Measurements and Uncertainties

The establishment or verification of a physical law or the experimental determination of a physical quantity usually involves measurement. A reading taken from the scale on a stopwatch, a meter stick or a voltmeter, for example, may be directly related by a chain of analysis to the quantity or law under study; any uncertainty in these readings would thus be directly reflected in an uncertainty in the final result. A measurement by itself, without at least a rough quantitative statement as to the uncertainty involved, produces a result of limited usefulness. It is therefore essential that any introductory physics laboratory course include a discussion of the nature of experimental uncertainties and the manner in which they should be assigned to experimental results.

1. UNCERTAINTY vs. DISCREPANCY

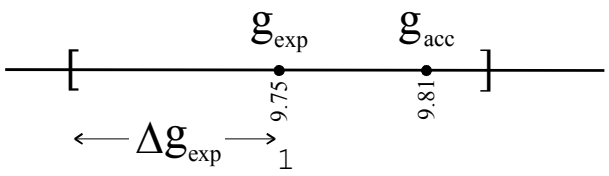
When you report the result of a measurement of a quantity x, you should also give the uncertainty Δx, e.g.,

```
5.0 m ± 0.1 m
⇗       ⇗
x       Δx
```

The uncertainty tells you how precise you think your measurement is. It is also often useful to compare your result with a "true" or accepted value; the difference between these is the discrepancy and is a reflection of the overall accuracy of the measurement.

An estimate of the uncertainty of a measurement should always be made; a calculation of the discrepancy can be made only if an accepted value or result happens to be available for comparison. The conclusion for an experiment should, whenever possible, address the question: ***Do the uncertainties involved in my measurements account for the discrepancies between my result and the accepted one?***

As an example, suppose you do an experiment to measure the acceleration of gravity and you report the experimental value ($g_{exp} \pm \Delta g_{exp}$) to be $9.75 \pm 0.08 \frac{m}{s^2}$ where the "accepted" value is $g_{acc} = 9.81 \frac{m}{s^2}$. As you can see from the graphic representation below, the *uncertainty* Δg_{exp} in the measurement accounts nicely for the *discrepanc*y between g_{exp} and g_{acc}.



2. ORIGINS OF UNCERTAINTIES

Problems which lead to discrepancies between experimental values and "true" values can be placed in two general categories:

I. Systematic Errors are inaccuracies due to identifiable causes and can, at least in principle, be eliminated. Errors of this kind result in values for the measured quantity which are **consistently** either too high or too low. Such errors can be

a) Theoretical - due to simplifications of the model system or approximations in the equations describing it.

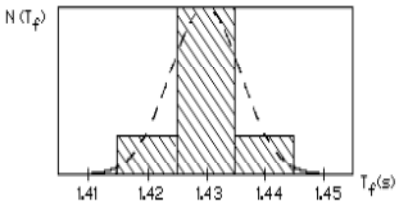
b) Instrumental - e.g., a poorly calibrated instrument.

c) Environmental - e.g., factors such as inadequately controlled temperature and pressure.

d) Observational - e.g., parallax in reading a meter scale.

II. Random Uncertainties are the result of small fluctuating disturbances which cause about half the measurements of any quantity to be too high and half to be too low. It is often not possible in practice to identify all the sources of such errors, to evaluate their effects individually or to completely eliminate them. They can, however, often be well characterized mathematically.

To illustrate the difference between systematic errors and random uncertainties, we consider the measurement of the length of time T_f taken for a ball to fall some fixed distance, say 10 m. Suppose we drop the ball about 1000 times and obtain as many values of T_f, rounding each time to the nearest .01 s. If $N(T_f)$ is a function of T_f that represents the number of times a particular measured value of T_f occurs, then a histogram of N versus T_f might look like



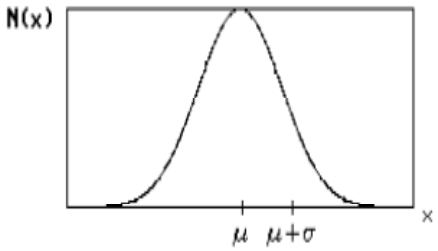
The "best" value for T_f is just the average (mean) value $\overline{T}_f$:

$$\overline{T}_f = \frac{\sum_{i=1}^{n} N_i T_{fi}}{n} \qquad (1)$$

where N_i is the number of times the value T_{fi} appears as a measurement and n is the total number of times the ball is dropped. In this example we see that $\overline{T}_f$ = 1.43 s; the discrepancy between this number and the "true" value of T_f, if known, would provide a measure of the accuracy of the measurement. Systematic errors (e.g. the reaction time involved in starting and stopping the clock) will affect $\overline{T}_f$ and hence the accuracy. The spread of T_f values, indicated by the width of the curve, is a reflection of the precision with which T_f is being measured. Random uncertainties contribute to this width and may be attributable, to e.g., small fluctuations of the height from which the ball was dropped or to the difficulty in determining the exact moment of impact with the ground. In what follows we discuss the mathematical treatment of such random uncertainties and the role they play in reporting laboratory measurements.

3. CHARACTERIZING A SET OF DATA: THE NORMAL DISTRIBUTION

It is most often the case that repeated measurements of the same quantity will, as in the timing experiment example described above, exhibit a spread about the average value related to the random errors associated with the measurement process. If we make "many" (say 10^6) measurements of a quantity x and plot the frequency of occurrence N(x), we quite often obtain a curve that approximates a Gaussian, or normal distribution, as pictured below.



This curve N(x) represents the relative probability with which values of x are obtained as the result of any single measurement of this quantity, which may be, for example, the reading of a meter or stopwatch. The analytical expression for such a curve is

$$N(x) = \frac{N_0}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (2)$$

where the parameters μ and σ determine the position and width of the peak, respectively. The "normalization" parameter N_O would correspond, in our timing example, to the total number of readings taken from the stopwatch.

Uncertainties

The curve representing Eq. (2) is of importance for the analysis of experimental data because in many cases this data is distributed normally and thus has a frequency distribution that can be "fit" to this curve. The more data taken, the better the fit will be. But for any set of data, regardless of the number of data points or the nature of the distribution, we can define quantities which characterize the "best value" and "spread" of the data.

For a set of data points x_i, the quantity $\overline{x}$ is the mean of all values obtained for x, and is defined by

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n} \qquad \textbf{(3)}$$

where n is the number of measurements. For data in which each value of x_i generally occurs more than once, Eq. (1) may provide a more convenient way to calculate the mean value of x. The "best value" that one finally reports for the quantity x is generally its mean, as defined above. If a large number of normally distributed data points are taken, then $\overline{x}$ should be close to the value of μ for which the curve representing Eq. (2) best fits the data.

We can characterize the spread of any finite set of n data points by its standard deviation, symbolized by s.d., which we will define as

$$\text{s.d.} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n}} \qquad \textbf{(4)}$$

The s.d. for a set of data thus represents the square root of the average square of the deviation (i.e., the "rms" deviation) that the data points exhibit with respect to their average value. For a large number of normally distributed data points, s.d. $\approx \sigma$, where σ is the value of the parameter of Eq. (2) which produces the best description of the data.

4. REPORTING THE VALUE OF A MEASURED QUANTITY

Because repeated measurements of the same quantity give us, as discussed above, a range of values reflecting the random fluctuations inherent in the measurement process, we need to report the result of these measurements in a way which reveals not only the best value of the quantity, but also the precision to which we are stating it: we report **both** the best value **and** the uncertainty. The distributions discussed above refer to a collection of single measurements. So if we make a single measurement of a quantity x and had some knowledge of the single measurement distribution curve, then we could report this value of x as

$$x \pm \text{s.d.} \qquad \textbf{(5)}$$

where x is the single measured value and s.d. is the standard deviation of the single measurement distribution. To determine an appropriate value for s.d. without actually making several

measurements, we would have to know something, a priori, about the single measurement curve; we would either have to be given a value of s.d. or we would have to guess.

Suppose, as is more common, we take n measurements instead of just one. Intuitively, it seems reasonable that the mean value should be reported with a smaller uncertainty than the single measurement discussed above. In fact, if one took **many** sets of n measurements and then calculated the mean for **each** set, the distribution of these **means** would have a smaller standard deviation (denoted S.D.) than that for a single measurement (s.d.). It is not hard to show that

$$\text{S.D.} \approx \frac{\text{s.d.}}{\sqrt{n}} \qquad \textbf{(6)}$$

What this means is that one could report the result of a set of n measurements as

$$\bar{x} \pm \text{S.D.} \approx \bar{x} \pm \frac{\text{s.d.}}{\sqrt{n}} \qquad \textbf{(7)}$$

so that the uncertainty associated with the average of 5 measurements, e.g., could be reduced by about a factor of 2 from the single measurement s.d. **NOTE**: The expression above for reporting an average is only meant to serve as a guideline in reporting results; it assumes that all uncertainties involved are purely random in nature. Systematic errors must always, at some level, be considered, so that in most cases we **cannot** create an arbitrarily small uncertainty simply by taking more and more measurements!

Relative vs. Absolute Uncertainty. One often reports the relative (or percentage) uncertainty associated with the best value in lieu of the absolute uncertainty Δx = s.d. or S.D. No units are associated with these relative uncertainties, so that these may be used to directly compare the precision of two different measurements. On the other hand, the absolute uncertainty is more convenient for comparing two measured values.

Example. Suppose we make 10 measurements of the period T of a pendulum, and that the results are tabulated to the nearest millisecond (ms) as

TRIAL #	T (s)	TRIAL #	T (s)
1	1.411	6	1.468
2	1.452	7	1.437
3	1.403	8	1.446
4	1.414	9	1.425
5	1.459	10	1.434

We need to report the best value for T along with the uncertainty, which we assume is due to random fluctuations. For the best value, we have the mean,

$$\bar{T} = \frac{\sum_{i=1}^{10} T_i}{10} = 1.435 \text{ s.}$$

To estimate the uncertainty, we calculate s.d. from Eq. (4):

$$\text{s.d.} = \sqrt{\frac{\sum_{i=1}^{10}(T_i - \overline{T})^2}{10}} = 0.02 \text{ s.}$$

But since we made 10 measurements, we can report the uncertainty as

$$\Delta T \approx \text{S.D.} \approx \frac{\text{s.d.}}{\sqrt{10}} \approx 0.006 \text{ s.}$$

Our result for T thus appears as

T = 1.435 s ± 0.006 s (with absolute uncertainty)
or
T = 1.435 s ± 0.5 % (with relative uncertainty)

Significant Figures. A decision must always be made, consciously or otherwise, as to how many significant figures should be used to report a quantity x and its uncertainty Δx. Since uncertainties are merely estimates, *it is seldom necessary to report Δx to more than 1 significant figure!* The number of significant figures with which you report x is often taken to imply its uncertainty, therefore this number *must be consistent with the explicitly stated uncertainty Δx*.

Rule of thumb for significant figures: *The measurement and the uncertainty should have their last digits in the same location relative to the decimal point.* Example:
$(1.06 \pm 0.01) \times 10^3$ m.

Thus, a length stated as 1.0 m ± 0.002 m would present a problem because the stated uncertainty would imply that the observer is withholding information by not reporting more decimal places in the length. A result stated as 1.06 m ± 0.2 m reports a length with too much accuracy, given the uncertainty. If x and Δx imply different uncertainties, one of them should be adjusted so that they are both consistent with the largest of the two uncertainties.

Guesstimating Uncertainties. The uncertainty to assign a given single measurement of x depends on the technique employed in the measurement, the quality of the equipment used, and the care with which the measurement is made. The number you come up with as Δx for a particular measurement is often quite subjective, because only **you** know what happened during **your** measurement. There are a couple of options for estimating Δx:

1. Take several measurements and calculate s.d., as in the example above for T;

2. Estimate Δx from the nature of the measurement. For example, when reading a scale, the uncertainty is **sometimes**, but **not always**, half of the smallest scale division, but you will encounter situations when the scale divisions have nothing to do with Δx. In other words, you will always have to use your judgment.

You will find yourself using option #2 quite often; it's a real timesaver.

5. PROPAGATION OF UNCERTAINTIES

Quite often the quantity of interest is measured indirectly, i.e., derived from one or more quantities which are each measured directly and reported with uncertainties, as discussed above. In these cases, the estimate of the uncertainty in the indirectly measured quantity must reflect the uncertainties associated with all of the relevant directly measured ones.

For example, suppose we wish to infer the acceleration of gravity (g) from direct measurement of both T_f (the time taken for an object to fall) and of L (the height from which it is dropped). The equations of kinematics tell us that g can be calculated from T_f and L:

$$g = \frac{2L}{T_f^2}.$$

How, then, do the uncertainties in L and T_f contribute to Δg, the uncertainty in g?

You will need to know the answer to this question, and to others just like it. We first give the general expression for propagating the uncertainties in directly measured quantities to find the uncertainty in the derived result. We then treat two special cases which occur quite often.

General Case: Let z be a quantity to be derived from the direct measurement of two independent quantities x and y by using the relationship z = f(x,y). Suppose we were to make small definite (known) errors Δx and Δy in the measurements of x and y respectively. These would show up as an error Δz in z that would be given by

$$\Delta z \approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y .$$

But because Δx and Δy are actually **random** uncertainties, a more realistic estimate of the uncertainty Δz is

$$\Delta z \approx \sqrt{\left(\frac{\partial f}{\partial x}\Delta x\right)^2 + \left(\frac{\partial f}{\partial y}\Delta y\right)^2}, \qquad \textbf{(8)}$$

as can be demonstrated from considerations of probability theory. Given uncertainties in the two quantities Δx and Δy, Eq. (8) can be used to estimate the propagated uncertainty Δz. The generalization to more than two directly measured quantities is straightforward.

Origin of Eq. (8) (Optional): Consider two independent random variables P and Q, each with an average of zero and with s.d.'s of P and Q, respectively. The s.d. of the sum P + Q (denoted (P+Q)) is, according to the definition of Eq. (4),

$$\Delta(P+Q) = \sqrt{\overline{(P+Q)^2}}$$

If we compute the square of this quantity in terms of ΔP and ΔQ, we get

from which the s.d. for P + Q is

$$\Delta(P+Q) = \sqrt{(\Delta P)^2 + (\Delta Q)^2}.$$

Now considering the function z = f(x,y): the uncertainty in z (Δz), will be due to the combined effects of the uncertainty in x (x) and the uncertainty in y (Δy). If P and Q represent the effect on z of changes in x and y, respectively, then ΔP and ΔQ, the s.d.'s for P and Q, will be given by

$$\Delta P = \frac{\partial f}{\partial x}\Delta x; \quad \Delta Q = \frac{\partial f}{\partial y}\Delta y,$$

where Δx and Δy are the s.d.'s, or uncertainties, in x and y. The spread, or uncertainty Δz in z will then be Δ (P+Q), given above. Substitution for ΔP and ΔQ results directly in Eq. (8) for Δz. It should be noted that this same argument can also be applied to derive Eq. (6) above for the uncertainty (S.D.) in the average.

Many times, it is not necessary to use Eq. (8) directly because the form of the expression for z will fit one of two **special cases**.

Special Case I: The quantity of interest (z) is given in terms of two directly measured quantities x and y, each with their respective uncertainties Δx and Δy, as **z = Ax + By**. In this case, the absolute uncertainty in z is given by

$$\Delta z = \sqrt{(A\Delta x)^2 + (B\Delta y)^2 + \cdots} \qquad \textbf{(8a)}$$

where the "+⋯" indicates that this expression is easily generalized to the case where z is given as the sum of three or more quantities.

Example: We wish to determine the mass of water, m_w, contained in a beaker from independent measurements of the mass of the beaker alone, m_b, and the combined mass of the beaker + water, m_t. Then $m_w = m_t - m_b$ and the uncertainty in the derived quantity m_w is

$$\Delta m_w = \sqrt{(\Delta m_t)^2 + (\Delta m_b)^2},$$

where Δm_t and Δm_b must either be estimated or determined from repeated measurements of these quantities.

Special Case II: The derived quantity (z) is the product of powers of measured quantities x and y: $\mathbf{z = K x^a y^b}$, where K is a constant. In this case we can give the fractional (relative) uncertainty in z as

$$\frac{\Delta z}{z} = \sqrt{\left[a\frac{\Delta x}{x}\right]^2 + \left[b\frac{\Delta y}{y}\right]^2} \qquad \textbf{(8b)}$$

where we can again generalize to any number of quantities, as indicated for Eq. (8a).

Example: If the acceleration of gravity g is given, as discussed above, in terms of L and T_f as $g = 2L/T_f^2$, what is Δg in terms of the uncertainties ΔL and ΔT_f? If g is rewritten as $g = 2(L^1)(T_f^{-2})$, then we have something that fits the form of special case II with x = L, a = 1 and $y = T_f$, b = -2. So for the relative uncertainty $\Delta g/g$ we get

$$\frac{\Delta g}{g} = \sqrt{\left[\frac{\Delta L}{L}\right]^2 + 4\left[\frac{\Delta T_f}{T}\right]^2}\,.$$

Note that because g depends on the second power of T_f but on only the first power of L, the uncertainty in g is much more "sensitive" to the uncertainty in T_f than to the uncertainty in L.

If neither of the special cases applies to the quantity being investigated, it is sometimes possible to derive the uncertainty by applying Eq. (8a) and (8b) in combination; otherwise, use the general expression of Eq. (8) directly.

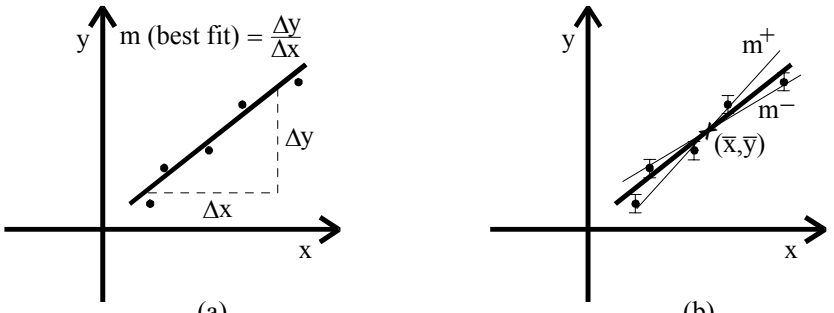
Graphs and Uncertainties. Physical laws often predict that two quantities (x and y) that can be measured (directly or indirectly) are proportional to each other. For example, Newton's second law in one dimension (F = Ma) predicts that the acceleration (a) of an object is proportional to the force F (the mass M being the constant of proportionality); Ohm's law (V = IR) predicts, for certain materials, that the voltage V and current I are proportional. It is often the constant of proportionality between these measured quantities (M or R in the above examples) that is to be determined by the experiment, rather than just individual pairs of values for x and y. The linear relationship between x and y can be expressed as

slope ↘

$$y = mx + b$$

↖ y-intercept

If y is plotted vs. x, as in graph (a) below, then the result **should** be a line with a slope m, calculated as $\Delta y/\Delta x$. (Note that here Δy and Δx refer to the rise and run, respectively, **not** to the uncertainties in y and x!).



Of course, experiments being what they are, the data points (as indicated in graph (a)) will not all be on a single line; it is up to you to draw the "best" line that "fits" the data so that the slope m (or y-intercept b) may be determined. There is a well-known method for determining the best line for a given set of points (x,y) -- the least squares fit (see below). For the purposes of this lab, however, it will often be sufficient to use a straightedge and your eyes to draw the line which represents the best compromise, i.e., "comes as close to as many points as possible."

NOTE: If the slope or the intercept is to be measured directly from the graph, it is most often advantageous to use *as much of the page as possible to draw the graph*!

If the slope m is the quantity to be reported as the result of the experiment, then the uncertainty Δm must be reported along with it: we need $m \pm \Delta m$. The uncertainty Δm can be thought of as arising from the uncertainties in the position of each plotted point on the graph, since each y value may be characterized by an uncertainty Δy. One way of representing these uncertainties is by means of the "error bars" drawn for each point on graph (b) above. As with the slope m, there is a well known (and somewhat complex) expression that can be used to determine Δm, but for our purposes it is sufficient to "eye ball" the range of slopes. As indicated in graph (b), this can be done by using the error bars to estimate m^+, the highest possible slope for a line that passes through or near most of the error bars, and m^-, the lowest slope. The uncertainty Δm is then **very roughly** given by

$$\Delta m \approx \frac{m^+ - m^-}{2}.$$

To estimate the slope (m) and its uncertainty (Δm) by eye:

1. Plot the data points (x,y) on graph paper. Use as much of the page as possible.

2. Place error bars through each point to indicate roughly the extent to which the location of that point on the graph is uncertain. The size of these bars can be estimated from the discussion under the "uncertainty" section in each experiment. An alternative to bars is to enlarge your dot to indicate its "fuzziness." If the uncertainty is too small to show up on your graph, state this in the report.

3. Plot the point $(\bar{x}, \bar{y})$ on your graph (where $\bar{x}$ and $\bar{y}$ are the means of the x and y values, respectively). It so happens that this point always lies on the "best line" defined below by the least squares fit procedure.

Uncertainties

4. Using a straight edge, draw the line passing through $(\bar{x}, \bar{y})$ that "comes closest" to the data points. Measure the slope m as $\frac{\text{rise}}{\text{run}}$, using as much of the line as possible and *paying attention to the units on the axis scales*.

5. Estimate the uncertainty Δm by drawing 2 dotted lines through $(\bar{x}, \bar{y})$ and passing within most of the error bars--one with the minimum slope, m^-, and one with the maximum slope, m^+. Estimate Δm from the expression in the paragraph above.

Least Squares Fit (Linear Regression): Though somewhat more involved and less intuitive than the graphical method described above for finding the slope and intercept of a best fit line, the analytic method of linear regression finds the best fit by minimizing the sum of squares of deviations of y values from the fitted straight line. For this reason, it is called the *method of least squares*. Given N data points $x_1, y_1; x_2, y_2; ...; x_N, y_N$, the slope m and y-intercept b of the best fit line (given by $y = mx + b$) are calculated as:

$$m = \frac{\sum_{i=1}^{N}(x - \bar{x})(y - \bar{y})}{N\sigma_x^{\,2}}; \quad b = \bar{y} - m\bar{x}, \qquad \textbf{(9)}$$

where $\bar{x}$ and $\bar{y}$ are the means of the x and y values respectively, given by Eq. (3), and σ_x is the s.d. of the x values as given by Eq. (4).

As an alternative to the graphical method described above for estimating the uncertainty Δm for the slope, you may use the least squares methodology to estimate this quantity along with the uncertainty in the intercept, Δb. The expressions for these are

$$\Delta m = \frac{1}{\sigma_x}\sqrt{\frac{\sum_{i=1}^{N} d^2}{N(N-2)}}; \qquad \Delta b = \Delta m \sqrt{\sigma_x^{\,2} + \bar{x}^2}\,, \qquad \textbf{(10)}$$

in which $d_i \equiv y_i - (mx_i + b)$ is the y deviation of the point (x_i,y_i) from the best fit straight line $y = mx + b$. If your calculator does linear regression calculations which result in the least squares fit slope m and correlation coefficient r, then it may be simpler to calculate the uncertainty in the slope (Δm) with the equation*

$$\Delta m = |m| \frac{\tan\left(\arccos(r)\right)}{\sqrt{N-2}} = \frac{|m|}{r}\sqrt{\frac{1-r^2}{N-2}}\,. \qquad \textbf{(11)}$$

References and further reading:

1. Lichten, William, Data and Error Analysis in the Introductory Physics Laboratory (Allyn and Bacon, Inc.)

2. Taylor, John R., An Introduction to Error Analysis (University Science Books)

- J. Higbie, *American Journal of Physics*, **59**(2) 1991

Summary on Uncertainties and Least Squares Fitting

I. Mean (average) and Standard deviation of means

a) Measurements : data

$x_1 \pm \Delta x,\ x_2 \pm \Delta x,\ x_3 \pm \Delta x, \ldots\ldots\ldots\ldots,\ x_N \pm \Delta x$

where Δx = uncertainty of measurement (random error)

b) Mean (average): the best value to represent this set of data

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_N}{N}$$

c) Standard deviation of a 'single' measurement Δx

$$s.d. = \Delta x = \sqrt{\frac{(\bar{x} - x_1)^2 + (\bar{x} - x_2)^2 + \ldots + (\bar{x} - x_N)^2}{N}}$$

d) Standard deviation of Means: S. D.

$$S.D. \approx \frac{s.d.}{\sqrt{N}} = \frac{\Delta x}{\sqrt{N}}$$

e) Report of the measurements;

$\bar{x} \pm S.D.$ $\quad$ $S.D.$ in one significant figure

II. Propagation of Uncertainties

Given: $A \pm \Delta A$, $B \pm \Delta B$ and $f = f(A,B)$

a) In general

$$\Delta f = \sqrt{\left(\frac{\partial f}{\partial A}\Delta A\right)^2 + \left(\frac{\partial f}{\partial B}\Delta B\right)^2}$$

b) Addition/ subtraction

$f = \alpha A + \beta B$

$$\Delta f = \sqrt{(\alpha\,\Delta A)^2 + (\beta\,\Delta B)^2}$$

c) Multiplication/ division/ powers

$$f = A^{\alpha}\ B^{\beta}$$

$$\Delta f = f\sqrt{\left(\alpha\frac{\Delta A}{A}\right)^2 + \left(\beta\frac{\Delta B}{B}\right)^2}$$

III. Least Squares Fit (Linear Regression):

a) Measurements: (x_1, y_1), (x_2, y_2),……, (x_N, y_N), N pairs of data**.**

b) In theory $y = mx + b$

m = *slope*, b = *y-intercept*

c) Result of Least Squares Fit

$$m = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{N\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$b = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{N\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$\Delta m = \frac{1}{\sigma_x}\sqrt{\frac{\sum_{i=1}^{N}\left[y_i - (mx_i + b)\right]^2}{N(N-1)}} = m\sqrt{\frac{\frac{1}{r^2} - 1}{N}} \approx m\sqrt{\frac{1 - r^2}{N}}$$

r = *correlation coefficient*

$$\Delta b = \Delta m\sqrt{\sigma_x^2 + \bar{x}^2}, \quad \sigma_x = \sqrt{\frac{\sum(\bar{x} - x_i)^2}{N}}, \quad \bar{x} = \frac{\sum x_i}{N}$$