APPENDIX: ERROR ANALYSIS

Sources of Error:

Experimental measurements are usually associated with three types of errors:

1. Personal error (mistakes you have made, time to redo!)
2. Systematic errors (consistent in repeated measurements)
3. Random error (uncertainty; an unexplainable variation in experimental conditions)

Errors in the first type can be reduced by carefully following directions, careful measurement techniques and data analysis. Errors of type 2 occur in the same direction in repeated measurements. These can be hard to detect but can be minimized by choosing appropriate testing conditions. An example would be a stopwatch running slow. This could be minimized by replacing the battery (to get it to run at the appropriate speed again) or further reduced by purchasing a better timer suited for the experiment or conditions. Errors of type 3 may occur in any direction, sometimes the measurement may be an overestimate, other times an underestimate. Illustrated below in figure A1 is an example of systematic error vs random error.

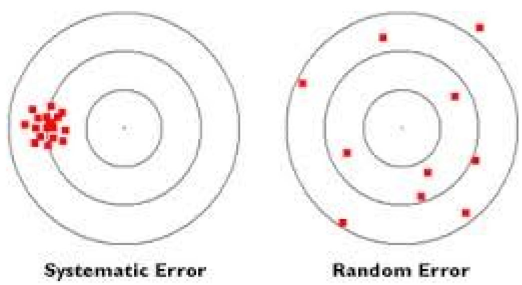

Figure A1: Systematic Error vs Random Error

To reduce errors of Type 3, you must repeat the experiment creating a mean or average value. This mean value is considered the best or most reliable value. How reliable is it? This can be determined by the standard deviation and/or the standard error of the measurements.

Mean and Standard Deviation

The mean value can be described as the average value expressed as:

$$\bar{x} = \frac{x_1+x_2+x_3+\cdots+x_N}{N} = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{\sum x_i}{N} \qquad \text{Equation A1}$$

The accuracy of the measurements for x_i is determined in terms of the standard deviation σ_x. Standard deviation is defined as:

$$\sigma_x = \sqrt{\left[\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2\right]} \qquad \text{Equation A2.}$$

Reporting Errors

There are three standard methods for reporting the uncertainty, Δx (also sometimes known as δx), in the measurement of a quantity x:

1. Using fraction uncertainty: $\Delta x/\bar{x}$
2. Using percentile uncertainty: $\frac{\Delta x}{\bar{x}} \times 100\%$
3. Using absolute uncertainty (Δx) as follows: $x(\text{measured}) = \bar{x} \pm \Delta x$

The third way of reporting error, absolute uncertainty, implies that the measurement falls between the average or best measurement plus or minus its standard deviation (its highest or lowest possible values). Using these, we get a more familiar way of reporting uncertainty:

$$\%\, error\ in\ x = \frac{\Delta x}{\bar{x}} \times 100\% = \frac{|x(measured) - x(accepted)|}{x(accepted)} \times 100\% \qquad \text{Equation A3}$$

The quantity Δx is take to be the standard deviation of the mean, and can be calculated by,

$$\Delta x = \frac{\sigma_x}{\sqrt{N}} = \sqrt{\left[\frac{1}{N(N-1)} \sum_{i=1}^{N} (x_i - \bar{x})^2\right]}. \qquad \text{Equation A4}$$

Propagation of error

Typically in lab, we measure more than one variable at a time or more than one variable is used at a time. This can lead to sources of error propagating or spreading to other variables more than one may realize. If we look at three variables, x, y, and z (in reality it can be as many as we need or would like) and their absolute uncertainty, $\Delta x, \Delta y,$ and Δz we will be left with 3 equations:

$$x(\text{measured}) = \bar{x} \pm \Delta x$$
$$y(\text{measured}) = \bar{y} \pm \Delta y$$
$$z(\text{measured}) = \bar{z} \pm \Delta z$$

These three variable (or more) may be used in an equation to solve for some variable known as Q. So what is the uncertainty, or ΔQ, of Q? The general equation for determining uncertainty, or error in propagation, in Q can be found as:

$$\Delta Q = \left[\left(\frac{\partial Q}{\partial x}\Delta x\right)^2 + \left(\frac{\partial Q}{\partial y}\Delta y\right)^2 + \left(\frac{\partial Q}{\partial z}\Delta z\right)^2 + \cdots\right]^{\frac{1}{2}} \qquad \text{Equation A5}$$

This equation is based on how the equation at hand is related to the variables used. The curved delta's are partial derivatives and can be solved with simple calculus in a few simple steps.

Step 1: Obtain $\frac{\partial Q}{\partial x}$. This can be done by differentiating Q with x holding all other variables (y, z, etc) constant. Next you'll need to obtain $\frac{\partial Q}{\partial y}$ while holding the other variables (x, z, etc) constant. You will continue this process until all of the experimented variables have been exhausted.

Step 2: Replace x with $\bar{x}$, y with $\bar{y}$, and z with $\bar{z}$ for all of the parts in step 1. Continue this process for all of the experimented variables. You will want to use the mean value in all of the variables.

Step 3: Substitute the partial derivatives you've found and modified in step 1 and 2 into Equation A5. From here you can usually multiply by "1" to rearrange the equation into $\frac{\Delta x}{\bar{x}}, \frac{\Delta y}{\bar{y}}, \frac{\Delta z}{\bar{z}}$, etc, this is optional (this is helpful to simplify calculations only).

Step 4: If you have the values for Δx, Δy, and Δz along with $\bar{x}, \bar{y}, \bar{z}$ substitute them into the equation and solve for the numerical answer.

An example of Steps 1 through 4 can be shown here:

To determine the uncertainty in an equation: $Q = x^2yz - xy^2 + xz^2$, a scientist (you) measured x, y, and z to be: $x = 3.0 \pm 0.1$ $y = 2.5 \pm 0.2$ and $z = 5.3 \pm 0.3$

Step 1:

$$\frac{\partial Q}{\partial x} = \frac{\partial}{\partial x}(x^2yz - xy^2 + xz^2) = 2xyz - y^2 + z^2$$

$$\frac{\partial Q}{\partial y} = \frac{\partial}{\partial y}(x^2yz - xy^2 + xz^2) = x^2z - 2xy + 0$$

$$\frac{\partial Q}{\partial z} = \frac{\partial}{\partial z}(x^2yz - xy^2 + xz^2) = x^2y - 0 + 2xz$$

Step 2:

$$\frac{\partial Q}{\partial x} = 2\bar{x}\bar{y}\bar{z} - \bar{y}^2 + \bar{z}^2$$

$$\frac{\partial Q}{\partial y} = \bar{x}^2\bar{z} - 2\bar{x}\bar{y} + 0$$

$$\frac{\partial Q}{\partial z} = \bar{x}^2\bar{y} - 0 + 2\bar{x}\bar{z}$$

Step 3:

$$\Delta Q = \left[((2\bar{x}\bar{y}\bar{z} - \bar{y}^2 + \bar{z}^2)\Delta x)^2 + \left((\overline{x}^2\bar{z} - 2\bar{x}\bar{y})\Delta y\right)^2 + ((\bar{x}^2\bar{y} + 2\bar{x}\bar{z})\Delta z)^2\right]^{\frac{1}{2}}$$

Step 4:

$$\Delta Q = \left[\left((2 * 3 * 2.5 * 5.3 - 2.5^2 + 5.3^2)0.1\right)^2 + \left((\overline{3}^2 * 5.3 - 2 * 3 * 2.5)0.2\right)^2 + ((3^2 * 2.5 + 2 * 3 * 5.3)0.3)^2\right]^{\frac{1}{2}}$$

$$= \left[\left((101.3)0.1\right)^2 + \left((32.7)0.2\right)^2 + ((54.3)0.3)^2\right]^{\frac{1}{2}}$$

$$= \sqrt{102.7 + 42.8 + 265.4} = 20.3$$

From our original equation, we know Q = 184.77 . So our final answer with proper significant figures and rounding is $Q = 180 \pm 20$.

Linear Regression (Linear Least Square Fit)

For your physics labs (and probably other labs as well), you often measure x_i (independent variable) and y_i (dependent variable) where these quantities are related through the usual linear equation:

$$y_i = Mx_i + C \qquad \text{Equation A5}$$

One can plot the y_i data points verse the x_i points and draw a "best fit straight line" with a ruler through the data points. This is only a qualitative best fit as there was no math involved. The method of linear regression provides a true and precise mathematical procedure for finding the best-fit line similar to what is found in computer programs such as excel or origin.

Due to uncertainty of measurements,

$$y_i - Mx_i + C = \Delta y_i \neq 0 \qquad \text{Equation A6}$$

In linear regression, the best fit straight line is obtained by optimizing the slope M and the intercept C, which by default will lead to a minimum value of Δy_i. This is done by minimizing a function termed "Chi-squared" χ^2, where

$$\chi^2 = \sum_{i=1}^{N} \frac{\Delta y_i^2}{\sigma_y^2} = \frac{(y_i - Mx_i + C)^2}{\sigma_y^2} \qquad \text{Equation A7}$$

Whereσ_y is the standard deviation mentioned in equation A2.

$$\sigma_y = \sqrt{\frac{\sum(y_i - \bar{y})^2}{N-1}}$$

Once reduced, we can solve for the unknown values which are important, M, the slope, and C, the y-intercept.

$$M = \frac{N\sum xy - \sum x \sum y}{D} \qquad \text{Equation A8}$$

$$C = \frac{\sum x^2 \sum y - \sum x \sum xy}{D} \qquad \text{Equation A9}$$

Where D, the denominator is:

$$D = N\sum x^2 - \left(\sum x\right)^2$$