Experiment

Error Analysis and Graph Drawing

I. Introduction:

I.1 It is impossible to do an experimental measurement with perfect accuracy. There is always an uncertainty associated with any measured quantity in an experiment even in the most carefully done experiment and despite using the most sophisticated instruments. This uncertainty in the measured value is known as the error in that particular measured quantity. There is no way by which one can measure a quantity with one hundred percent accuracy. In presenting experimental results it is very important to objectively estimate the error in the measured result. Such an exercise is very basic to experimental science. The importance of characterizing the accuracy and reliability of an experimental result is difficult to understate when we keep in mind that it is experimental evidence that validate scientific theories. Likewise, reliability and accuracy of measurements are also deeply relevant to Engineering.

The complete science of error analysis involves the theory of statistics (see Ref. 1,2) and is too involved to present here. This short presentation is intended to introduce the student to some basic aspects of error analysis and graph drawing, which it is expected that the student will then put into practice when presenting his/her results of the coming experiments.

I.2 When a measurement of a physical quantity is repeated, the results of the various measurements will, in general, spread over a range of values. This spread in the measured results are due to the errors in the experiment. Errors are generally classified into two types: *systematic* (or *determinate*) errors and *random* (or *indeterminate*) errors. A systematic error is an error, which is constant throughout a set of readings. Systematic errors lead to a clustering of the measured values around a value displaced from the "true" value of the quantity. Random errors on the other hand, can be either positive or negative and lead to a dispersion of the measurements around a mean value. For example, in a time period measurement, errors in starting and stopping the clock will lead to random errors, while a defect in the working of the clock will lead to systematic error. A striking example of systematic error is the measurement of the value of the electric charge of the electron '*e*' by Millikan by his

Oil Drop method. Millikan underestimated the viscosity of air, leading to a lower value for his result

$$e = (1.591 \pm 0.002) \times 10^{-19}\ \text{C}. \tag{1}$$

Compare this with a more modern and accurate value (Cohen and Taylor 1973, Ref 3)

$$e = (1.602\ 189 \pm 0.000\ 005) \times 10^{-19}\ \text{C}. \tag{2}$$

Systematic errors need to be carefully uncovered for the particular experimental set-up and eliminated by correcting the results of the measurements.

I.3 Random errors are handled using statistical analysis. Assume that a large number (N) of measurements are taken of a quantity Q giving values $Q_1, Q_2, Q_3, \ldots Q_N$. Let $\overline{Q}$ be the mean value of these measurements

$$\overline{Q} = \frac{1}{N}\sum_1^N Q_i, i = 1,2,\ldots..N \tag{3}$$

and let ‘d’ be the deviation in the measurements

$$d = \sqrt{\frac{1}{N}\sum_1^N \left(Q_i - \overline{Q}\right)^2}, i = 1,2,\ldots..N\ . \tag{4}$$

The result of the measurement is quoted (assuming systematic errors have been eliminated) as

$$Q = \overline{Q} \pm d. \tag{5}$$

The error ΔQ in the quantity Q is then taken to be the deviation d. (This is called the *standard error* in Q).

In a single measurement of a physical quantity, the error can be estimated as the least count (or its fraction) of the instrument being used.

As an example, the result of a measurement of the radius of curvature R, of a plaino-convex could be quoted as

$$R = 140 \pm 0.2\ \text{cm}. \tag{6}$$

This means that we expect that the value of R to be in the range 139.8 to 140.2 cm. Note however, that this does not mean that the “true” value of R necessarily lies in this range, only that there is a probability that it will do so.

The error in a measurement can also be quoted as a percent error,

$$\frac{\Delta Q}{\overline{Q}} \times 100 = \frac{d}{\overline{Q}} \times 100\ . \tag{7}$$

For example, the percent error in R is 0.143%.

I.4 **Combination of errors:**

Often the value of a quantity of interest may depend on other measured quantities. For example we could have a quantity Q which is a function F of a number of independent (actively controlled by us) variables say x,y and z i.e.,

$$Q = Q(x,y,z) \tag{8}$$

In general, the error in Q is related to errors in $x, y, z, \ldots$, as follows (for small errors)

$$(\Delta Q)^2 = (\Delta Q_x)^2 + (\Delta Q_y)^2 + (\Delta Q_z)^2 + \ldots \tag{9}$$

where

$$\Delta Q_x = \left(\frac{\partial Q}{\partial x}\right)\Delta x \quad ; \quad \Delta Q_y = \left(\frac{\partial Q}{\partial y}\right)\Delta y \quad ; \quad \Delta Q_z = \left(\frac{\partial Q}{\partial z}\right)\Delta z \quad etc.$$

The following table summarizes the results for combining errors for some standard functions. Try to derive some of these results.

Sr No	Function $Q(x,y)$	Error ΔQ or Fractional Error $\Delta Q/Q$
1	$Q = x + y$	$\Delta Q = \sqrt{(\Delta x)^2 + (\Delta y)^2}$
2	$Q = x - y$	$\Delta Q = \sqrt{(\Delta x)^2 + (\Delta y)^2}$
3	$Q = x\,y$	$\left(\frac{\Delta Q}{\overline{Q}}\right)^2 = \left(\frac{\Delta x}{\overline{x}}\right)^2 + \left(\frac{\Delta y}{\overline{y}}\right)^2 \Rightarrow \left(\frac{\Delta Q}{\overline{Q}}\right) = \sqrt{\left(\frac{\Delta x}{\overline{x}}\right)^2 + \left(\frac{\Delta y}{\overline{y}}\right)^2}$
4	$Q = x/y$	$\left(\frac{\Delta Q}{\overline{Q}}\right)^2 = \left(\frac{\Delta x}{\overline{x}}\right)^2 + \left(\frac{\Delta y}{\overline{y}}\right)^2 \Rightarrow \left(\frac{\Delta Q}{\overline{Q}}\right) = \sqrt{\left(\frac{\Delta x}{\overline{x}}\right)^2 + \left(\frac{\Delta y}{\overline{y}}\right)^2}$
5	$Q = x^n$	$\frac{\Delta Q}{\overline{Q}} = n\frac{\Delta x}{\overline{x}}$
6	$Q = \ln x$	$\Delta Q = \frac{\Delta x}{\overline{x}}$
7	$Q = e^x$	$\frac{\Delta Q}{\overline{Q}} = \Delta x$

II **Drawing of best fit straight line graph:**

Below we describe how to fit a straight line to a set of data. Relations that are not linear can be transformed to a linear one by an appropriate transformation of the variables (as you will learn from these assignments).

To draw the best fit straight line graph through a set of scattered experimental data points we will follow a standard statistical method, known as *least squares fit* method. Let us consider a set of N experimental data points (x_1,y_1), (x_2,y_2), $(x_3,y_3)\ldots(x_N,y_N)$. It is well known that a straight line graph is described by the equation

$$y = mx + c. \tag{10}$$

We ask the question: how are the slope m and the y-intercept c to be determined such that a straight line best approximates the curve passing through the data points? Let $S_i = y_i - mx_i - c$ be the deviation of any experimental point $P(x_i, y_i)$, from the best fit line. Then, the gradient 'm' and the intercept 'c' of the best fit straight line has to be found such that the quantity

$$S = \sum_i (y_i - mx_i - c)^2$$

is a minimum. We thus require

$$\frac{\partial S}{\partial m} = -2\sum_i x_i(y_i - mx_i - c) = 0 \quad and \quad \frac{\partial S}{\partial c} = -2\sum_i (y_i - mx_i - c) = 0,$$

which give,

$$m\sum x_i^2 + c\sum x_i = \sum x_i y_i \quad and \quad m\sum x_i + Nc = \sum y_i.$$

The second equation can be rewritten as $\bar{y} = m\bar{x} + c$, where $\bar{y} = \frac{1}{N}\sum y_i \quad and \quad \bar{x} = \left(\frac{1}{N}\sum x_i\right)$ showing that the best fit straight line passes through the centroid $(\bar{x}, \bar{y})$ of the points (x_i,y_i). The required values of m and c can be calculated from the above two equations to be

$$m = \frac{\sum (x_i - \bar{x})y_i}{\sum (x_i - \bar{x})^2} \quad and \quad c = \bar{y} - m\bar{x}\,. \tag{11}$$

The best fit straight line can be drawn by calculating m and c from above. A graphical method of obtaining the best fit line is to rotate a transparent ruler about the centroid so that it passes through the clusters of points at the top right and at the bottom left.

This line will give the maximum error in m, $(\Delta m)_1$ on one side. Do the same to find out the maximum error in m, $(\Delta m)_2$ on the other side. Now bisect the angle between these two lines and that will be the best fit line through the experimental data.

What are the errors in the gradient and intercept due to errors in the experimental data points? The estimates of the standard errors in the slope and intercept are

$$(\Delta m)^2 \approx \frac{1}{D}\frac{\sum S_i^{\,2}}{N-2} \quad and \quad (\Delta c)^2 \approx \left(\frac{1}{N}+\frac{\bar{x}^2}{D}\right)\frac{\sum S_i^{\,2}}{N-2},$$

where $D=\sum(x_i-\bar{x})^2$ and S_i is the deviation $S_i = y_i - mx_i - c$.

II.1 Presentation of error associated with experimental data in a graph:

Let us consider a function, $y = f(x)$, where x is an independent parameter which in the hand of the experimentalist during performing the experiments and y is the experimental data which is having a value depending upon the x and the instruments. Let the error associated with x be $\pm\Delta x$ and that for y be $\pm\Delta y$. One can represent $\pm\Delta x$ and $\pm\Delta y$ with the experimental data point $P(x,y)$ on the graph paper. To do that, first plot $P(x,y)$ on the graph paper, then draw a vertical line parallel to y axis about the point $P(x,y)$ of length $2\Delta y$. So upper half of the line represents the error $+\Delta y$ and the lower half represents $-\Delta y$ error. To present $\pm\Delta x$, draw horizontal lines at the two ends of the vertical line of length $2\Delta x$ each. The whole presentation is now giving the errors associated with the experimental point $P(x,y)$.

Figure 1 is an example of experimental data of resonance absorption of γ-ray experiment (Mössbauer spectroscopy) with error associated with each experimental data. The solid line gives the fitted curve through the experimental data. Note that the error in the variable along horizontal axis is not shown.

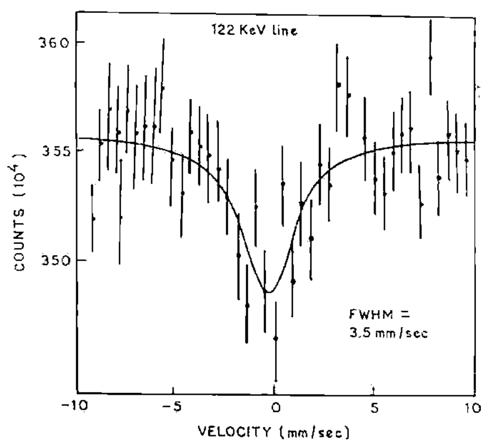

Fig.1

II.3 Use of graphs in experimental physics:

In practical physics, the graph of the experimental data is most important in improving the understanding of the experimental results. Moreover from the graphs one can calculate unknowns related to the experiments and one can compare the experimental data with the theoretical curve when they are presented on same graph. There are different types of graph papers available in market. So, one should choose the appropriate type of graph paper to present their experimental results in the best way depending upon the values of the experimental data and the theoretical expression of the functions. To understand all those some of the assignments are given below in addition to those we discussed before.

III Using calculator to fit curves

After doing these assignments you also have to verify the best fit parameters, you have obtained, by using your calculator. Calculators do the kind of curve fitting we want. Find out the procedure, specific to your calculator, to fit curves to a set of data. You will find it in the calculator's manual under heading like "statistical" and sub heading "Regression". If the formula of interest is one of the standard relations in the calculator then you simply enter the data and read of the best fit parameters that the calculator responds with. For other formulae you can use the trick of transforming to a new set of variables. In all your future work you can do the regression analysis using a calculator.

IV. Exercises and Viva Questions

1. What is the general classification of errors? Give an example of each. How are they taken care of?
2. What is the meaning of standard error? Calculate the standard error for the hypothetical data given in the adjacent table. Express the quantity as in eq. (5), i.e. $R = \overline{R} \pm d$.

Radius of curvature (cm)
130.121
130.136
130.139
130.148
130.155
130.162
130.169

3. What is the percent error in Millikan's measurement of the charge of the electron: $e = (1.591 \pm 0.002) \times 10^{-19}$ C?
4. What is the error in the volume of a cube $V = L^3$ if the error in L is 0.01m? If L is measured as $L = 2 \pm 0.01$, express the value of V in a similar manner.
5. A small steel ball-bearing rests on top of a horizontal table. The radius (R) of the ball is measured using a micrometer screw gauge (with vernier least count 0.05 mm) to be 2.15 mm. The height of the table is found using an ordinary meter scale

to be 90 cm. What is the height of the center of the steel ball from the floor (include the error)?

6. Let $Q = x - y$, where $x = 100 \pm 2$ and $y = 96 \pm 2$. Calculate Q (express the result with the error included).
7. Consider the quantity $Q = x/y$. If $x = 50 \pm 1$ and $y = 3 \pm 0.2$. Calculate Q (express the result with the error included).
8. In an experiment involving diffraction of sodium light using a diffraction grating, the doublet lines are unresolved at first order and a single spectral line is seen at an angle of 13^{o}. If the least count of the vernier of the telescope is $1'$, what will be the error in the calculated value of the grating constant d? (Principal maxima of a grating occur at angles θ such that $d sin\theta = m\lambda$. The wavelength separation between the sodium doublet lines is 6 A^{o}).
9. Consider an experiment to measure the gravitational acceleration g by measuring the time period of a simple pendulum. What are the possible sources of systematic error in this experiment?
10. "If there are always errors in any measurement then there is nothing like the 'true' value of any measured quantity". Comment on this statement. In what sense then do you understand the values of 'physical constants' to be constants?

References:

1. "Practical Physics", G.L. Squires, Cambridge University Press, Cambridge, 1985.
2. "Laboratory Experiments in College Physics", C.H. Bernard and C.D. Epp, John Wiley and Sons, Inc., New York, 1995.
3. Cohen, E.R. and Taylor, B.N., J. Phys. and Chem. Reference Data, Vol 2, page 663, 1973.