NAME: ________________________

COURSE I

LABORATORY GUIDE TO MEASUREMENT, ANALYSIS AND REPORTING

DEPARTMENT OF PHYSICS

UNIVERSITY OF CAPE TOWN

LABORATORY GUIDE TO MEASUREMENT, REPORTING AND ANALYSIS

CONTENTS

SECTION A – MEASUREMENT

A1. THE MEASUREMENT PROCESS

Traditionally, the methods used to determine the value of some physical quantity and the uncertainty associated with the determination of that quantity were characterised by the terms "random error" and "systematic error". The traditional method, often referred to as a "frequentist approach", gave rise to inconsistencies in the way in which experimental results were quoted. To correct the deficiency, new standards for quoting measurement results have been published and these have been accepted worldwide.

The framework adopted in the standards is known as the "probabilistic approach" and you are encouraged to acquaint yourself with the methods described in those specifications, especially if you expect to be involved in laboratory work in your future career. The relevant standards are expressed in ISO/IEC 17025 and the associated Guide to the Expression of Uncertainty in Measurement (GUM); the ISO International Vocabulary of Basic and General Terms in Metrology (VIM); and NIST Technical Note 1297 of 1994.

It is important for you to be aware that the term "error" is often used quite loosely in conversation in laboratory work, particularly by older persons, but the word is <u>not</u> synonymous with "uncertainty"! The term "uncertainty" is clearly defined while use of "error" is often misleading. The use of the word "error" in the context of laboratory work is discouraged.

In this section, a number of new words and terms are introduced. As you work through each experiment and report, it is expected that you will refer constantly to this section and think carefully about what each of these words and terms implies, because their meaning is often not the same as that to which we have become accustomed in everyday use. A useful site for further information is: ***http://www.bipm.org/utils/common/documents/jcgm/****.*

<u>Introduction</u>

The purpose of undertaking a measurement in science is to provide knowledge about some physical quantity. The physical quantity that we may wish to investigate through an experiment is called the **measurand**. Examples of measurands are the length of a box or the speed of light in some medium.

It is important to realise that the best that can be achieved through a process of measurement in science is to improve our knowledge about some specified measurand, knowing all the while that the value of a measurand can never be determined with absolute certainty! We cannot think of a measurand as having some "true" or "exact" value that can actually be found. We have to think of the knowledge that we have of any measurand always as being limited and incomplete.

The measurement process

In Figure 1 we identify the steps that need to be taken when going through the measurement process in order to determine the required result. Thereafter terms used in the process are explained.

Step 1:

Decide what the **measurand** is (what is to be measured). Then choose the best instrument available to you for this measurement process.

↓

Step 2:

Use the appropriate instrument to take one or more **readings** that will provide sufficient data from which to infer knowledge about the measurand (see Figure 2).

↓

Step 3:

From the reading (or readings) taken, decide on the **best approximation of the value of the measurand**.

↓

Step 4:

Consider all the sources of uncertainty in the knowledge of the measurand that may be inferred from the measurement process, and then **evaluate the uncertainty** (see Figure 3).

↓

Step 5:

Quote the final result of the measurement process.

Figure 1: The measurement process

Note:

It is quite common for people to refer to "taking a measurement" when what they actually mean is "taking a reading". A reading is simply a number that is obtained by using and observing an instrument of some sort, which is different from measurement, which is a process; the outcome of which is to improve our knowledge of a measurand.

Taking readings

The following flowchart shows the process that should be followed when taking readings.

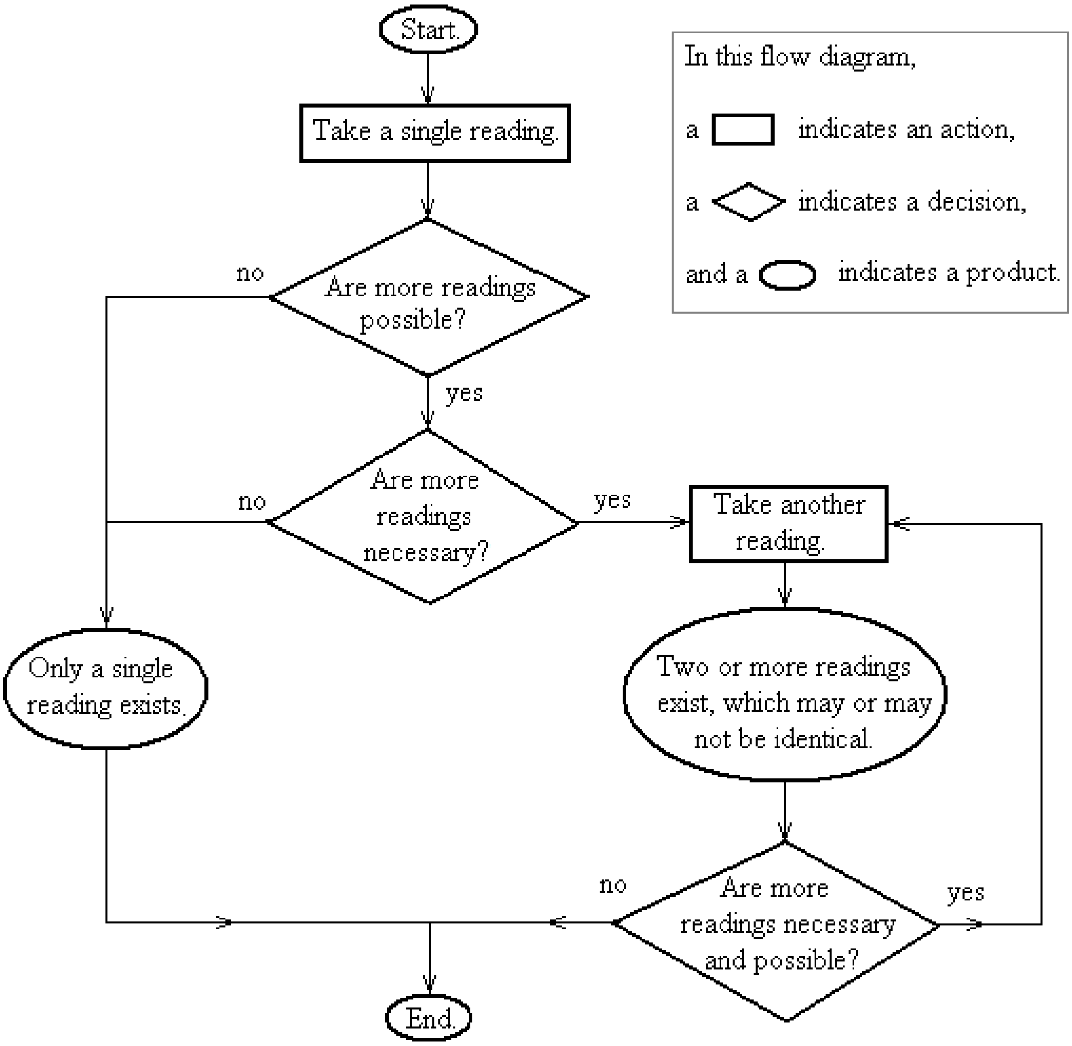


Figure 2: Flowchart for taking readings

Deciding on the best approximation of the measurand

In the case where only **one** reading has been taken, then that reading is the "best approximation of the value of the measurand". However, where **more than one** reading has been taken, then the "best approximation of the measurand" is most often the mean (average) of those readings.

Uncertainty

Figure 3 presents an overview of the two GUM approaches (**TYPE A** and **TYPE B**) that are used to evaluate the uncertainty in a measurement process. The detail of these analyses is covered in the next section of this manual.

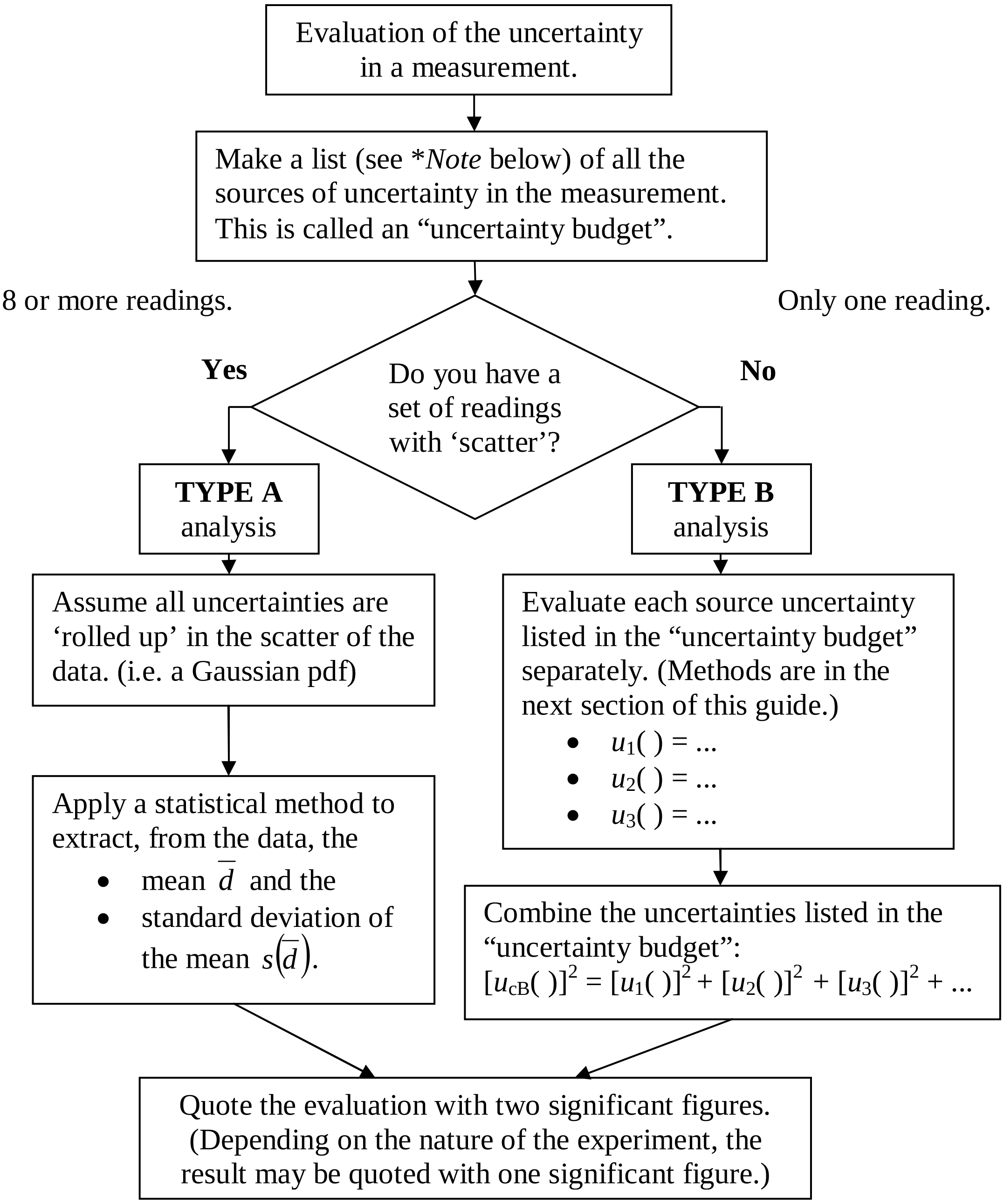


Figure 3: Flowchart for determining uncertainty in measurement

***<u>Note:</u>**

It is never possible to obtain a perfect result when working through a measurement process because there are <u>always</u> sources of uncertainty in the knowledge that can be inferred about a measurand.

Quoting the result (an interval)

A detailed discussion on how to quote the result is given in the Reporting section of this guide. There are however some important things of which you should take note.

When the uncertainty associated with a measurement is quantified, the outcome of a measurement defines an interval. As you will see in the guide, the outcome of a process of measurement is written as the (best approximation ± uncertainty).

For example, the value of '*g*' may be given as: (9.790 ± 0.052) $m.s^{-2}$

Note:
You are reminded that an interval is a set of numbers with the property that any number that lies between the two numbers defining the interval is included in the set of numbers.

You will recognise that the 'greater' the uncertainty, the 'wider' the interval that quantifies the result of the measurement process and because there will always be an uncertainty in every measurement process, the width of the interval that quantifies the result of a process of measurement will never be zero!

Caution*:*
The uncertainty in the result of a measurement is not an indication of mistakes that may have been made while working through the measurement process. If you are aware of your having made a mistake, then you should repeat the experiment; "Human error" is not a valid source of uncertainty in a result!

A2. QUANTIFYING UNCERTAINTY- TYPE A and TYPE B

Introduction

The problem being dealt with is to find a systematic method whereby we can start with data (one or more readings), and then sensibly infer knowledge about the measurand in question. The way in which this is done is to apply statistical methods concerning the probability of what the knowledge may be. More specifically, the method involves choosing a probability density function (pdf) that best describes the knowledge that can be inferred from the data (given the measurement process), and then applying appropriate statistical calculations to quantify the relevant uncertainties.

For the purposes of this course however, we will not delve into probability theory or the details of the different pdf's that may be used to model our knowledge of measurands in different situations - we will merely quote the results as we need them. You are however encouraged to improve your understanding of this aspect of laboratory reporting in your own time.

There are two distinct approaches (see Figure 3) by which uncertainties applicable to a process of measurement are quantified, namely:

- **TYPE A** evaluation of uncertainty, and
- **TYPE B** evaluation of uncertainty.

TYPE A evaluation of uncertainty

A **TYPE A** evaluation of the uncertainty is applicable to the uncertainty associated with the 'spread' or 'scatter' in values of a number of readings that have been taken in the course of a measurement process.

No doubt you are familiar with the idea that a 'Gaussian' is a symmetrical bell-shaped curve that represents the normal distribution of a random variable. It turns out that this property of the Gaussian (or Normal) probability density function makes it an appropriate probability density function upon which to model our knowledge of a measurand **where the relevant data is a set of readings with some 'scatter' in the value of those readings**.

No rigorous explanation for the above statement is given in this manual, except that we illustrate the point by way of an example to convince ourselves that the Gaussian probability density function is indeed an appropriate choice.

Example to illustrate the choice of the Gaussian pdf in **TYPE A** evaluations

Consider an example in which the measurand is the time for you to travel from wherever you live to the university campus. So, for the next 40 trips to the university you use the stopwatch on your cell phone to determine your travel time in minutes, i.e. you take a set of readings.

The data are tabulated as follows:

Table 8.1: Time taken to get from home to campus (in minutes)

17.1	16.9	17.4	16.4	16.8	17.9	16.0	18.2
18.1	19.3	16.9	17.4	17.8	17.1	17.6	17.8
16.4	17.4	15.9	15.8	17.7	17.3	17.4	16.8
16.6	16.8	17.6	17.4	18.9	18.4	17.9	17.8
18.3	17.2	18.1	18.7	16.3	17.4	18.4	16.5

A simple calculation shows that the average time taken is 17. 4 minutes.

We wish to present this data in a histogram, so the next step is to draw up a frequency table (or distribution table). Note that in this example a bin width of 0.5 minutes was chosen. The choice of bin width is arbitrary and is normally chosen to suit the data.

Table 8.2: Frequency table for data in Table 8.1

Bin (minutes)	Number of readings per bin
15.5 to 15.9	2
16.0 to 16.4	4
16.5 to 16.9	7
17.0 to 17.4	10
17.5 to 17.9	8
18.0 to 18.4	6
18.5 to 18.9	2
19.0 to 19.4	1
	Total : 40

Next we plot a histogram showing how the data are distributed over the values of the readings taken by plotting the number of readings per bin vs the time interval of bins, see Figure 4.

Finally, to show that the general shape of the histogram suggests that the distribution of the data is Gaussian shaped, we superimpose a trend line, shown as the light grey line in Figure 4.

While this illustration is in no way a rigorous explanation, it should be noted that as more and more readings are added to the set of data, and the bin width is narrowed, the distribution of any set of data of this sort will tend towards the formation of a symmetrical Gaussian distribution.

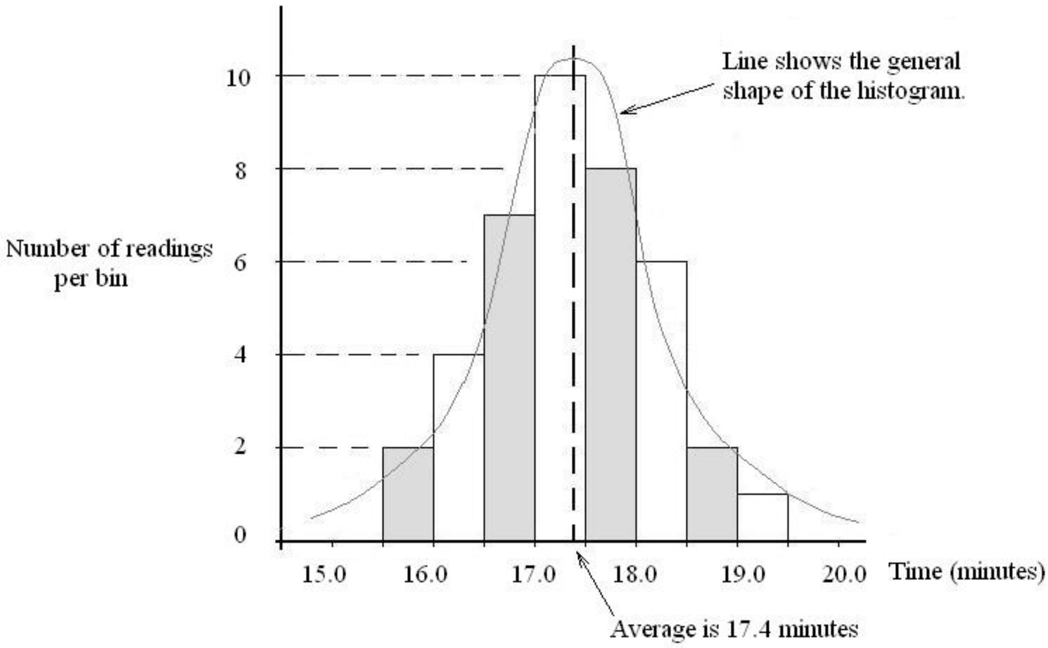


Figure 4: Histogram of travel times

<u>The Gaussian probability density function (pdf)</u>

Having decided that the Gaussian pdf is a suitable probability density function by which to model our knowledge of the measurand, we need to consider the implications of this application.

You are reminded that a probability density function of a continuous random variable is a function that can be used to obtain the probability that the random variable takes a value in a given interval. Further, when the probability density function is portrayed graphically, the area under the graph (subtended by the interval) indicates the probability that the random variable takes a value in that interval. The total area under the pdf (the probability that the value exists) is therefore unity.

The important features of the Gaussian pdf - the standard deviation $s(d)$ in particular - are illustrated graphically in Figure 5.

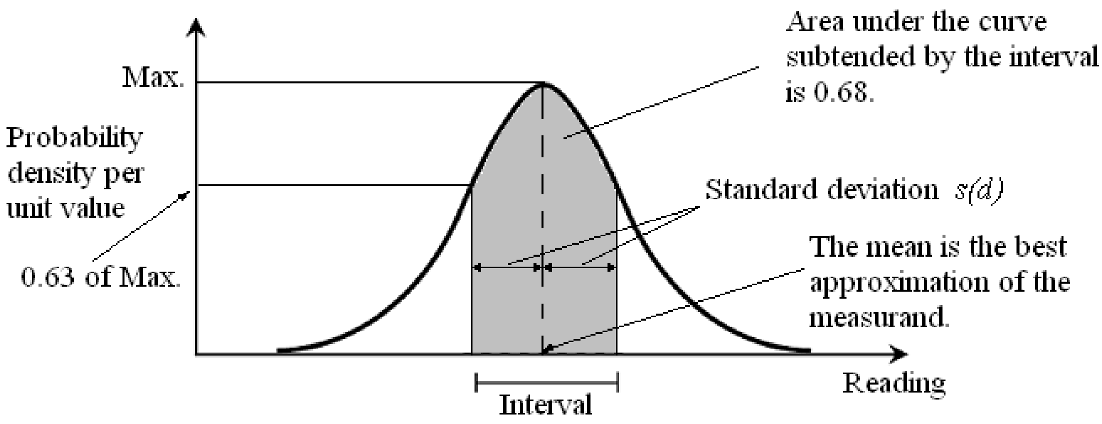


Figure 5: Gaussian pdf showing one standard deviation

The standard deviation $s(d)$ is defined to be half the 'width' of the Gaussian at 0.63 x (maximum of the Gaussian pdf). The equation, Eq. (2), used to calculate the standard deviation directly from the data, is given in the next section.

In this context, the **standard deviation** of a set of data is a measure of the 'spread' of the set of readings around the mean of those readings. In laboratory work, the use of the standard deviation not only gives an accepted measure of the consistency of the measurement process that was used in an experiment, but also makes it possible to specify the level of confidence with which the result of the measurement is quoted.

Of note is that the interval shown in Figure 5, which is defined by the 'mean ± one standard deviation', subtends an area of approximately 0.68; which means that the probability of finding any single reading in that interval is 68%.

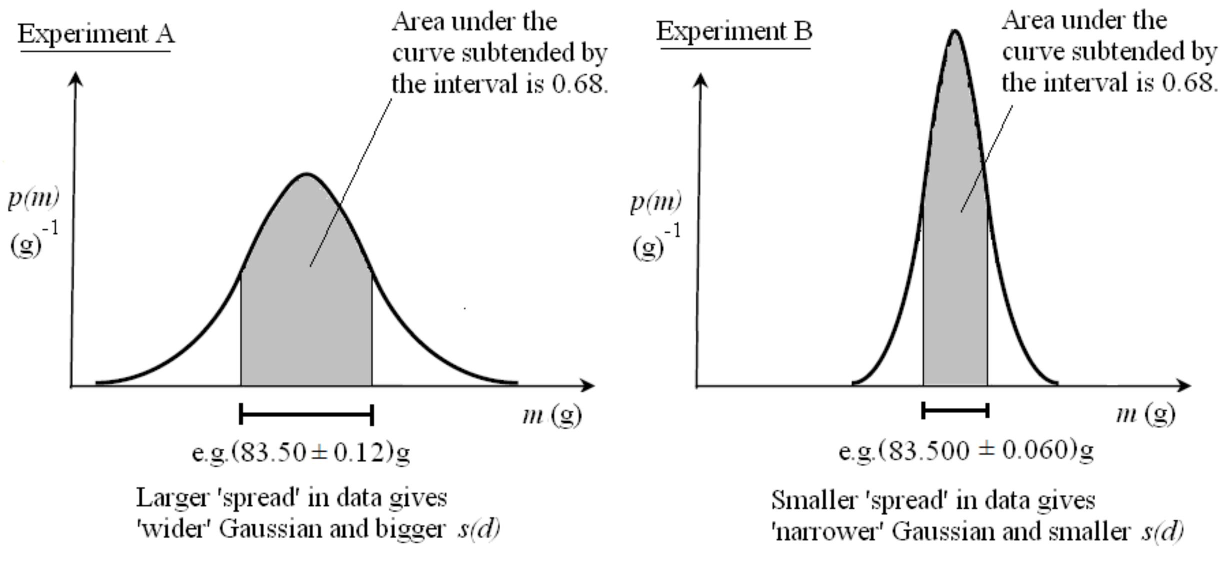


Figure 6: Effect of spread in data on standard deviation $s(d)$

From the example given in Figure 6, it can be seen that when the spread in the data is relatively large, the applicable Gaussian probability density function will be relatively wide; meaning that the standard deviation $s(d)$ will be a relatively large number. However, when the spread in the data is relatively small, the applicable Gaussian probability density function will be relatively narrow; meaning that the standard deviation $s(d)$ will be a relatively small number.

<u>How to do a **TYPE A** calculation of the uncertainty associated with a set of readings</u>

Having decided that the uncertainty in the measurement can be represented by a spread in the readings and that the spread in the readins 'fits' a Gaussian probability function, there are three parameters required to calculate the uncertainty in the measurement.

- The first is the **mean** $\overline{d}$,
- the second is the **standard deviation**, $s(d)$,
- and third is the **standard deviation of the mean** $s(\overline{d})$.

The reason for using the **standard deviation of the mean** in the final result is beyond the scope of this course, but consider that if the whole experiment were repeated a number of times, then the spread in the means derived from each experiment would also be represented by a Gaussian probability density function with a new width related to the **standard deviation of the mean**. Since the mean is quoted as the best approximation, the uncertainty quoted should also relate to the mean, and not to a single reading.

The following equations are used to calculate the necessary parameters:

Mean (average) $\overline{d}$ of the data
$$\overline{d} = \frac{1}{N}\sum_{i=1}^{N} d_i \qquad (1)$$

Standard deviation $s(d)$
$$s(d) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(d_i - \overline{d}\right)^2} \qquad (2)$$

Standard deviation of the mean $s(\overline{d})$
$$s(\overline{d}) = \frac{s(d)}{\sqrt{N}} \qquad (3)$$

where N is the number of readings and d_i is the i^{th} reading.

Example: **TYPE A** evaluation of uncertainty associated with a set of data

Consider an example in which the measurand is the period T of an oscillating pendulum. A stopwatch has been used to take ten readings of the time taken for twenty oscillations of the pendulum.

Table 9.2: Data recorded using digital stopwatch

Time for 20 oscillations T_{20} (s)	Period, $T = T_{20}/20$ (s)
19.43	0.9715
20.49	1.025
20.76	1.038
20.63	1.032
20.56	1.028
21.65	1.083
20.82	1.041
19.77	0.9885
20.39	1.020
19.02	0.9510

Figure 7: A set of readings for TYPE A evaluation of uncertainty

Using Eq. (1), the **mean** $\bar{d}$ is 1.0176 s. Because this measurement involved more than one reading, the mean is the best approximation of the measurand; in this case, $T_{scatter} = 1.0176$ s.

Using Eq. (2), the **standard deviation** $s(d)$ is 0.0379 s. However, this is not the parameter used to quote the uncertainty associated with the spread in the data. For that we still need to calculate the standard deviation of the mean $s(\bar{d})$.

Using Eq. (3), the **standard deviation of the mean** $s(\bar{d})$ is 0.0119 s. This is the **standard uncertainty,** $u(T_{scatter}) = 0.012$ s.

The result of this measurement is that period T of the pendulum is (1.018 ± 0.012) s.

Note

As a general rule the uncertainty $u(T_{scatter})$ is quoted to two significant figures and the best approximation of the measurand $T_{scatter}$ is written with the same number of decimal places as the uncertainty.

TYPE B evaluation of uncertainty

It has need noted that whenever a set of data has been recorded and there is a ‘scatter’ (spread) in the set of data, a TYPE A evaluation is done; however, in every other case, a **TYPE B** evaluation of the uncertainty is applicable.

Because of the very many different circumstances under which **TYPE B** evaluations are applicable, there are many different probability density functions (pdf’s) by which it is possible to model our knowledge of a measurand. It is beyond the scope of this course to delve into the theory and use of these probability density functions, so once again, the results of the application will be used as and when necessary.

In the methods described below, we will concentrate on evaluating just two sources of uncertainty, but it is emphasised that there are many other methods that may be used. We concentrate on:

- Evaluating the act of actually reading a scale (on a dial, display or rule) as a source of uncertainty, and
- Evaluating the instrument rating as a source of uncertainty.

<u>Uncertainty associated with reading the display of a digital instrument (**TYPE B**)</u>

Consider an example in which the measurand is the mass *m* of a piece of metal. A scale with a digital readout is selected to determine the mass and when the piece of metal is placed on the scale the reading shown in Figure 8 appears on the display:

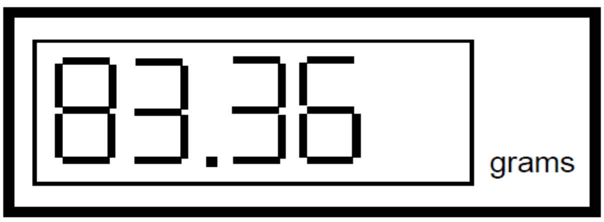


Figure 8: Reading on a digital display

This is the only reading that can be taken in this measurement process because any repeat of the procedure gives the same reading, so the displayed reading is the best approximation of the measurand; in this case 83.36 g.

As has been noted earlier, the formal way of evaluating the uncertainty associated with this reading is to identify an appropriate probability density function to model the knowledge that can be inferred about this measurand. However, the use of probability density functions is a fairly technical procedure and so we will apply a much simpler, straightforward way in which to quantify this uncertainty.

Simply consider that the value of the measurand is somewhere between 83.35 g and 83.37 g, and so the uncertainty $u\ (m)$ associated with this measurement is therefore 0.01 g. And if we considered taking of the reading as the only source of uncertainty in this measurement, we would quote the result of the process as "the mass of the piece of metal is (83.36 ± 0.01) g."

> ***Note:***
> *As a general rule, the uncertainty* $u(m)$ *should be quoted to two significant figures. However, it is acceptable where appropriate, to make a 'common sense' decision and to quote the uncertainty* $u(m)$ *to one significant figure, as in this case. Clearly the instrument cannot infer any information beyond the second decimal place.*

Uncertainty associated with reading the scale of an analogue voltmeter (**TYPE B**)

Consider an example in which the measurand is the voltage V across a torch battery. An analogue voltmeter is chosen to determine the voltage, and when the leads of the voltmeter are placed across the battery the reading shown in Figure **9** appears on the display of the voltmeter:

Figure 9: Reading on an analogue display

Once again if any repeat of the procedure gives the same outcome, we decide that the displayed reading is the best approximation of the measurand; in this case 1.55 V.

The uncertainty being evaluated is the uncertainty associated with the reading of the scale only and, once again, we will use a rule of thumb to quantify that uncertainty.

Simply consider that the value of the measurand V_{batt} is somewhere between 1.5 V and 1.6 V, and so the uncertainty $u\ (V_{batt})$ associated with this measurement is therefore 0.05 V. And if we considered taking of the reading as the only source of uncertainty in this measurement, we would quote the result of the process as "the voltage V of the torch battery is (1.55 ± 0.05) V."

> ***Caution:***
> *It would be a mistake to assume that whenever a reading is taken, that the uncertainty associated with that reading is simply half of the width of the graduation on the scale of the instrument; as you will see in the next two examples.*

Uncertainty associated with reading the scale of a pressure gauge (**TYPE B**)

Consider an example in which the measurand is a pressure p. A typical bourdon tube gauge is used to determine the pressure, and the reading shown in Figure 10 appears on the dial of the pressure gauge:

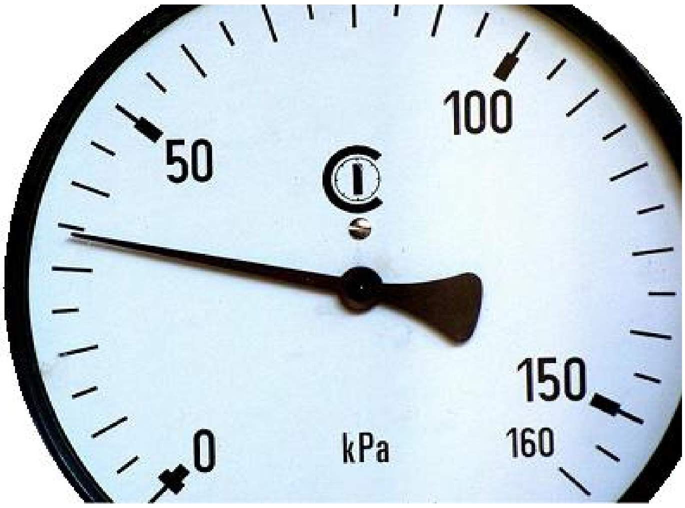


Figure 10: Reading on a pressure gauge

In this case it is necessary to exercise judgement about the reading. If this is the only reading that can be taken or any repeat of the procedure gives the same outcome, it is reasonable to say that it is the best approximation of the measurand; p_{read} = 34 kPa.

But what about the uncertainty associated with the reading? It is inappropriate simply to say that the reading is between 30 kPa and 35 kPa, which are the graduated points on the scale. It is much more reasonable to estimate that the value of the measurand is between 33 kPa and 35 kPa, which means that the uncertainty $u\ (p_{read})$ associated with this reading is 1 kPa.

When considering only the uncertainty associated with reading the dial of this pressure gauge, the pressure p may be declared to be (34 ± 1) kPa.

Uncertainty associated with reading a metre stick (**TYPE B**)

Consider an example in which the measurand is the effective length L of a pendulum. Note that the measurand is not the physical length of the pendulum, but the effective length upon which the period of oscillation of the pendulum depends.

Imagine that the pendulum is made up from a long thin rod with a pivot at one end and a ball fixed at the other. The effective length of the pendulum is given by the difference between the positions of the centre of mass of the ball at one end, and the centre of the pivot at the other, which means that it is necessary to make two position measurements.

A metre stick is selected to determine the length of the pendulum and once the zero of the metre stick is lined up with the pivot end of the pendulum, you see the following:

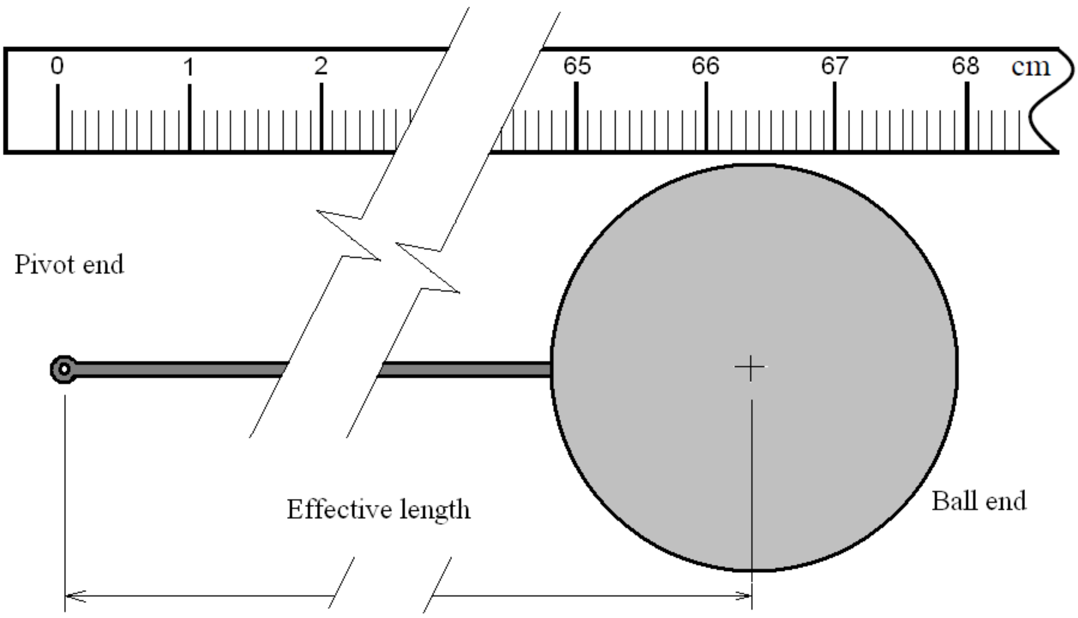


Figure 11: Reading on a metre stick

As with every analogue instrument, the user may have to deal with the question of parallax and in this case it may be a significant contributor to the uncertainty associated with the two measurements. Moreover, there is also some uncertainty as to where the centre of mass of the ball might actually be, as can be seen in Figure 11.

After taking due care, it may be decided that the best approximation of the position of the ball end is 664 mm. Now, given the practical difficulties of taking this reading, it would be unreasonable simply to assume that the position of this end is somewhere between 663 mm and 665 mm. A far more likely assessment would be that the position is somewhere between 661 mm and 667 mm, and so the uncertainty associated with this position measurement is therefore 3 mm. The result of the measurement of the position at the ball end is therefore (664 ± 3) mm.

The best approximation of the position of the pivot end is 0 mm, and a reasonable estimate of the uncertainty associated with this position measurement is 2 mm. The result of the measurement of the position at the pivot end is therefore (0 ± 2) mm.

The effective length of the pendulum is given by the difference in the position measurements of the ball and pivot ends. i.e. 644 – 0 and so the best approximation of the measurand *L* is 664 mm.

Now you have two uncertainties - both as a result of the reading of a scale – and it is necessary to fine the combined uncertainty. Combining uncertainties is dealt with in more detail in the next section of this manual, but it is important to note that these uncertainties are not added together algebraically, but in quadrature (see Section A3):

$$u_c(L) = \sqrt{3^2 + 2^2} \text{ mm.}$$

The combined uncertainty is therefore 3.6 mm and if we quote this result to one significant figure, we have a final result for the measurand of *L* = (664 ± 4) mm.

<u>Uncertainty associated with an instrument rating (**TYPE B**)</u>

Apart from the uncertainty associated with the actual readings taken from instrument scales, dials and displays, an important contributor to the combined uncertainty associated with measurement is the uncertainty associated with the **rating** of the instruments themselves.

Manufacturers of instruments that have been calibrated to some standard will always state the "**rating**" of the instrument by specifying the uncertainty associated with the instrument as a percentage of whatever reading the instrument is displaying.

The oscilloscopes that are used in the Course I laboratory have a rating of:

- 5% on the *x* –axis (when set for a time-base between 0.1 μs and 50 ms per division),
- 3% on the *y* –axis (when reading an input voltage between 5 mV and 5 V per division).

Consider an example in which the measurands are the period T and the amplitude V of a sine wave produced by a signal generator. You connect the signal generator to an oscilloscope and you see on the screen the wave shown in Figure 12:

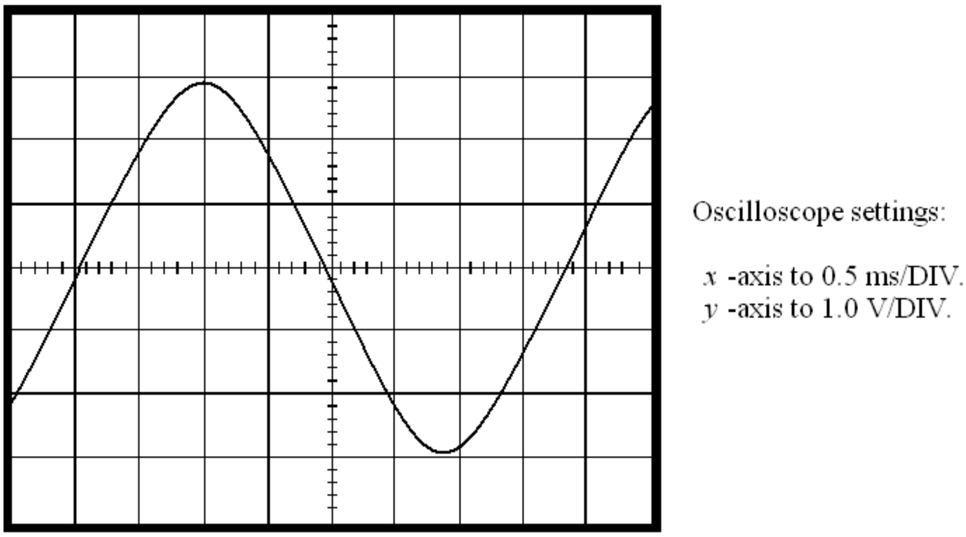


Figure 12: Reading on an oscilloscope

After careful consideration, you decide that the period spans 7.6 divisions, and so the best approximation of the period is 3.8 ms.
Likewise for the voltage, where the peak-to-peak voltage spans 5.8 divisions, the best estimate of $V_{p\text{-}p}$ is 5.8 V.

Now there are two sources of uncertainty to consider: the uncertainty associated with the reading of the display <u>as well as</u> the uncertainty associated with the instrument rating.

a) The (**TYPE B**) uncertainty associated with <u>reading</u> of the oscilloscope display: You may decide that the uncertainty associated with the reading of the display is 0.1 of a division for both the *x*- and the *y* –axis. In this case, the uncertainties associated with reading the display are:

$$u(T_{read}) = 0.1(0.5) = 0.05 \text{ ms.}$$
$$u(V_{p\text{-}p\ read}) = 0.1(1.0) = 0.1 \text{ V.}$$

b) The (**TYPE B**) uncertainty associated with the rating of the oscilloscope:
Using the specified instrument ratings, the uncertainties associated with the instrument ratings are:

$$u(T_{rating}) = 0.05(3.8) = 0.19 \text{ ms.}$$
$$u(V_{p\text{-}p\ rating}) = 0.03(5.8) = 0.17 \text{ V.}$$

The method used to **combine these uncertainties** is discussed in the next section of this guide.

A3. COMBINING UNCERTAINTIES

You will be aware by now that there may be many sources of uncertainty that may affect the result of a measurement, and this section of the guide considers how, once all the uncertainties associated with identified aspects of a measurement have been evaluated, they are merged to give a **combined uncertainty**.

Note that the combined uncertainty of a measurement is not simply the arithmetic sum of each of the uncertainties, but rather is the root of the sum of the squares of the respective uncertainties. For completeness, it is worth noting that it does not matter whether an uncertainty was initially evaluated by using a **TYPE A** or a **TYPE B** evaluation of uncertainty; they may all be combined in this way as may be required in more complex measurements.

The uncertainty budget

It is customary to show all of the uncertainties that have been identified and quantified in a measurement process in a table, which is referred to as the **uncertainty budget**, an example of which can be seen in Figure 13.

Table 9.1: Uncertainty budget for the single voltmeter reading

Source of uncertainty	**Standard uncertainty (volts)**	**Type of evaluation**
$u(V_{battery})$	0.05	Type B
$u(V_{zero})$	0.05	Type B
$u(V_{rating})$	0.031	Type B
Combined standard uncertainty: $u_c(V) = \sqrt{(0.05)^2 + (0.05)^2 + (0.031)^2} = 0.077$ volts.		

Figure 13: Example of an uncertainty budget

The data shown in Figure 13 comes from the example that was used to illustrate a **TYPE B** evaluation of the uncertainty associated with the reading of the voltmeter shown in Figure 9. In that case, the best approximation of the measurand V_{batt} was 1.55 V, and the uncertainty $u\,(V_{batt})$ associated with the reading of the scale for V_{batt} was 0.05 V.

Added to the uncertainty budget, as shown in Figure 13, is the uncertainty associated with the 'zero' reading $u\,(V_{zero})$, which is assumed to be the same as that for any other reading on the instrument, as well as the uncertainty associated with the rating of the instrument. In this case, the manufacturer of the voltmeter had specified the rating of the instrument to be 2%, which is a typical rating for moving coil instruments, so the uncertainty $u\,(V_{rating})$ associated with the rating would be:

$$u\,(V_{rating}) = (0.02)(1.55\text{ V}) = 0.031\text{ V}.$$

General formula for combining uncertainties

Where each of the uncertainties associated with a measurement process to determine a measurand have been identified and evaluated, the general equation for combining the uncertainties is as follows:

$$u_c(k) = \sqrt{\left[u_1(k)\right]^2 + \left[u_2(k)\right]^2 + \left[u_3(k)\right]^2} \qquad (4)$$

where the measurand is some quantity k, and the sources of uncertainty were identified and quantified giving $u_1(k)$, $u_2(k)$ and $u_3(k)$. The **combined uncertainty** is $u_c(k)$.

It should be noted that for any one measurement process, there is no limit to the number of sources of uncertainty that may be identified, evaluated, and then combined in this way.

SECTION B - ANALYSIS

B1. DRAWING GRAPHS

Linearising equations to simplify graphs

One of the aims in investigating physical phenomena is to establish the relationship between the variables that are being measured.

For example:

> If we were investigating an object that is experiencing uniform acceleration a then we will find the function describing the relationship between velocity v and time t to be a straight line of the form $y = mx + c$.
>
> In this case $v = at + u$ where u is the initial velocity.
>
> On the other hand, the relationship between the position s and the time t will be of the form $y = a x^2 + b x + c$, i.e. a parabola.
>
> In this case $s = ½\, at^2 + ut + s_0$ where s_0 is the initial position.

Clearly, the straight line is the simplest form to analyse. It is much more difficult to be sure that a particular shape is indeed a parabola, a hyperbola, a log or an exponential - and then to find the constants that will allow us to write an equation relating the variables – than it is to analyse a straight line.

For this reason we try to plot the recorded data in the form of a straight line, a process known as “linearising”. We then use this “linearised” plot to determine the dependence of the variables on each other.

For example, the quadratic equation

$$s = ½\, at^2 + ut,$$

is easily linearised by dividing throughout by t ($t \neq 0$) to give

$$s/t = ½\, at + u.$$

General rules for plotting a graph

- Use a pencil!
- When asked to plot the graph of say ***P*** vs ***Q***, it means that ***P*** must be on the vertical axis (y - axis) and ***Q*** must be on the horizontal axis (x - axis).
- If it has not been specified as to which variable should be on which axis (as in the case of "plot ***P*** vs ***Q***") then the dependent variable should be on the vertical axis (y - axis) and the independent variable should be on the horizontal axis (x - axis).
- Each graph must have a caption (title) describing why you plotted the graph.
- Each axis should be labelled with the name of the variable and the units.
- Choose appropriate scales on the axes so that the graph will not be too small, but will cover a fair portion of the page in each direction.
- Use a ⊙ or × for the data points, and not a ● (blob).
- Decide whether or not each axis should start from zero (it is not always necessary to show the origin).
- Axes should be marked in scaling factors of 1, 2, 5, or these to powers of ten. Scaling factors such as 3 or 4 usually make the scales difficult to read and should be avoided.

Drawing a "Best Fit" by hand

You will often be asked to draw the line of "best fit" by hand. The line that you have to draw is your best, reasoned guess by eye as to what the "best fit" may be.

An example is shown in Figure 14.

Remember, the "best fit" has to model the trend of the data points; your graph should not join the points! The "best fit" may not even go through any of the data points.

Don't force your line through the origin if you think that the intercept should be zero – remember that it's a best fit to all the data points – and if the origin is not a data point, it shouldn't affect your choice of the "best fit" at all.

> ***Important****:*
> *When determining the slope m of the "best fit" line, do not use data points to find Δy and Δx, but choose co-ordinates on the line to do the calculation. Remember that what you need to find the slope of the "best fit" line, not the slope of a line between a pair of specific data points!*

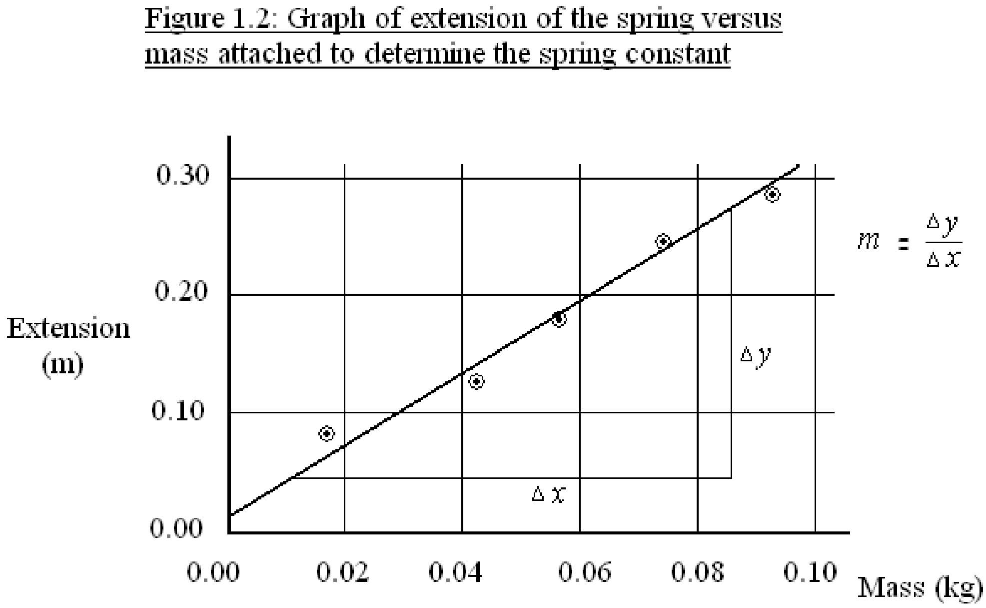


Figure 14: Example of a graph with a "best fit" line

Uncertainty associated with the "best fit" by eye

It is important to note that when a "best fit" is done by eye, there is a simple way to get a rough estimate of the uncertainty associated with the line you have chosen. Simply consider what the values of *m* and *c* would be if you had chosen reasonable "minimum" and "maximum fits", as illustrated in Figure 15.

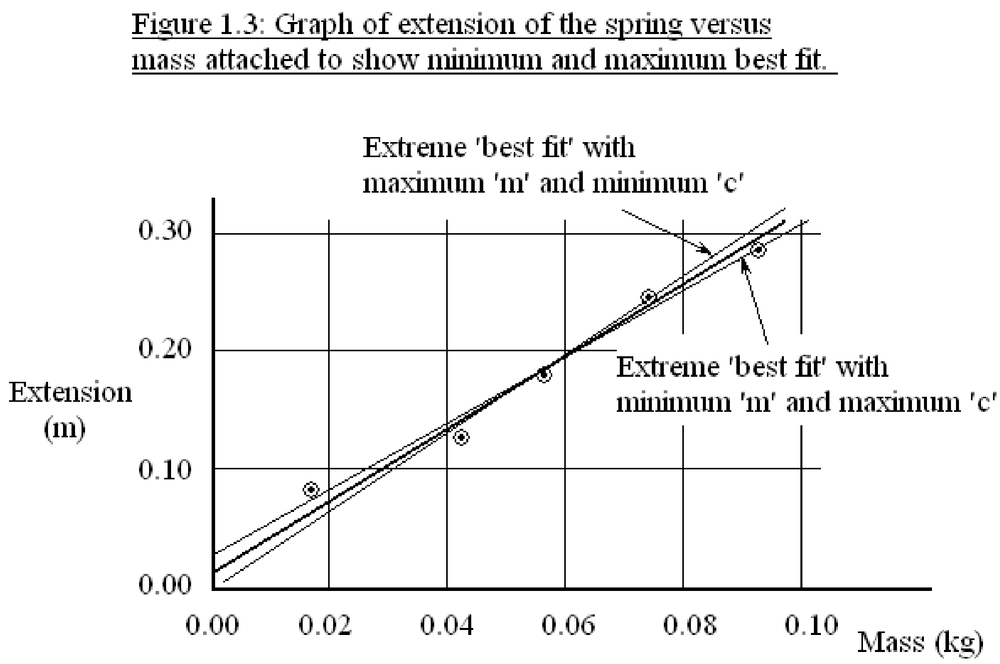


Figure 15: Estimating uncertainty associated with the "best fit"

To evaluate the uncertainty associated with the 'm' which was determined from the best fit line, we will use a simple rule of thumb. Determine the extreme minimum 'm_{min}' and the extreme maximum 'm_{max}' from the graph, and then let the uncertainty $u(m)$ associated with the best fit be equal to half the difference between the two extremes:

$$u(m) = ½ \, | m_{max} - m_{min} | \, .$$

B2. METHOD OF LEAST SQUARES

Since straight line graphs are convenient and important in data analysis, there is a better way to find the slope m and the intercept c of a line that is the "best fit" of a set of data. The procedure for calculating these values is the method of least squares. The method of least squares provides the values for the slope m and the intercept c, as well as the associated uncertainties $u(m)$ and $u(c)$.

Given a set of data points (x_i, y_i) that represent a straight line relationship, the slope and the intercept of the "best fit" line can be determined by the following equations:

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

$$c = \frac{\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} ,$$

and the uncertainty associated with m and c, $u(m)$ and $u(c)$ respectively can be determined:

Where $d_i = y_i - \left(m x_i + c\right)$,

$$u(m) = \sqrt{\frac{\sum d_i^2}{n\sum x_i^2 - \left(\sum x_i\right)^2} \left(\frac{n}{n-2}\right)}$$

$$u(c) = \sqrt{\frac{\sum d_i^2 \sum x_i^2}{n\left(n\sum x_i^2 - \left(\sum x_i\right)^2\right)} \left(\frac{n}{n-2}\right)} .$$

A computer is generally used to do the least squares calculations and this facility is available on computers in the Physics Course I Laboratory. When a computer is not available, it is easy to use your calculator and the equations given above. Section D of this guide also gives the instructions for some common calculators to perform these calculations.

B3. UNCERTAINTY WHEN RESULTS ARE CALCULATED

Often, having completed two (or more) measurement processes, where the results are, say $A \pm u(A)$ and $B \pm u(B)$, you need to use the results of the measurements to calculate some other value, say R and its associated uncertainty $u(R)$.

We will not derive the general formulae given in Figure 16 for dealing with this problem as they involve the use of partial derivatives which you will meet in your mathematics course. We do caution however, that the uncertainty $u(R)$ of the **propagated** (calculated) value of R is <u>not</u> simply the sum of the uncertainties $u(A)$ and $u(B)$.

Form of equation from which result R is calculated	*Formula for calculating the standard uncertainty u(R)*
$R = aA \pm bB \pm c$	$\left[u(R)\right]^2 = \left[a\,u(A)\right]^2 + \left[b\,u(B)\right]^2$ or $u(R) = \sqrt{\left[a\,u(A)\right]^2 + \left[b\,u(B)\right]^2}$
$R = cA^aB^b$	$\left[\frac{u(R)}{R}\right]^2 = \left[a\frac{u(A)}{A}\right]^2 + \left[b\frac{u(B)}{B}\right]^2$ or $u(R) = R\sqrt{\left[a\frac{u(A)}{A}\right]^2 + \left[b\frac{u(B)}{B}\right]^2}$

(Coefficients *a, b & c* are numbers with zero uncertainty)

Figure 16: General formulae used in propagation of uncertainties

Note that it is sometimes necessary to consider the effect of correlation of uncertainties – as when the same instrument is used twice in the same experiment – before propagating an uncertainty as a result of a calculation. Ask the lab demonstrators about this.

B4. COMPARING RESULTS

The results of any two measurement processes can only be meaningfully compared if the intervals associated with each of the measurements are known. More specifically:

- If the intervals that represent the results of two measurement processes overlap, then the two results "**agree within experimental uncertainty**."
- If the intervals that represent the results of two measurement processes do not overlap, then the results "**do not agree within experimental uncertainty**."

For example, three groups in the laboratory measure the period T of a pendulum. The result, i.e. the best estimate and the standard uncertainty, quoted by each group is as follows:

$$T_1 = (5.73 \pm 0.41) \text{ s}$$
$$T_2 = (5.62 \pm 0.10) \text{ s}$$
$$T_3 = (6.28 \pm 0.25) \text{ s}$$

It is a good idea to then represent the three measurements in the form of intervals on a number line:

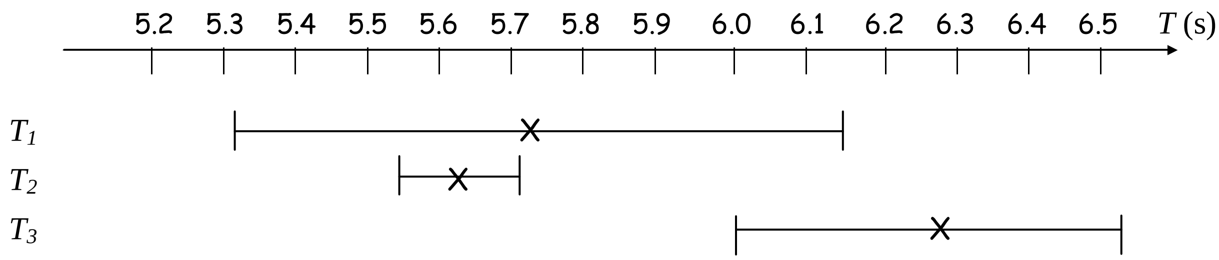


Figure 17: Comparison of results

The intervals associated with T_1 and T_2 overlap, and therefore it can be said that "these two results agree within experimental uncertainty". The same can be said about the intervals associated with T_1 and T_3 (i.e. that they overlap and the results agree).

However, the interval associated with T_3 does not overlap with the interval associated with T_2 and therefore these results "do not agree within experimental uncertainty."

> ***Note:***
> *This is the only meaningful way to compare two measurements or the results of calculated measurands!*

If you do not know the interval associated with each of two measurement processes, it is not possible to compare the results, no matter how "close" or "far" they seem to be to each other. **To say that two things are "nearly the same" is meaningless in the context of laboratory work!**

Comparing experimental results with "book" values

You may have occasion to decide whether or not your measured value for a quantity agrees with the generally accepted "book" value. For example, in an experiment you may have measured the Universal Gas Constant R to be 8.341 ± 0.054 J K^{-1} mol^{-1}. You look in a textbook and see that the value given there is:

$$8.31451(72)\ \mathrm{J\ K^{-1}\ mol^{-1}}.$$

The seven and two in brackets indicate that the last two digits (51) have an uncertainty of 72 associated with them; in other words, the "book" value is:

$$(8.31451 \pm 0.00072)\ \mathrm{J\ K^{-1}\ mol^{-1}}.$$

It is clear that the accepted "book" interval overlaps the interval that is the result of your measurement process, therefore you can say that your "result agrees with the accepted "book" value within experimental uncertainty".

SECTION C - REPORTING

C1. COMPONENTS OF THE LABORATORY REPORT

There are many acceptable ways to structure a laboratory report and each possible component of a report listed below need not have its own heading - some components can be combined with others – but all of these components should be covered in the report.

- **Title, Author & Partners (where applicable) and Date**
- **Aim**: Prepares the reader for what he or she is about to read.
- **Introduction and Theory**: Provides the theoretical underpinning of the method used and the underlying physics.
- **Apparatus**: Describes the apparatus used.
- **Method**: Describes the procedures followed.
- **Data**: Presents the readings taken - generally in tables - without analysis.
- **Graphs and "Best Fit" Line**: Presents information derived from the data.
- **Calculations**: Shows all relevant calculations to determine the result.
- **Uncertainty Analysis:** Shows the calculation of the uncertainty in the result.
- **Interpretation, Discussion and Conclusion**: Explains the findings.

C2. WRITING THE LABORATORY REPORT

<u>Purpose of the report</u>

The purpose of a laboratory report is to communicate the aim, process and result of an investigation to an outside audience. It is a record of your direct ("hands-on") experience in the laboratory. In most cases, a scientific investigation is considered to be incomplete without a report.

By synthesising (putting together) the different aspects of your laboratory experience in a structured and coherent report, the essence of your investigation becomes clearer in your own mind. In the process, you develop your skills of reasoning and ability to communicate in writing.

Title of the report

The title must be short but factual and descriptive. It must summarise the major aspects to be dealt with in the report. The key words used in the title will often come from the laboratory task that has been set and you need to identify these.

Aim

Often you will be asked to carry out an investigation in order to determine some quantity or to investigate some behaviour. If that is the case, state as clearly and as concisely as possible what it is that is to be determined, or what behaviour is to be investigated.

Introduction and Theory

This component of the report is to give the reader the theoretical underpinning of the method used in the experiment. Any relevant theory, usually expressed in the form of equations, should be included in this section. Note that the theory given here lays the foundation for the later interpretation, analysis and discussion of the data that was collected during the experiment.

Apparatus

This component of the report is to give the reader a clear, concise description of the apparatus used in the experiment. Make neat, labelled sketches in which salient features of the apparatus are highlighted. Where appropriate, state what the function of the apparatus may be. It is not necessary to show minor details like knobs on instruments and the precise shapes of ancillary boxes and stands.

Method

This component of the report briefly describes, in your own words, the procedures that were actually followed when doing the experiment. (**Do not simply rewrite the laboratory instructions in past tense!**) Note that this includes the procedures used to obtain data as well as to analyse that data. Wherever appropriate, give a reason for each step you took in a procedure, e.g. if your aim is to measure the force constant of a spring you could include a statement of the form "*a graph of* T *vs* $\sqrt{M}$ *was plotted* *and the force constant of the spring was determined from the slope of the graph"*. Do not omit any significant steps and note that the method as you have described it lays the foundation for your discussion and analysis that is to follow.

- It is preferable to use an impersonal style when writing the method: i.e. "Five measurements were recorded", and not "I took five measurements".
- Write in the past tense: i.e. "This was done, then that was done" <u>not</u> "Do this, then do that", as if you were presenting a recipe.

<u>Data</u>

This component of the report records a key part of the experiment and careful planning is needed to present the data clearly and sensibly. The data is usually presented, without comment or analysis, in tables. Tables are to be numbered in sequence e.g. Table 1, Table 2, etc., and each table has to have a title (e.g. *Table 1: Measurements of the width of the cylinder)*.

<u>Calculations</u>

This component of the report shows any relevant calculations that you have made. These may be calculations required to generate information to be graphed or recorded, or they may be calculations done to extract information from data. In any event, the nature and purpose of doing the calculation should be made clear to the reader. When doing calculations, remember to quote appropriate units and significant figures.

<u>Graphs and "Best Fit" Line</u>

A picture tells a thousand words, and this is it. This component of the report is central to the successful communication of your findings. Graphs are to be numbered in sequence e.g. Graph 1, Graph 2, etc., and each graph has to have a caption which describes the purpose for which it has been presented (e.g. *Graph 3: To determine the viscosity of the sample of oil*).

<u>Uncertainty Analysis</u>

The determination of the uncertainties associated with measurement and the derived results is a critical aspect of laboratory work. Note that the concepts in this section are not simple and it will take more than a cursory glance at this manual every so often to come to terms with this work.

Interpretation, Discussion and Conclusion

These components of the report are used to present and discuss your results. In the interpretation and discussion, you are expected to draw all the components of the report together.

For example, if your aim was to measure *g* you should state the outcome of the experiment clearly such as, '*the measurand g was determined to be (9.790 ± 0.052) m* s^{-2}", and then discuss how this result compares with the expected "book" value.

You may give explanations for your results but be careful not to mix facts with opinions. Avoid meaningless phrases such as "This was a successful experiment" and "... was caused by human error."

Finally, you are required to comment on the factors that may have contributed to the uncertainty in the quoted result, and you need to suggest changes that could be made so as to improve the outcome of the experiment.

C3. TABULATING DATA

Whenever you take a set of readings, record those readings in a table. Here are key points to consider when doing so:

(i) Plan your table, especially how many columns and rows you will need. Remember to include columns for the results of calculations. You may also sometimes repeat a reading a number of times and then calculate the mean of those readings.

(ii) Each table must have a title which should reflect your reason for recording the readings. If you use more than one table in your report, number them Table 1, Table 2, etc.

(iii) Each column must have a heading and appropriate units. The units should be included with the heading and not written alongside each reading in the table.

(iv) Record the readings carefully. Remember, for example, that writing 36.0 cm and 36 cm is not the same (See RECORDING SIGNIFICANT FIGURES). In the first case you are reading to the nearest millimetre, while in the second case you are only reading to the nearest centimetre.

Title → Table 1: Readings taken to determine the spring constant

Headings with units →

Applied force, F (N)	Extension of spring, x (mm)			$\bar{x}$ (mm)	Spring constant $k = F / x$ (N m^{-1})
0.100	41	41	42	41.3	2.42
0.201	82	83	82	82.3	2.44
0.298	126	129	128	127.7	2.33
0.400	160	159	163	160.7	2.49

Readings   Calculations

Figure 18: Example of a table

C4. RECORDING SIGNIFICANT FIGURES

What are Significant Figures?

The number of figures that convey meaningful information (in a recorded datum) is known as the "**number of significant figures"** of that datum.

Significant figures in a number are:

- All non-zero digits (e.g. 54 has two significant digits while 456.78 has five significant digits).
- All zeros appearing anywhere between two non-zero digits (e.g. 302.05 has five significant digits).
- All trailing zeros, where the number has a decimal point (e.g. 36.65000 has seven significant digits).
- Leading zeros are not significant (e.g. 0.0003 has only one significant figure).

Do not confuse significant figures with decimal places. The number of decimal places refers to the number of figures after the decimal point. (E.g. 46.320 has five significant figures, but three decimal places; while 0.0040 has two significant figures, but four decimal places.)

Rounding off significant figures

You will often need to round off your calculations to the correct number of significant figures. The general rules for rounding off are:

- The last significant figure to be retained remains unaltered if the next digit is less than 5 (e.g. 3.434 rounds off to 3.43).
- The last significant figure to be retained is increased by one if the next digit is greater than 5 (e.g. 3.436 rounds off to 3.44).
- The last significant figure to be retained may remain unaltered or be increased by one if the next digit is equal to 5. Note that this is not a hard and fast rule and you are free to choose whichever convention you prefer. One choice is to round in such a way that the last digit to be retained is even.
- Don't do a double round off (e.g. rounding 3.4348 off to three significant figures becomes 3.43; don't round 3.4348 to 3.435 and then to 3.44).

Recording significant figures

When taking a reading (datum), the number of significant figures to be recorded is determined by the precision of the instrument and the ability of the observer to use the instrument. You should strive to record as many significant figures as possible, but it is incorrect to record a reading to more significant figures than can be determined from the instrument.

For example, if the scale on a particular instrument makes it possible to read a mass to 1/10th of a gram, then it is incorrect to record the reading of some mass to 1/100th of a gram. For example, a recording of say 12.4 g on such an instrument may be correct, but a recording of 12.40 g taken from that instrument is certainly incorrect! Such a recording is beyond the capability of the instrument.

Similarly when using a metre rule to measure distance, you can only take readings to the nearest millimetre. For example, you may record (correctly) the reading of a distance as say 127 mm, but a recorded reading of 127.0 mm will be incorrect because that degree of precision cannot be achieved when using a meter rule.

Avoid creating false precision through calculation

It will sometimes happen that you have to use readings recorded in a table in a calculation. The answer that you get on the calculator will generally give many more significant figures and/or decimal places than the numbers you used in the first place. This is known as creating 'false precision' through calculation.

When doing such a calculation, you should be aware that it is incorrect to generate a number that has more significant figures and/or decimal places than those with which you started out. Note that this is not applicable when you have to calculate the average of a number of readings! An average may have more significant figures than the readings (see Figure 18 for example).

For example, if you were to combine two readings in a calculation where one reading is recorded to say three significant figures and two decimal places, while the other is recorded to say six significant figures and one decimal place, the answer that you get from the calculation must be rounded off to the worst case. See examples that follow.

Examples:

Addition (or subtraction):

If we add a mass of 2004 kg (i.e. read to the nearest 1 kg) to a mass of 83.61 kg (i.e. read to the nearest 0.01 kg), then:

	2.004×10^3 kg
	0.08361×10^3 kg
sum:	2.088×10^3 kg

Note that the precision of the result of the calculation cannot be better than the precision of the least precise reading involved (i.e. the 2004 kg mass measured to the nearest 1 kg in this case).

Multiplication (or division):

Suppose we obtain readings for the lengths of the sides of a rectangle and find them to be 10.2 cm and 15.6 cm respectively. Then the area of the rectangle is $10.2 \times 15.6 = 159.12$ cm^2. But are all five figures of the answer significant? Can we multiply two numbers, each having three significant figures, to obtain an answer with five significant figures?

Clearly this cannot be the case. The convention when multiplying numbers of different significant figures is that the result should be quoted with the number of significant figures equal to the number with the least significant figures (i.e. 159 cm^2 in this case).

- If we have calculated a velocity from a time reading of 2.357 s (4 significant figures) and a distance reading of 25.3 m (3 significant figures) then our answer must be quoted as 10.7 m s^{-1} (3 significant figures).

- Multiplying two readings, 1.23×10^3 and 1983.3, gives 2.44×10^6 (try it).

C5. QUOTING A RESULT

When quoting the result of a measurement process, you have to provide the following information:

- The measurand (what was measured?),
- The best approximation of the measurand,
- The standard uncertainty,
- The units of the measurand,
- The level of confidence with which the result is quoted.

<u>Quoting the result using one standard deviation</u>

In general, when one standard deviation is used to calculate the uncertainties in a measurement process, there is a probability of about 68% that the value of the measurand exists within the interval $A \pm u(A)$, where A is the best estimate of the value of the measurand and $u(A)$ is the combined standard uncertainty.

For example, the result of a measurement in which the measurand was the mass of a piece of metal may be quoted as:

$$m = (83.45 \pm 0.34)\ \text{g} \quad \text{(with a 68\% level of confidence)},$$

meaning that the best estimate of the mass of the piece of metal was determined to be 83.45 g with a standard uncertainty of 0.34 g and a level of confidence of 68%.

In other words, there is a 68% probability that the mass of metal falls in the interval between 83.11 g and 83.79 g.

> "… the mass of the piece of metal was determined to be (83.45 ± 0.34) g with a level of confidence of 68%".

<u>Quoting the result using two or three standard deviations</u>

In this course we will quote all results using one standard deviation, but it is common practice for laboratories to quote results using two or even three standard deviations, in which case the quoted level of confidence will be about 95% and 99% respectively.

For example, if a result is quoted with a level of confidence of 95%, then you know that the uncertainty was evaluated by using two standard deviations.

Quoting the number of significant figures in the uncertainty

It is common practice to quote the uncertainty in the result of a measurement to **two significant figures.** However, each of the significant figures quoted in the result should carry a practical meaning in that they have to convey to the reader a sense of the care with which the measurement has been conducted.

In your Course I laboratory work you are therefore expected to quote the uncertainties in the results of your measurements to two significant figures unless doing so would create the impression that the result of the measurement is more precise than in fact it is (e.g. TYPE B evaluations from analogue scales).

Finally, once you have rounded the standard uncertainty off to two significant figures (or one significant figure as the case may be), you have to round off the best estimate to the same number of **decimal places** as the uncertainty.

Quoting the result using fractional (%) uncertainty

It may be convenient on occasion to express the standard uncertainty $u(A)$ as a fraction of the best approximation of the measurand A. Thus, we may sometimes refer to the **fractional uncertainty** given by $u(A) / A$ or to the **percentage uncertainty** given by $(u(A) / A) \times 100$ %.

For example, if the quoted result of a measurement is (3.60 ± 0.32) units, then the fractional uncertainty is 0.32 / 3.60 = 0.09 and the percentage uncertainty is 9%.

This form of quoting uncertainty is commonly used by manufacturers of electronic components where the uncertainty in the quoted value is called the "tolerance" and it is quoted as a percentage of the nominal value of the component. For example, if the value of a resistor is given as 220 Ω with a tolerance of 5%, it means that the value of the resistor is (220±11) Ω.

> ***Caution:***
> *This form is quoting the result is not always valid and user discretion is advised. Say you have completed a measurement and have determined the result to be (0.00 ±0.15) units. Then to quote the result using fractional uncertainty would be quite meaningless.*

SECTION D – USING COMPUTERS AND CALCULATORS

D1. USING EXCEL

(All of the following instructions are based on Excel 2003, as running in the Course I Lab.)

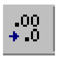
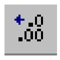
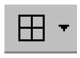
Tables with Excel

Initiate Excel by clicking on the Excel shortcut on the computer's desktop. If you want to change the width and height of cells, highlight the cells and then use **FORMAT**, **ROW**, **HEIGHT** and **FORMAT**, **COLUMN**, **WIDTH** to change their sizes.

To add headings (and units) to columns, click on the cell and TYPE as you would normally do. Unfortunately one can't produce superscripts and subscripts in Excel, so one needs to use the ^ sign to indicate power of a unit, and the _ sign to denote a subscript. Your data is entered in the same way: just click on the cell and enter the value.

If you want to get Excel to do a calculation just click on the cell where you want the answer to go, and TYPE = followed by the required mathematical expression. For example, if the value in the particular cell is obtained by multiplying the value in cell C3 by that in D3 just TYPE **= C3*D3** (in fact, you can just TYPE **=,** then click on cell C3, TYPE *****, click on D3 and hit the enter key). In Excel, the * symbol is used to indicate multiply, the / symbol divide and the ^ symbol "raise to the power".

It is very easy to apply the same formula to the rest of the cells in a column; just select the cell with the formula in it, click on the bottom right hand corner and drag it down the column. Excel will adjust cells referred to in copied formulae by incrementing their row numbers and column letters as they are copied down rows and across columns respectively. In most cases this is exactly what one wants, but in some cases one wants the same cell referred to in the calculations of an entire row or column. If this is the case, then use a $ sign (e.g. E5 included in a formula is left unchanged when copied to a new cell).

Remember to take care to display the correct number of significant figures in your data. Since Excel calculates the answer to as many places as can fit in the cell, you often need to adjust the number of decimal places in columns to reflect the correct number of significant figures. Decreasing the number of decimal places is easiest achieved using the button marked  ; increasing by that marked . To add gridlines or borders to your table use the button marked .

Plotting with Excel

When you are ready for Excel to plot your data you need to highlight the two columns of data you wish to plot and use **INSERT**, **CHART** to call up the chart (plotting) procedure. Choose the **XY (SCATTER)** option and then select the chart sub-TYPE option which allows you to plot the points without lines. Then click on **NEXT**, and **NEXT** again until you bring up the CHART OPTIONS window. You can use the **TITLE** option to write a title and label the *x*- and *y*-axes, the **GRIDLINE** option to switch off the gridlines and the **LEGEND** option to switch off **SHOW LEGEND**. Having done this, click on **FINISH**.

Move your plot into the desired position by clicking just inside the borders and dragging it to where you want it positioned. Expand it by pulling on the small black squares in the middle of the borders.

"Best fit" with Excel

To draw the best straight line through your points click somewhere inside your graph to select it, then click on **CHART, ADD TRENDLINE** and select the **LINEAR** option. If you need to know the slope and intercept of this straight line click on **OPTIONS** and check the **DISPLAY EQUATION ON CHART** box. Before plotting your graph it will probably be best to remove the grey background – easiest done by clicking twice somewhere on the background and either choosing the colour white, or choosing "none".

The final alteration you may wish to attempt is to change the origin and scale of the plot so that your graph covers most of the page. To change the scale on the x-axis, click on the axis to select it. Next click on **FORMAT, SELECTED AXIS** and choose the **SCALE** option to change the minimum for your *x*-axis. Repeat this exercise for the *y*-axis if desired.

D2. USING CURFIT

CURFIT can be used to calculate the slope and intercept of the best fit straight line through data, together with their uncertainties. To start the program, click on the CURFIT shortcut on the desktop, as you do with any program running on Windows. The program will open (see Figure 19). Click **OK** and TYPE in the **X** and **Y** values in the appropriate columns, as shown in Figure 20.

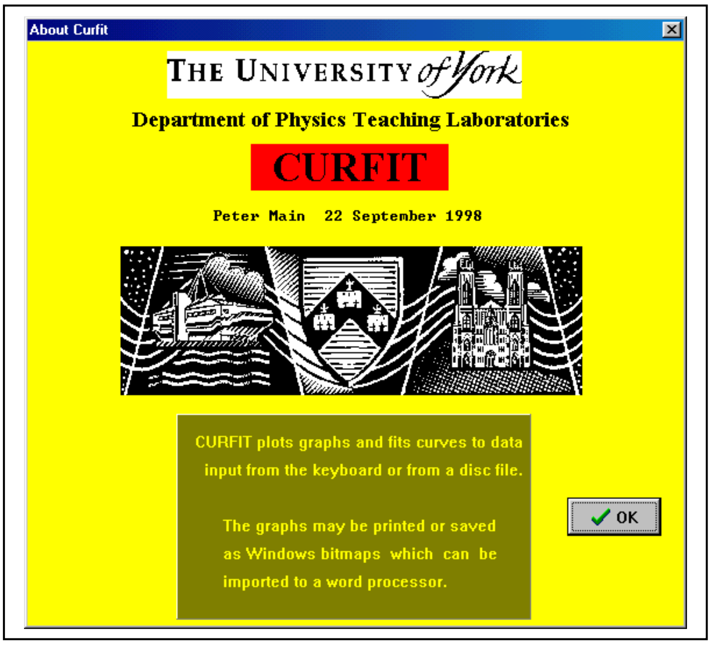


Figure 19: Title screen of CURFIT

Select **STRAIGHT LINE** on **FIT CURVE TO DATA**; a small window will open.
In the small window, select **CALCULATE SLOPE AND INTERCEPT** (default option) and click **OK** on the small window; the small window will close.

Click on **DISPLAY GRAPH** to see the graph and **DISPLAY PARAMETERS** to get the value of the uncertainties in the slope and intercept (see Figure 20).

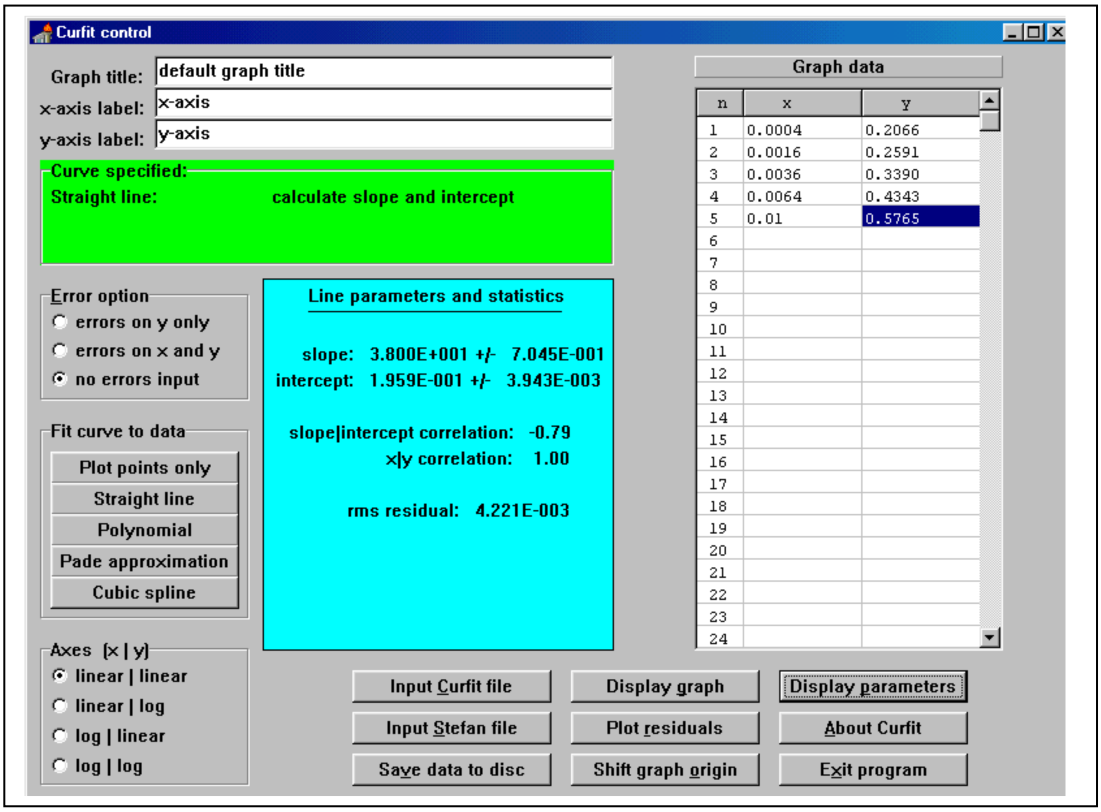


Figure 20: Data input and fitting screen of CURFIT

D3. USING CALCULATORS

SHARP
(compiled from SHARP EL-531WH User Manual
with Course I Lab Comments in Shadowed Text Boxes)

STATISTICAL CALCULATIONS [13]

Statistical calculations are performed in the statistics mode. Press [MODE] [1] to select the statistics mode. This calculator performs the seven statistical calculations indicated below. After selecting the statistics mode, select the desired sub-mode by pressing the number key corresponding to your choice.
When changing to the statistical sub-mode, press the corresponding number key after performing the operation to select the statistics mode (press [MODE] [1]).

select SD

- [0] (SD) : Single-variable statistics
- [1] (LINE) : Linear regression calculation
- [2] (QUAD) : Quadratic regression calculation
- [3] (EXP) : Exponential regression calculation
- [4] (LOG) : Logarithmic regression calculation
- [5] (PWR) : Power regression calculation
- [6] (INV) : Inverse regression calculation

The following statistics can be obtained for each statistical calculation (refer to the table below):

Single-variable statistical calculation

Statistics of ①

Linear regression calculation

Statistics of ① and ② and, in addition, estimate of y for a given x (estimate y') and estimate of x for a given y (estimate x')

Exponential regression, Logarithmic regression, Power regression, and Inverse regression calculation

Statistics of ① and ②. In addition, estimate of y for a given x and estimate of x for a given y. (Since the calculator converts each formula into a linear regression formula before actual calculation takes place, it obtains all statistics, except coefficients a and b, from converted data rather than entered data.)

Quadratic regression calculation

Statistics of ① and ② and coefficients a, b, c in the quadratic regression formula ($y = a + bx + cx^2$). (For quadratic regression calculations, no correlation coefficient (r) can be obtained.) When there are two x' values, press [2ndF] [←→].
When performing calculations using a, b and c, only one numeric value can be held.

①	$\bar{x}$	Mean of samples (x data)
	sx	Sample standard deviation (x data)
	σx	Population standard deviation (x data)
	n	Number of samples
	Σx	Sum of samples (x data)
	Σx^2	Sum of squares of samples (x data)
②	$\bar{y}$	Mean of samples (y data)
	sy	Sample standard deviation (y data)
	σy	Population standard deviation (y data)
	Σy	Sum of samples (y data)
	Σy^2	Sum of squares of samples (y data)
	Σxy	Sum of products of samples (x, y)
	r	Correlation coefficient
	a	Coefficient of regression equation
	b	Coefficient of regression equation
	c	Coefficient of quadratic regression equation

- Use [ALPHA] and [RCL] to perform a STAT variable calculation.

Data Entry and Correction [14]

Entered data are kept in memory until [2ndF] [CA] are pressed or mode selection. Before entering new data, clear the memory contents.

[Data Entry]

Single-variable data
Data [DATA]
Data [(x,y)] *frequency* [DATA] (To enter multiples of the same data)

Two-variable data
Data x [(x,y)] *Data y* [DATA]
Data x [(x,y)] *Data y* [(x,y)] *frequency* [DATA] (To enter multiples of the same data x and y.)

- Up to 100 data items can be entered. With the single-variable data, a data item without frequency assignment is counted as one data item, while an item assigned with frequency is stored as a set of two data items. With the two-variable data, a set of data items without frequency assignment is counted as two data items, while a set of items assigned with frequency is stored as a set of three data items.

[Data Correction]

Correction prior to pressing [DATA] immediately after a data entry:
Delete incorrect data with [ON/C], then enter the correct data.

Correction after pressing [DATA]:
Use [▲] [▼] to display the data previously entered.
Press [▼] to display data items in ascending (oldest first) order. To reverse the display order to descending (latest first), press the [▲] key.
Each item is displayed with 'Xn=', 'Yn=', or 'Nn=' (n is the sequential number of the data set).
Display the data item to modify, input the correct value, then press [DATA]. Using [(x,y)], you can correct the values of the data set all at once.

- When ▲ or ▼ appears, more data items can be browsed by pressing [▲] or [▼].
- To delete a data set, display an item of the data set to delete, then press [2ndF] [CD]. The data set will be deleted.
- To add a new data set, press [ON/C] and input the values, then press [DATA].

***sx* corresponds to standard deviation**

[15]

$$\bar{x} = \frac{\Sigma x}{n} \qquad \sigma x = \sqrt{\frac{\Sigma x^2 - n\bar{x}^2}{n}}$$

$$sx = \sqrt{\frac{\Sigma x^2 - n\bar{x}^2}{n-1}}$$

$$\Sigma x = x_1 + x_2 + \cdots + x_n$$
$$\Sigma x^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

$$\bar{y} = \frac{\Sigma y}{n} \qquad \sigma y = \sqrt{\frac{\Sigma y^2 - n\bar{y}^2}{n}}$$

$$sy = \sqrt{\frac{\Sigma y^2 - n\bar{y}^2}{n-1}}$$

$$\Sigma xy = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$
$$\Sigma y = y_1 + y_2 + \cdots + y_n$$
$$\Sigma y^2 = y_1^2 + y_2^2 + \cdots + y_n^2$$

Examples

【13】 [DATA] [(x,y)] [$\bar{x}$] [Sx] [σx] [n] [Σx] [Σx²] [$\bar{y}$] [Sy] [σy] [Σy] [Σy²] [Σxy] [r] [a] [b] [c] [x′] [y′] [←→]

DATA		
95	[MODE] [1] [0]	*0.*
80	95 [DATA]	*1.*
80	80 [DATA]	*2.*
75	[DATA]	*3.*
75	75 [(x,y)] 3 [DATA]	*4.*
75	50 [DATA]	*5.*
50		
$\bar{x}$=	[RCL] [$\bar{x}$]	*75.71428571*
σx=	[RCL] [σx]	*12.37179148*
n=	[RCL] [n]	*7.*
Σx=	[RCL] [Σx]	*530.*
Σx²=	[RCL] [Σx²]	*41’200.*
sx=	[RCL] [Sx]	*13.3630621*
sx²=	[x²] [=]	*178.5714286*
$\frac{(95-\bar{x})}{sx}\times10+50=$	[(] 95 [−] [ALPHA] [$\bar{x}$] [)] [÷] [ALPHA] [Sx] [×] 10 [+] 50 [=]	*64.43210706*

【14】 [DATA] [▲] [▼]

DATA		
30	[MODE] [1] [0]	*0.*
40	30 [DATA]	*1.*
40	40 [(x,y)] 2 [DATA]	*2.*
50	50 [DATA]	*3.*

↓

DATA		
30	[▼] [▼] [▼]	
45	45 [(x,y)] 3 [DATA]	X2 = *45.*
45	[▼]	N2 = *3.*
45		
60	[▼] 60 [DATA]	X3 = *60.*

Statistical Calculation (STAT)

All calculations in this section are performed in the STAT Mode ([MODE] [2]).

Selecting a Statistical Calculation Type

In the STAT Mode, display the statistical calculation type selection screen.

■ Statistical Calculation Types

Key	Menu Item	Statistical Calculation
[1]	1-VAR	Single-variable
[2]	A+BX	Linear regression
[3]	_+CX^2	Quadratic regression
[4]	ln X	Logarithmic regression
[5]	e^X	e exponential regression
[6]	A•B^X	ab exponential regression
[7]	A•X^B	Power regression
[8]	1/X	Inverse regression

select 1

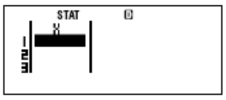
■ Inputting Sample Data

Displaying the STAT Editor Screen

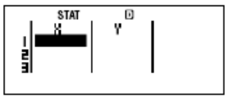
The STAT editor screen appears after you enter the STAT Mode from another mode. Use the STAT menu to select a statistical calculation type. To display the STAT editor screen from another STAT Mode screen, press [SHIFT] [1] (STAT) [2] (Data).

STAT Editor Screen

There are two STAT editor screen formats, depending on the type of statistical calculation you selected.



Single-variable Statistics



Paired-variable Statistics

- The first line of the STAT editor screen shows the value for the first sample or the values for their first pair of samples.

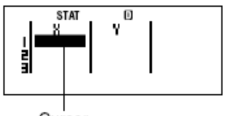
FREQ (Frequency) Column

If you turn on the Statistical Display item on the calculator's setup screen, a column labeled "FREQ" will also be included on the STAT editor screen.

You can use the FREQ column to specify the frequency (the number of times the same sample appears in the group of data) of each sample value.

Rules for Inputting Sample Data on the STAT Editor Screen

- Data you input is inserted into the cell where the cursor is located. Use the cursor keys to move the cursor between cells.

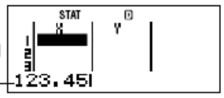


Cursor

- The values and expressions you can input on the STAT editor screen are the same as those you can input in the COMP Mode with Linear format.
- Pressing [AC] while inputting data clears your current input.
- After inputting a value, press [=]. This registers the value and displays up to six of its digits in the currently selected cell.

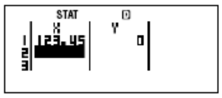
Example: To input the value 123.45 in cell X1
(Move the cursor to cell X1.)

[1] [2] [3] [.] [4] [5]



The value you input appears in the formula area.

[=]



Registering a value causes the cursor to move down one cell.

STAT Editor Screen Input Precautions

- The number of lines in STAT editor screen (the number of sample data values you can input) depends on the type of statistical data you selected, and on the Statistical Display setting of the calculator's setup screen.

Statistic Type \ Statistical Display	OFF (No FREQ column)	ON (FREQ column)
Single-variable	80 lines	40 lines
Paired-variable	40 lines	26 lines

- The following types of input are not allowed on the STAT editor screen.
  - [M+], [SHIFT] [M+] (M–) operations
  - Assignment to variables (STO)

Precautions Concerning Sample Data Storage

Sample data you input is deleted automatically whenever you change to another mode from the STAT Mode or change the Statistical Display setting (which causes the FREQ column to be shown or hidden) on the calculator's setup screen.

Editing Sample Data

Replacing the Data in a Cell

(1) On the STAT editor screen, move the cursor to the cell you want to edit.

(2) Input the new data value or expression, and then press [=].

Important!

- Note that you must totally replace the existing data of the cell with new input. You cannot edit parts of the existing data.

Deleting a Line

(1) On the STAT editor screen, move the cursor to the line you want to delete.

(2) Press [DEL].

Inserting a Line

(1) On the STAT editor screen, move the cursor to the line that will be under the line you will insert.
(2) Press [SHIFT] [1] (STAT) [3] (Edit).
(3) Press [1] (Ins).

Important!

- Note that the insert operation will not work if the maximum number of lines allowed for the STAT editor screen are already used.

Deleting All STAT Editor Contents

(1) Press [SHIFT] [1] (STAT) [3] (Edit).
(2) Press [2] (Del-A).
  - This clears all of the sample data on the STAT editor screen.

Note

- Note that you can perform the procedures under "Inserting a Line" and "Deleting All STAT Editor Contents" only when the STAT editor screen is on the display.

■ STAT Calculation Screen

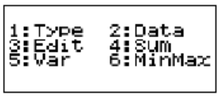
The STAT calculation screen is for performing statistical calculations with the data you input with the STAT editor screen. Pressing the [AC] key while the STAT editor screen is displayed switches to the STAT calculation screen.

The STAT calculation screen also uses Linear format, regardless of the current input/output format setting on the calculator's setup screen.

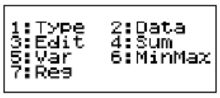
■ Using the STAT Menu

While the STAT editor screen or STAT calculation screen is on the display, press [SHIFT] [1] (STAT) to display the STAT menu.

The content to the STAT menu depends on whether the currently selected statistical operation type uses a single variable or paired variables.



Single-variable Statistics



Paired-variable Statistics

STAT Menu Items

Common Items

Select this menu item:	When you want to do this:
[1] Type	Display the statistical calculation type selection screen
[2] Data	Display the STAT editor screen
[3] Edit	Display the Edit sub-menu for editing STAT editor screen contents
[4] Sum	Display the Sum sub-menu of commands for calculating sums
[5] Var	Display the Var sub-menu of commands for calculating the mean, standard deviation, etc.
[6] MinMax	Display the MinMax sub-menu of commands for obtaining maximum and minimum values

Paired-variable Menu Item

Select this menu item:	When you want to do this:
[7] Reg	Display the Reg sub-menu of commands for regression calculations • For details see "Commands when Linear Regression Calculation (A+BX) Is Selected" and "Commands when Quadratic Regression Calculation (_+CX²) Is Selected".

Single-variable (1-VAR) Statistical Calculation Commands

The following are the commands that appear on the sub-menus that appear when you select [4] (Sum), [5] (Var), or [6] (MinMax) on the STAT menu while a single-variable statistical calculation type is selected.

See **Appendix** <#039> for information about the calculation formula used for each command.

Sum Sub-menu ([SHIFT] [1] (STAT) [4] (Sum))

Select this menu item:	When you want to obtain this:
[1] Σx^2	Sum of squares of the sample data
[2] Σx	Sum of the sample data

Var Sub-menu ([SHIFT] [1] (STAT) [5] (Var))

Select this menu item:	When you want to obtain this:
[1] n	Number of samples
[2] $\bar{x}$	Mean of the sample data
[3] $x\sigma n$	Population standard deviation
[4] $x\sigma n{-}1$	Sample standard deviation

MinMax Sub-menu ([SHIFT] [1] (STAT) [6] (MinMax))

Select this menu item:	When you want to obtain this:
[1] minX	Minimum value
[2] maxX	Maximum value

Appendix Single-variable Statistical Calculation

<#040> Select single-variable (1-VAR) and input the following data:
{1, 2, 3, 4, 5, 6, 7, 8, 9, 10} (FREQ: ON)

<#041> Edit the data to the following, using insert and delete:
{0, 1, 2, 3, 4, 5, 6, 7, 9, 10} (FREQ: ON)

<#042> Edit the FREQ data to the following:
{1, 2, 1, 2, 2, 2, 3, 4, 2, 1} (FREQ: ON)

- Examples <#043> through <#045> all use the same data as Example <#042>.

<#043> Calculate sum of squares of the sample data and sum of the sample data.

<#044> Calculate number of samples, mean, and population standard deviation.

<#045> Calculate minimum value and maximum value.

$x\sigma n$-1 corresponds to standard deviation

#039

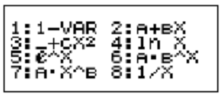
$$\bar{x} = \frac{\Sigma x}{n}$$

$$x\sigma n = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n}}$$

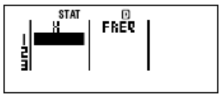
$$x\sigma n - 1 = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n - 1}}$$

#040

[SHIFT] [MODE] [▼] [3] (STAT) [1] (ON)
[MODE] [2] (STAT)



[1] (1-VAR)



[1] [=] [2] [=] [3] [=] [4] [=]
[5] [=] [6] [=] [7] [=] [8] [=]
[9] [=] [1] [0] [=]

[AC]