Introduction

In science, you must always state your final result **with a confidence limit**.

$$x \pm \triangle x \text{ units}$$

Your *final result* is your best estimate for the value being investigated, whereas the *confidence limits* are the upper and lower limits between which you declare that you confidently expect the true value to lie.

You are required to understand the errors, estimate their magnitude using some relatively simple formulae, reduce them by developing good experimental technique, and discuss them at length throughout your report. The failure to consider, estimate, reduce or discuss the error associated with any quantity you obtain is very poor experimental technique.

Error analysis really is quite simple and is of great importance to your report and your development as an experimental physicist. Topics covered include sources and types of error, how to calculate and combine errors and how to display your final result. Errors are often thought of as troublesome, but given the appropriate background knowledge they are really quite easy.

1 Sources of error in experimental work

1.1 Mistakes (!!)

First, let us mention momentarily errors that are just silly blunders. Misreading a scale, recording a number incorrectly or a making a mistake in calculation are all 'errors', but they can't be estimated for inclusion in a result, so they are not really errors in the scientific sense. The best advice with these errors is to **double check** and **re-take** questionable measurements.

1.2 Random errors

Random errors become evident when one takes repeated readings of a physical quantity and a small spread of readings is obtained. The cause could be fluctuations in experimental conditions, unavoidable instrument variations, or your own unconscious subjective bias in the setting or reading of instruments.

Random errors can be investigated by taking **repeated measurements**.

1.2.1 Repeated readings

If you need to know the value of a quantity as accurately as possible, it is obvious that you need to check the result a few times. Repeating the measurement allows you to calculate

the best value for the quantity being investigated and investigate the magnitude of the random errors by considering the statistical variation in the data.

The best way to estimate the **best value** of the result is the arithmetic mean of the measured result,

$$\langle X \rangle = \frac{1}{N} \sum_i X_i. \tag{1}$$

where N is the number of points in the data set.

An estimate of the **uncertainty of the result** comes from the standard deviation σ of the data,

$$\sigma^2 = \frac{\sum_i (X_i - \langle X \rangle)^2}{(N-1)}. \tag{2}$$

For errors with a normal distribution we expect 68% ($= 1 - 1/e$) of the recorded measurements to be within one standard deviation of the mean.

1.3 Reading errors

Both the accuracy and precision[1] of any real measuring device is limited.

In first year laboratories you learnt that the precision is limited to perhaps to half of the smallest scale on the instrument (ruler). More sensitive equipment improves the precision with which values can be obtained; however, a reading error is still **always** present.

1.3.1 Parallax

A quick little reminder of a thing called **parallax**. It is important to align yourself with the instrument appropriately if you wish your results to be accurate.

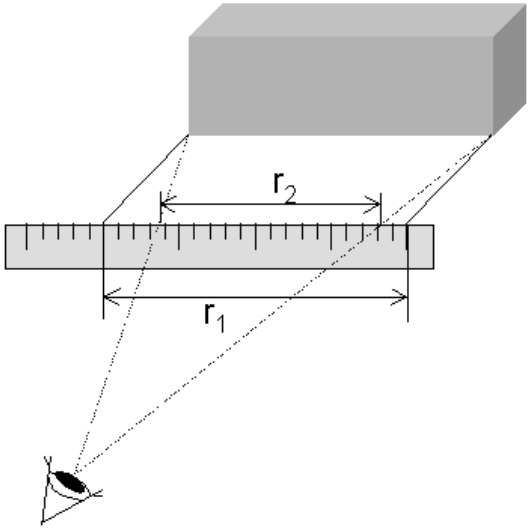

Figure 1: Parallax errors arising from the measurement of the length of an object. The apparent length, r_2, in this case is shorter than the real length, r_1. The apparent length would be the real length if the measurement was taken with the ruler in contact with the object.

1.3.2 Limit of accuracy

$$\boxed{\text{Limit of accuracy} = \text{half of the smallest division}}$$

The accuracy of the result will depend upon the accuracy of the measuring device you use. In figure 2 we consider the length of a block of material. It is quite obvious that the accuracy of our result is greater in the second case. In the Second Year laboratories the accuracy of our result should be down to a half of the smallest division. This applies to both rulers and calipers.

[1]The distinction between accuracy and precision can be illustrated by making an analogy with target archery. Precision is how close the arrows are together; accuracy is how close they are to the bullseye.

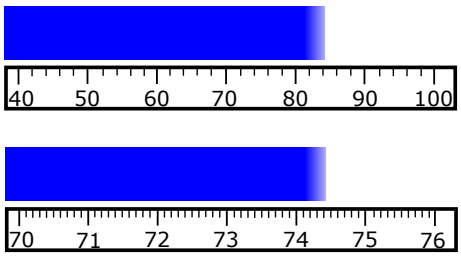

Figure 2: The first scale can be read to the nearest 0.5 mm, while the second scale can be read to the nearest 0.05 mm.

1.3.3 The Vernier scale

A Vernier scale[2] is used whenever one needs to make a measurement of a distance or angle to a greater accuracy than that obtainable through direct visual reading of a linear scale. The Vernier scale uses the linear reading scale of the measuring apparatus (i.e. a ruler) along with another scale which is scaled by a factor of 9:10 compared to the linear scale.

Reading a Vernier scale is best understood with an example (figure 3), but it'll take you a few measurements of your own before you can really use it confidently.

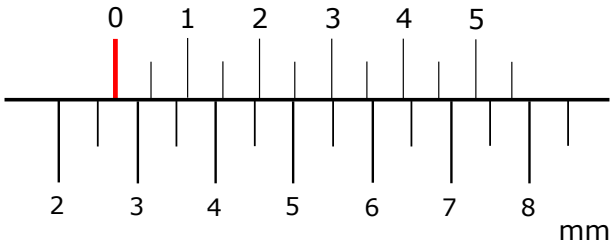

Figure 3: Vernier scale

In order to read this scale:

- The top line of numbers is the Vernier scale itself, with the bottom line the measurement itself. The Vernier scale scrolls across the measurement scale.
- We treat the 'zero' line on the Vernier as indicating the region of the measurement. This is the first approximation to the measurement and yields a result of **2.5 mm**.
- Next, we determine which marking on the Vernier scale most closely lines up with the markings on the measurement scale. We need to judge whether the Vernier scale is to the left or the right of the linear scale. When we cannot say clearly left or right, we have alignment.

[2]Invented by the French mathematician and inventor Pierre Vernier (1850-1637).

- In this case we find that **2.5** on the Vernier scale has the best alignment. We add this to the first approximation to find the reading, giving a result of **2.75 mm**.
- Finally, the smallest scale on the Vernier scale is **0.005 mm**. This is the uncertainty in the measurement.

Thus our final result would be

$$(2.50 + 0.25 \pm 0.005)\ \text{mm} = (2.750 \pm 0.005)\ \text{mm}$$

1.3.4 Digital voltmeter

If you are using a digital voltmeter to record a voltage, it might be reading 1.004 volts, with the smallest unit 0.001 volts. However the accuracy of the digital metres is typically 1 digit plus some fraction of the reading (typically 0.3%). In this example the error is calculated as follows:

$$\boxed{\triangle V = \tfrac{3}{1000} \times V + \text{limit of reading}}$$

So in the above case we find:

$$\triangle V = 3/1000 \times 1.004 + 0.001 = 0.004012$$

therefore the error is 0.004 V after accounting for significant figures.

Thus the final result plus confidence limit is (1.004 ± 0.004) V.

This error may seem surprisingly 'large', but **this is good scientific analysis**. If you remember a few years ago, there was a measurement in Italy that suggested some neutrinos had travelled faster than light, which was eventually shown to be a timing error... Can you imagine if proper error analysis hadn't been performed?

1.4 Systematic errors

As mentioned above, random errors can generally be controlled by averaging a number of readings. There is however another type of error - known as *systematic errors*. Systematic errors are errors that *systematically* shift the measurements in a particular direction away from the best value. Thus repeated readings **do not show** the presence of systematic errors, and no amount of averaging will reduce their effects.

Both systematic and random errors are often present in the same measurement, and the different effects they have are contrasted in figure 4.

There are no rules for eliminating systematic errors, or even detecting them, but the following points may be of some assistance.

- **Apparatus**

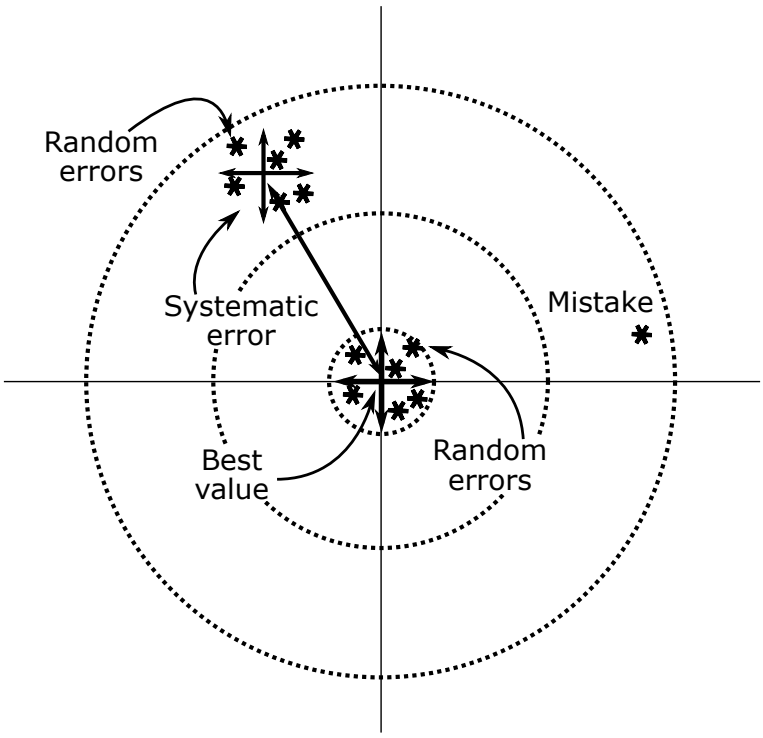

Figure 4: Schematic diagram showing the different types of errors and how they are manifest in the results.

  - If at all feasible, meters should be checked against a standard, or against another and hopefully better meter. Zero settings should be checked and adjusted if necessary.
  - So-called 'given values' should be treated with suspicion, and checked if possible. Resistors especially can vary 10% or more from recently measured values.
  - If possible, vary the conditions at measurement slightly to avoid systematic errors. Measure using different cables, or starting from different ends of a set of measurements.

- **Observations**
  - Measurements should be repeated by different observers to detect experimenter bias.
  - Corrections for instrument bias should be made, such as incorrect zero settings on meters.
  - Any necessary precautions (warm up times, temperature controls, minimizing vibration, etc.) should be observed and recorded.

- **Analysis**

– The experiment should be carefully analysed for likely sources of error and steps should be taken to minimise these.
– Results should be checked for feasibility, and **wildly outlying** readings re-taken or recalculated.

2 Graphs and errors

2.1 Error bars

Estimates of uncertainties in measured values should be indicated on graphs as error bars. These are bars passing through the estimated point running from the lower limit to the upper limit of the range covered by the uncertainties.

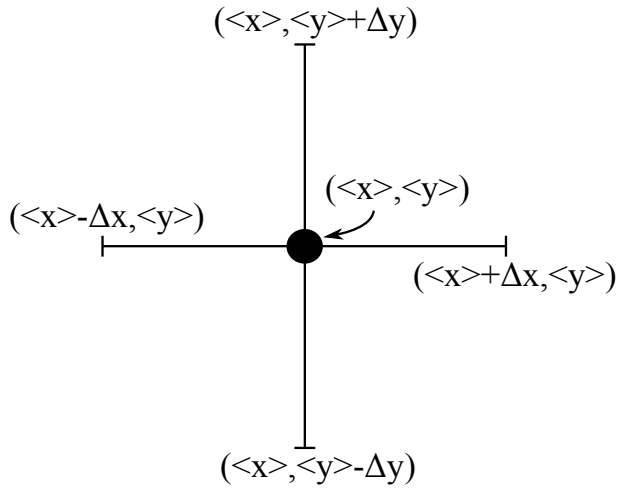

Figure 5: Error bars

Thus if $(\langle x\rangle, \langle y\rangle)$ is the estimated point and $\triangle x$ and $\triangle y$ are the uncertainties then the error bar in the x direction runs from $\langle x\rangle - \triangle x$ to $\langle x\rangle + \triangle x$, and the bar in the y directions runs from $\langle y\rangle - \triangle y$ to $\langle y\rangle + \triangle y$ as shown in figure 5.

2.2 Estimates of gradients and uncertainties

In many experiments we obtain results from *gradients* of plotted data. Although the trend of the data may be clear, random errors will cause the plotted points not to lie precisely on any given function.

There are thus two questions to be answered:

- What is the best estimate for the gradient of the line?
- What is the uncertainty in the estimated gradient?

There are a number of ways to estimate the gradient of a straight line from a plot of experimental points. We shall consider two such methods: best fit by eye, and least squares fitting, while mentioning a general equation that can be applied as a double-check.

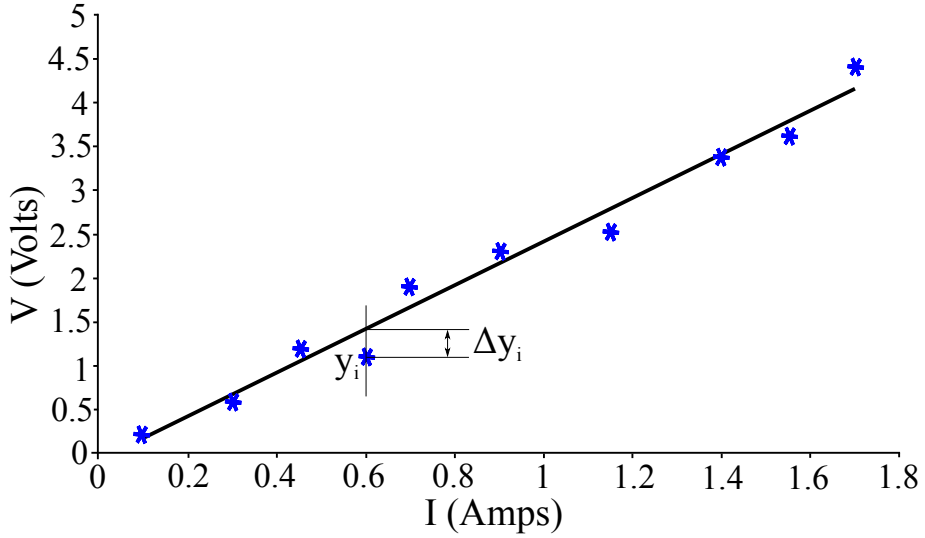

Figure 6: Line of best fit by eye.

2.2.1 Best fit by eye

In many cases it is sufficient to draw the single straight line which seems to lie closest to the largest number of experimental points.

If you were to draw error bars of length one standard deviation on each point, then a line of best fit should pass through **at least 2/3** of these error bars. If the error bars are two standard deviations long, the line should pass through 90% of the bars.

The gradient can of course be calculated using "rise over run" and the equation of a straight line:

$$y = mx + c \qquad m = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1}$$

The uncertainty in the gradient can be calculated using the 'maximum and minimum gradient approach'. Two lines are drawn, a steepest and most shallow line that fit the data, (with fit meaning that they pass through 67% or 90% of the error bars for one and two standard deviation error bars respectively). The gradient is calculated for each line and the average of these is used as the estimate of the gradient.

$$\boxed{\text{gradient} = \tfrac{1}{2}\ (\text{steepest gradient} + \text{shallow gradient})}$$

The uncertainty can be taken as

$$\boxed{\text{gradient} = \tfrac{1}{2}\ (\text{steepest gradient - shallow gradient})}$$

This steep/shallow gradient approach is a good way to get an indication of whether your individual errors are reasonable. Is the error in your gradient reasonable? If not, look at your data and ask yourself why.

2.2.2 Least squares fit

This technique is one of the most accurate techniques for obtaining the gradient of a line of best fit. The technique was developed by Gauss and assumes that all of the error is concentrated in the y co-ordinate, and asks

> "For a line of the form $y = mx + c$, what values of m and c yield the smallest value for the sum of errors (differences between the point it predicts and the experimental point) squared?"

This technique minimises the total distance between the points on a line of best fit and the experimental point to obtain the best fit.

While it may be possible to manually use the method of least squares for data, we have limited time, so can instead turn to Excel to perform this technique for us. Excel can find the line of best fit in its **curve fit**algorithm. You should **always** discuss the results of any automated error-finding or line-fitting algorithm.

Never assume physical values excel offers are 'correct' simply by virtue of being a computer program!

2.2.3 Equation for secondary check

You may have noticed occasionally that some graphs give an R^2 value along with an equation for the line of best fit. We can use the following equation to have a more 'mathematical' idea of the error in the gradient, rather than the more physical method of steepest/shallowest gradient.

$$\Delta m = m\sqrt{\frac{\frac{1}{R^2} - 1}{N - 2}} \tag{3}$$

where m is the gradient value, R^2 is the R^2 value given with the line of best fit, and N is the number of data points.

3 Combining errors

It's exceptionally rare to find an experiment with only a single source of error. Several random and systematic errors will likely be present in measurements of any given quantity. It is further even more likely that each experimental quantity will have to be combined to calculate the required results. Thus it is important to know how errors work with each other if we want to estimate the overall error in any result.

The basis of this combination is the following formula, which really makes a lot of sense. Suppose that your final result Z is a function of some experimentally determined quantities $(x, y, \ldots)$. It seems reasonable that if x and y have small errors in them, then the final result, being a function of those variables will also have some error in it. In the mid 1600's Sir Isaac Newton developed calculus (and partial derivatives) to

consider such variations. The general formulae for the error in Z due to experimental uncertainties $(\triangle x, \triangle y, \ldots)$ in the parameters is

$$(\triangle Z)^2 \approx \left(\frac{\delta Z}{\delta x}\right)^2 (\triangle x)^2 + \left(\frac{\delta Z}{\delta y}\right)^2 (\triangle y)^2 + \ldots \tag{4}$$

Every error we quote can be calculated using the parameters in equation 4. The errors $\triangle x, \triangle y, \ldots$ can be random errors such as the standard deviation from repeated measurements, reading errors from a Vernier scale, tolerance limits from a digital voltmeter, uncertainties in the gradient of a graph... All these 'errors' combine together to give an error in the final result.

The example below illustrates the use of this equation and calculation of the final error.

3.1 Worked example

Consider the use of a diffraction grating to determine the wavelength, λ, of an unknown emission line. The formula is:

$$m\lambda = d\sin\theta. \tag{5}$$

Thus

$$\lambda = \frac{d\sin\theta}{m}. \tag{6}$$

We are trying to find out, if there are small errors (standard deviation, reading errors ...) in the variables m, d and θ, what the resultant error in λ is. We calculate the partial derivatives of λ with respect to d, m and θ and then use equation 4)

$$\frac{\delta\lambda}{\delta d} = \frac{\sin\theta}{m}, \qquad \frac{\delta\lambda}{\delta m} = \frac{-d\sin\theta}{m^2}, \qquad \frac{\delta\lambda}{\delta\theta} = \frac{d\cos\theta}{m} \tag{7}$$

So we can now write:

$$(\triangle\lambda)^2 = \frac{\sin^2\theta}{m^2}(\triangle d)^2 + \frac{d^2\sin^2\theta}{m^4}(\triangle m)^2 + \frac{d^2\cos^2\theta}{m^2}(\triangle q)^2\left(\frac{\pi}{180}\right)^2 \tag{8}$$

Where the last term was converted into radians, which allows you to enter the angle in degrees. Alternatively you could not include the $(\pi/180)^2$ term and enter the angle in radians. If given the quantities (with m a DISCRETE variable with no associated error):

$$\begin{aligned} d &= (1710 \pm 2) \text{ nm}, \\ m &= 1, \\ q &= (17.30 \pm 0.01) \text{ degrees}, \end{aligned}$$

we arrive at the result

$\lambda = (508.511 \pm 0.6595)$ nm $= (508.5 \pm 0.7)$ nm.

Note also that we only quote the error to one significant figure. This leads us to a discussion about...

4 Significant figures and result presentation

The correct presentation of the final result from the previous problem is (508.5 ± 0.7) nm.

This means we are confident that our result is between 507.8 nm and 509.2 nm.
Please take note of the following:

- Results can be rounded up or down depending upon the magnitude of the error. Ask your demonstrator if you are unsure.
- The 'best guess' can only be as accurate as the error. There is no point in quoting (508.5110 ± 0.6) nm. The 0.0110 is made redundant by the greater figure on the error.
- Usually only the first significant figure in the error is of any use. Stating the wavelength is (508.5 ± 0.7)nm is good; quoting $\pm$ 0.6595 shows you haven't considered how the error affects the result.
- Displaying the result as $(5.085 \times 10^2 \pm 6.595 \times 10^{-1})$ nm is far from elegant and indicates that we can calculate the error but aren't confident relating it to our result. We should instead express the result and error as $(5.085 \pm 0.007) \times 10^2$ nm.

Calculating the error is only half the exercise. We must also *consider both how to minimise and the proper expression of error.*

5 General expressions for error

5.1 General expression

If the result Z depends upon $(x, y, \ldots)$, the corresponding error is calculated as follows:

$$(\triangle Z)^2 \approx \left(\frac{dZ}{dx}\right)^2 (\triangle x)^2 + \left(\frac{dZ}{dy}\right)^2 (\triangle y)^2 + \ldots \tag{9}$$

5.2 Sums and differences

If $Z = x + y + \ldots$ or $Z = x - y - \ldots$ then[3]

$$(\triangle Z)^2 \approx (\triangle x)^2 + (\triangle y)^2 + \ldots \tag{10}$$

5.3 Products and quotients

If $Z = x \times y \times \ldots$ or $Z = x \div y \div \ldots$, then

$$\left(\frac{\triangle Z}{\langle Z \rangle}\right)^2 = \left(\frac{\triangle x}{\langle x \rangle}\right)^2 + \left(\frac{\triangle y}{\langle y \rangle}\right)^2 + \ldots \tag{11}$$

5.4 Powers and functions

If $Z = x^N$, then

$$\left(\frac{\triangle Z}{\langle Z \rangle}\right) = |N| \left(\frac{\triangle x}{\langle x \rangle}\right) \tag{12}$$

5.5 Simple and trigonometric functions

If $Z = f(x)$, then[4]

$$\left(\frac{\triangle Z}{\langle Z \rangle}\right) = \left|\frac{df}{dx}\right| \left(\frac{\triangle x}{\langle x \rangle}\right) \tag{13}$$

[3]Independent errors add in quadrature if normally distributed. If they are not normally distributed then they may NOT add in quadrature. If they are not independent then they WILL generally NOT add in quadrature.

[4]This last category includes all the trigonometric functions. Angles must be in radians unless we have included the $(\pi/180)^2$ term in the error.