{cm}^2$ (or $5.40 \times 10^2 \text{cm}^2$). Similarly, the volume of the textbook is written as $V = 2100\ \text{cm}^3$ cm^3 (or $2.1 \times 10^3 \text{cm}^3$), which should not contain more than two significant figures.

What about the uncertainties of *A* and *V*? Since the area and the volume of the textbook are calculated from the measured dimensions, the errors in *A* and *V* are propagation of the errors of *L*, *W* and *H*. Table 1.2, lists the rules used to calculate the propagated errors in various arithmetic operations. Notice that, whether we are adding or subtracting numbers, the combined error simply adds up. Similarly, we use the same rule for combining error whether the operation is multiplication or division. Notice that in the last rule, *k* is an exact constant factor that does not have uncertainty (i.e. $\Delta k = 0$).

For example, the uncertainty of the textbook area is given by

$$\Delta A = |A| \left(\frac{\Delta L}{L} + \frac{\Delta W}{W} \right) = 540 \ \left(\frac{0.1}{26.2} + \frac{0.1}{20.6} \right) = 4.7\ \text{cm}^2,$$

which can be rounded to $\Delta A = 5\ \text{cm}^2$, because it is not an actual measurement. So, we see that the third digit in *A* is uncertain indicating that the number of significant figures in *A* is indeed 3. The final answer is then written as $A = (5.40 \pm 0.05) \times 10^2 \text{cm}^2$. Similarly the volume of the textbook is written as $V = (2.1 \pm 0.1) \times 10^3 \text{cm}^3$. It is often useful to speak of the relative error $\Delta x / x$ to compare the size of error to the size of the value measured. Another advantage is that the

Table 1.2

Relation	Equation	Limit error
Addition	$z = x + y$	$\Delta z = \Delta x + \Delta y$
Subtraction	$z = x - y$	$\Delta z = \Delta x + \Delta y$
Multiplication	$z = x\ y$	$\Delta z = \lvert z \rvert \left(\frac{\Delta x}{\lvert x \rvert} + \frac{\Delta y}{\lvert y \rvert} \right)$
Division	$z = x / y$	$\Delta z = \lvert z \rvert \left(\frac{\Delta x}{\lvert x \rvert} + \frac{\Delta y}{\lvert y \rvert} \right)$
Power	$z = x^n$	$\Delta z = n\ x^{n-1} \Delta x$
Exact constant factor	$z = k\ x$	$\Delta z = k\ \Delta x$

relative error of a product or dividend is just the sum of the relative errors of what make them up.

Example: Derive a formula for the uncertainty in the volume of a cylinder with radius *r* and height *h*?

Solution: The volume (v) of the cylinder is given by $v = \pi\ r^2 h$. Using the rules in Table 1.2, the uncertainty in v is given by

$$\begin{aligned} \Delta v &= \pi\ \Delta(r^2 h) \\ &= \pi\ (r^2 h)\left[\frac{\Delta r^2}{r^2} + \frac{\Delta h}{h}\right] \\ &= \pi\ (r^2 h)\left[\frac{2r\Delta r}{r^2} + \frac{\Delta h}{h}\right] \\ &= \pi\ (r^2 h)\left[\frac{2\Delta r}{r} + \frac{\Delta h}{h}\right]. \end{aligned}$$

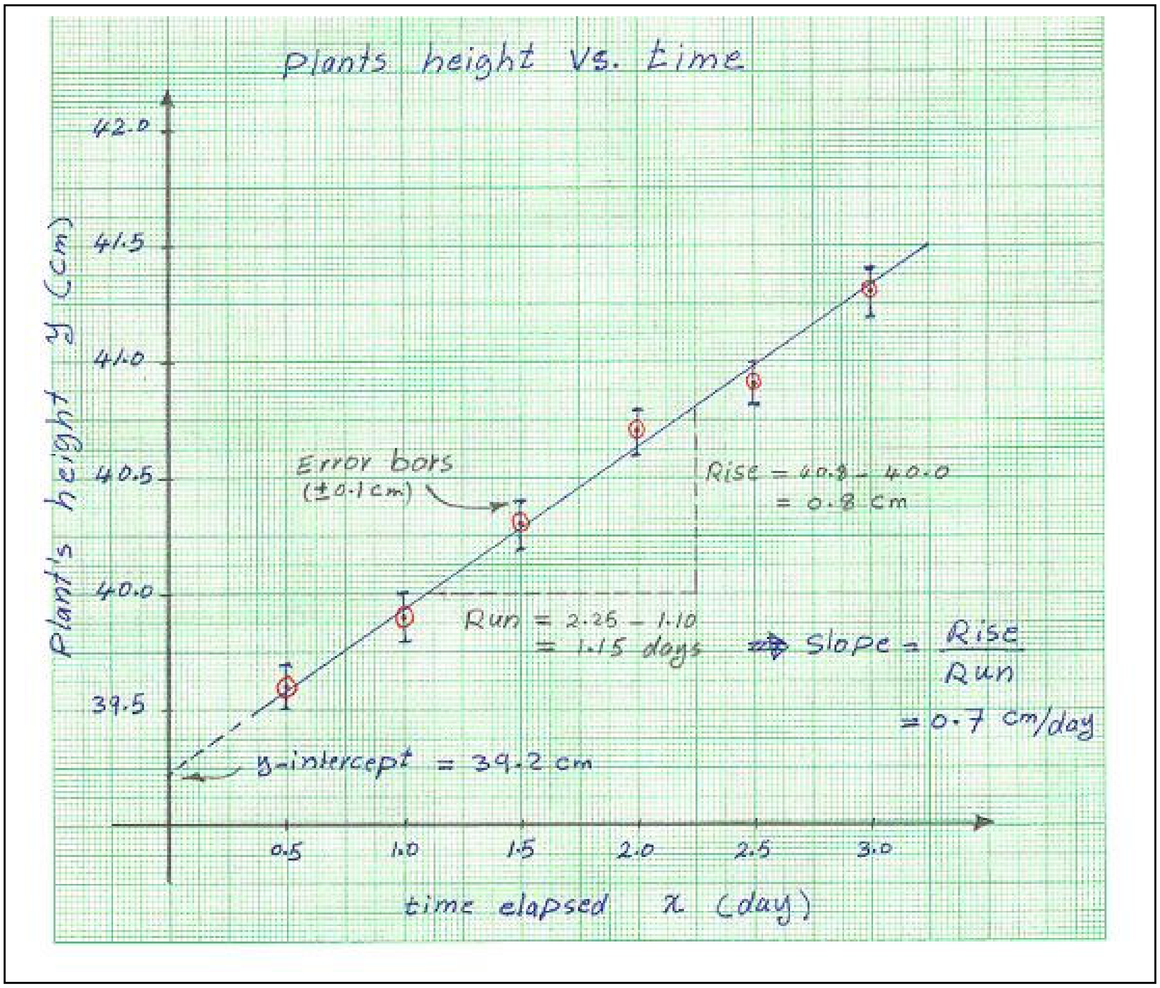

Figure 1.1

The Linear Graph

Many scientific studies involve finding how one quantity is related to another. One way to explore relationships is to collect and analyze experimental data. In a typical experiment, the value of the first quantity (called the independent variable) is varied, and the value of the second quantity (called the dependent variable) is measured. The result is two sets of values corresponding to both variables. For example, assume that you wanted to monitor the growth of your favourite indoor plant, which you got as a birthday gift from your mom. So, over a period of three days you recorded the plant's height every 12 hours as shown in Table 1.3. Since the measurements were made using a regular ruler, you estimated the error in the plant's height to be about 1 mm.

Table 1.3

Elapsed time x (day)	Plant's height $y \pm 0.1$ (cm)
0.5	39.6
1.0	39.9
1.5	40.3
2.0	40.7
2.5	40.9
3.0	41.3

It is very convenient to display tabulated data in a graph form. The independent variable is usually (but not necessarily) plotted on the horizontal axis while the vertical axis is reserved for the dependent variable. In Figure 1.1, the circled dots represent the *x* and *y* values in Table 1.3. The vertical bars on each data point show the range of error in each measurement. You may notice that the *y* value goes up with the *x* value to make a graph that is almost a straight line. This indicates that the linear function $y = mx + b$ will likely represent the data well. Finding the best straight line, that represents the relation between *x* and *y*, is called the linear fit to the data. The parameter *m* in this equation is called the slope of the line, and it represents the rate at which *y* increases with *x*. The other parameter *b* is called the y-intercept, and represents the value of *y* when $x = 0$.

To find the slope of the linear graph, we need first to mark two points on the straight line and write down their *x* and *y* coordinates. Note that these coordinates should be read directly from the linear graph and not from the data table. If the first point has the coordinates (x_1, y_1) and second points has the coordinates (x_2, y_2), then we can calculate the change in *y* (rise = $y_2 - y_1$) caused by the change in *x* (run = $x_2 - x_1$). The slope is then calculated as the ratio of rise over run such that

$$m = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1}.$$

In Figure 1.1, the slope is calculated to be $m = 0.7 \text{ cm/day}$, which is a relatively high growth rate!

The *y*-intercept (*b*) corresponds to the point at which the linear graph intersects the *y*-axis (i.e. when $x = 0$). In Figure 1.1, $b = 39.2 \text{ cm}$, which represents the height of the plant 12 hours before you took the first measurement. The best linear fit to the data in Table 1.3 is then given by the equation $y = 0.7x + 39.2$.

For additional readings you may refer to Section 1-4 and Appendix A-3 in the textbook.