

## Module P1.3 Graphs and measurements

- 1 [Opening items](#)
- 2 [Axes and coordinates](#)
- 3 [Plotting a graph](#)
  - 3.1 [A simple experiment](#)
  - 3.2 [A slightly more complicated experiment](#)
  - 3.3 [Guidelines for plotting graphs](#)
- 4 [Graphs of some common relationships](#)
  - 4.1 [Linear relationships](#)
  - 4.2 [Determining the gradient and intercept of a linear graph](#)
  - 4.3 [Non-linear graphs and linearization](#)
  - 4.4 [Linearizing with logarithms](#)
- 5 [Incorporation of data errors on a graph](#)
  - 5.1 [Showing the likely data errors on the graph](#)
  - 5.2 [Estimating the errors in the gradient and the intercept](#)
- 6 [Numerical analysis of slopes and intercepts](#)
  - 6.1 [Choice of the 'best' line by numerical means — the method of least squares](#)
  - 6.2 [A cautionary note](#)
- 7 [Closing items](#)
  - 7.1 [Module summary](#)
  - 7.2 [Achievements](#)
  - 7.3 [Exit test](#)

[Exit module](#)

# 1 Opening items

## 1.1 Module introduction

Plotting a *graph* is a very useful way of showing the relationship between two measured variables; this module explains how this is done, using two simple experiments as examples. It gives some general guidelines for preparing graphs and then goes on to examine *linear relationships* and the straight-line graphs that represent them. It also explains how to estimate the *gradient* and *intercept* that characterize such graphs.


Some commonly encountered non-linear relationships are dealt with in Subsection 4.3, and the use of logarithms to transform certain non-linear relationships into linear ones is discussed in Subsection 4.4.

Section 5 deals with the question of experimental errors on graphs: how these can be plotted, and how they can be used to estimate the error in the gradient and intercept of a straight-line graph.

Finally, in Section 6, formulae are given that can be used to calculate directly the values of the gradient and intercept of the straight-line graph that best fits a given set of numerical data.

**Study comment** Having read the introduction you may feel that you are already familiar with the material covered by this module and that you do not need to study it. If so, try the [Fast track questions](#) given in Subsection 1.2. If not, proceed directly to [Ready to study?](#) in Subsection 1.3.

## 1.2 Fast track questions

**Study comment** Can you answer the following *Fast track questions*?. If you answer the questions successfully you need only glance through the module before looking at the *Module summary* (Subsection 7.1) and the *Achievements* listed in Subsection 7.2. If you are sure that you can meet each of these achievements, try the *Exit test* in Subsection 7.3. If you have difficulty with only one or two of the questions you should follow the guidance given in the answers and read the relevant parts of the module. However, *if you have difficulty with more than two of the Exit questions you are strongly advised to study the whole module.* 

### Question F1

The mass  $M$  of a flask and its contents, is measured as a function of the volume  $V$  of liquid in the flask, and the results are given in Table 1.

Plot a graph of  $M$  against  $V$ .

Assuming that  $M$  and  $V$  are related by the equation  $M = M_0 + \rho V$ , use your graph to estimate the values of  $M_0$  and  $\rho$ . Take care to include error estimates in your answer.



**Table 1** See Question F1.

Mass of the flask and its contents $M/\text{g}$	Volume of the liquid $V/\text{cm}^3$
$330 \pm 20$	$200 \pm 20$
$480 \pm 20$	$400 \pm 20$
$660 \pm 20$	$600 \pm 20$
$850 \pm 20$	$800 \pm 20$
$1010 \pm 20$	$1000 \pm 20$

## Question F2

Show that the data in Table 2 can be sensibly represented by an equation of the form  $y = kx^\alpha$ , and find approximate values for  $k$  and  $\alpha$ .



**Table 2** See Question F2.

$x$	$y$
2	63
6	91
10	107
25	146
62	198

### *Study comment*

Having seen the *Fast track questions* you may feel that it would be wiser to follow the normal route through the module and to proceed directly to [Ready to study?](#) in Subsection 1.3.

Alternatively, you may still be sufficiently comfortable with the material covered by the module to proceed directly to the [Closing items](#).

## 1.3 Ready to study?

**Study comment** In order to study this module you will need to be familiar with the following terms: error (in a measurement), mean (i.e. average), random error, systematic error, table of values and uncertainty (a synonym for error). You should be able to use a calculator to perform arithmetical operations, and to work in scientific (powers-of-ten or standard) notation. The module also makes use of logarithms; so you will need to know what the log of a number is, and how to obtain it using your calculator. (Most aspects of this topic, including the use of natural logarithms ( $\log_e$ ), are reviewed within the module.) If you are uncertain of any of these terms, you can review them by referring to the *Glossary* which will indicate where in *FLAP* they are developed. The following *Ready to study questions* will allow you to establish whether you need to review some of the topics before embarking on this module.

## Question R1

Draw up and complete a [table of values](#) for each of the following equations: (a)  $2y = 3x - 1$ , and (b)  $y = x^2 + 3$ , where  $x$  is between  $-3$  and  $+3$  in both cases.



## Question R2

Evaluate the following in [scientific notation](#):

$$(a) \frac{(10^6)^{2/3} \times (4 \times 10^4)}{0.001 \times 10^2}$$


$$(b) \frac{(65 \times 10^{-4}) \times (10^8)^{9/2}}{(1.3 \times 10^4) \times (200)}$$

$$(c) \frac{900 \times 0.0002 \times (12 \times 10^2)^2}{30 \times 0.1}$$

$$(d) \left[ \frac{0.0025 \times 10^4 \times 400 \times (10^2)^4}{500\,000 \times (100)^{1/2}} \right]^3$$




### Question R3

Evaluate the following logarithms (without using your calculator): 

- (a)  $\log_{10}(10)$ , (b)  $\log_{10}(1000)$ , (c)  $\log_{10}(1000\ 000)$ , (d)  $\log_{10}(0.0001)$ ,  
(e)  $\log_{10}(0.0000001)$ , (f)  $\log_{10}(1)$ .



### Question R4

Given that  $\log_{10} 2 = 0.301$ , evaluate the following logarithms (without using your calculator): 

- (a)  $\log_{10}(200)$ , (b)  $\log_{10}(20\ 000)$ , (c)  $\log_{10}(0.2)$ , (d)  $\log_{10}(0.00002)$ ,  
(e)  $\log_{10}(4)$ , (f)  $\log_{10}(32)$ , (g)  $\log_{10}(40)$ , (h)  $\log_{10}(64)$ .





## Question R5

A spring is suspended vertically and is stretched by a mass hung from its free end. This is repeated several times with equal masses, and the extension of the spring measured on each occasion. The values obtained (in mm) are:

77, 74, 74, 71, 73, 75, 72, 75

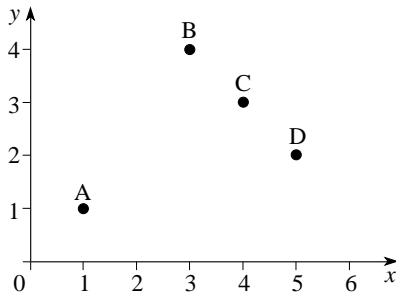
Calculate the [mean](#) value, and estimate the size of the [random error](#).  
What sources of [systematic error](#) might there be?



## 2 Axes and coordinates

In this module we will use the term **graph** to mean a plot of two quantities (or *variables*) on two perpendicular lines called the **axes**, as in Figure 1. In the case of Figure 1 the horizontal and vertical axes represent the variables  $x$  and  $y$ , respectively, so they are labelled accordingly, and may be referred to as the  **$x$ -axis** and the  **$y$ -axis**. If we were to use the axes to represent two different variables, say  $m$  and  $t$ , then they would be known as the  $m$ -axis and the  $t$ -axis. The arrow on each axis indicates the direction in which the corresponding variable increases. The axes of this particular graph intersect at a point called the **origin** which corresponds to the zero value for each variable. Sometimes it is convenient to draw a graph so that the axes intersect at values other than zero.

We can now describe the position of a point on the graph by specifying the corresponding value of each variable. For example the point B in Figure 1 is specified by  $x = 3$  and  $y = 4$ . These values are called the **coordinates** of point B. Coordinate values are often given in the form of an **ordered pair**, such as  $(3, 4)$  in which case the first value, 3, corresponds to a value on the horizontal axis, while the second value, 4, corresponds to a value on the vertical axis. If we wish to refer to a point by a letter as well as by its coordinates we may, for example, refer to it as the point B(3, 4).



**Figure 1** A simple graph.

### Question T1

Determine the  $x$ -coordinate and the  $y$ -coordinate of each of the points A, C and D shown in Figure 1. ☐



### Question T2

On the axes given in Figure 1 plot the points: P(2, 4), Q(1, 4), R(0, 2), S(3, 0). ☐



## 3 Plotting a graph

### 3.1 A simple experiment

Graphs are very useful for displaying the results of experiments, and are particularly good at revealing any systematic relationship between the variables that have been measured. We will illustrate this by analysing a very simple experiment — the stretching of a rubber band. This may seem a trivial experiment, but it serves our purpose quite well. The rubber band that we will examine is the sort used in a vacuum cleaner. It needs to be sufficiently flexible to be drawn easily over the pulleys, but not so flexible that it slips when the motor runs. To fit the band on the pulleys, it has to be stretched by about 5 cm, but it has to be able to sustain at least twice that extension without breaking or suffering any permanent distortion.

In a preliminary trial experiment we cut the rubber band and suspend it vertically, then attach various masses to the free end. We find that the band stretches by 12 cm when loaded with a mass of 8 kg. We then use a ruler to measure the extension of the band as we increase the load from 0 to 8 kg, in equal increments of 1 kg.

The results of this experiment are shown in Table 3. There will of course be an *uncertainty*, or *error*, in the measurements of extension and mass — we will return to this point later.

Is there a way of expressing the relationship between extension and load, and how can this relationship be found?

**Table 3** Experimental results for the extension of a loaded rubber band.


Mass $m/\text{kg}$	Extension $e/\text{mm}$
0	0
1	13
2	28
3	44
4	58
5	74
6	89
7	105
8	120

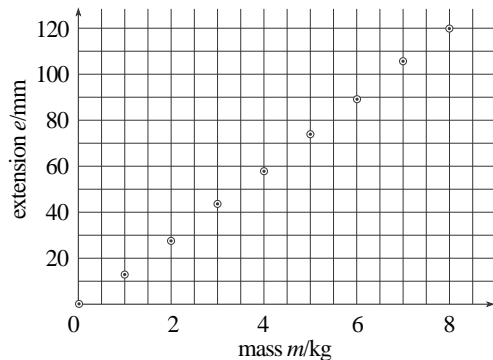
This is where a graph comes in useful. Figure 2 shows the data of Table 3 plotted as a graph.

Extension is plotted along the vertical axis and the mass along the horizontal axis. This is because the graph is meant to show how the extension changed as the mass of the load varied.

In this experiment both the extension and mass vary, both are variables; but the mass  $m$  is independently *chosen* by the experimenter and may be varied at will, so we call it the **independent variable**. The experimenter has no such direct control over the extension  $e$ , which is determined by the load. The extension is therefore called the **dependent variable**.

It is customary to plot all graphs with the independent variable along the horizontal axis and the dependent variable along the vertical axis; this convention has been followed in Figure 2.

Notice that the axes in Figure 2 are labelled in a similar way to the headings in Table 3.  In each case the variable quantity is divided by the appropriate unit of measurement. This allows us to write dimensionless numbers along the axes and saves the bother of writing units each time.



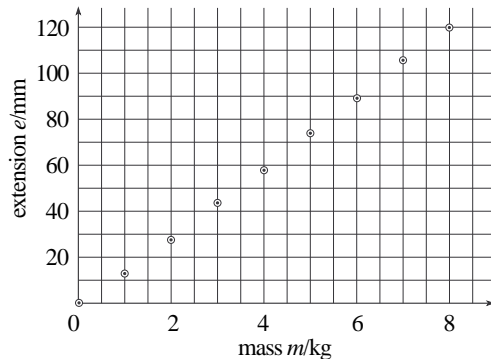
**Figure 2** Experimental results for the extension of a loaded rubber band.

Each of the data points is represented by a point within a circle, which allows some precision in the placing of the point, while at the same time making its location obvious at a glance.

Table 3 contains the same information as Figure 2, but the relationship is much easier to understand when it is displayed as a graph.

It looks very much like a *linear* relation — a straight line can be drawn through the points. So how should the ‘best’ straight line be drawn? The simplest method — which is perfectly satisfactory for most purposes — is to do it ‘by eye’. Choose a line that goes as nearly as possible through all the points, with the points scattered as equally as possible on both sides of the line.

In this particular case we would be justified in assuming that zero load will cause zero extension (so that the point (0, 0) really should be on the ‘best’ straight line).



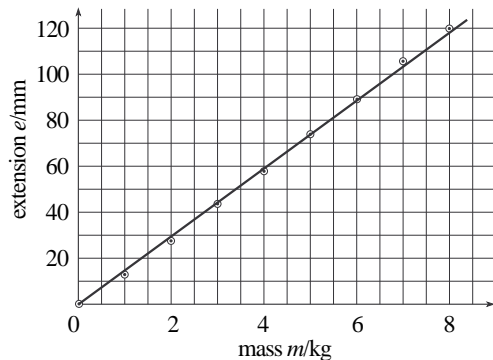
**Figure 2** Experimental results for the extension of a loaded rubber band.

Figure 3 shows the straight line through the origin that ‘best’ fits the data.

Having drawn the straight line, we can readily estimate what the extension of the rubber band might be for any given load; for example, if the load is 2.5 kg, we can see from the graph that the extension will be about 37 mm.

Obtaining the estimated value of the dependent variable for any given value of the independent variable *within the range of the measurements made* is called [interpolation](#).

Sometimes we want to estimate the value of the dependent variable ( $e$  in this instance) *outside* the range of the data; for example, we might want to know the value of the extension for a mass of 10 kg. This procedure is referred to as [extrapolation](#) and is generally a riskier business than interpolation because it relies on an *assumption* that a trend can be continued, as the example described in the next subsection illustrates.



**Figure 3** The best straight line drawn through the points.




### 3.2 A slightly more complicated experiment

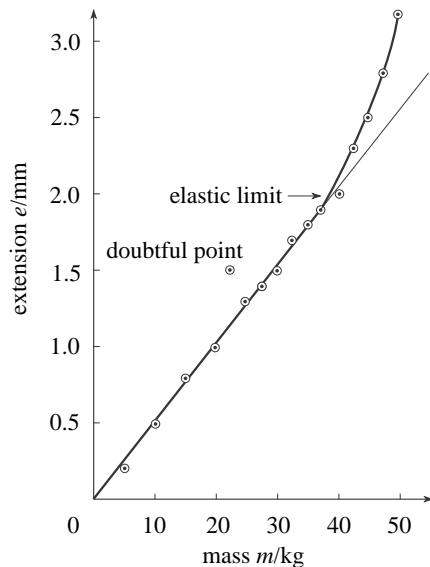
Suppose we now perform an experiment similar to the one we described in the previous subsection, but this time using a piece of copper wire in place of the rubber band. Table 4 and Figure 4 summarize the results obtained.

**Table 4** Experimental results for the extension of a loaded copper wire.

Mass $m/\text{kg}$	Extension $e/\text{mm}$	Mass $m/\text{kg}$	Extension $e/\text{mm}$
5.0	0.2	32.5	1.7
10.0	0.5	35.0	1.8
15.0	0.8	37.5	1.9
20.0	1.0	40.0	2.0
22.5	1.5	42.5	2.3
25.0	1.3	45.0	2.5
27.5	1.4	47.5	2.8
30.0	1.5	50.0	3.2

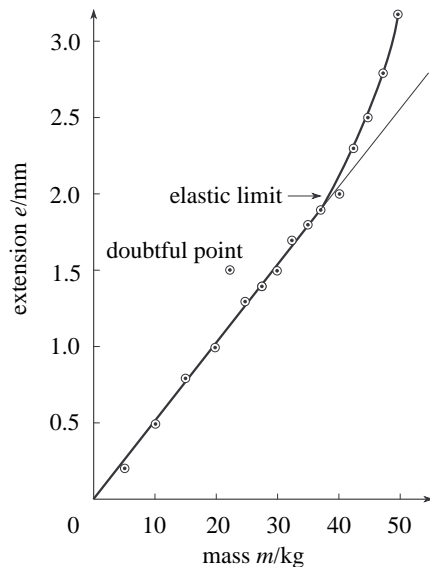
After glancing at the graph you can see immediately that there is something odd about the point (22.5, 1.5), so we will leave that to one side for the moment. From the best curve through the remaining points you can see that:

- (a) For extensions up to about 2.0 mm, the points on Figure 4 lie fairly close to the straight line so the extension is proportional to the mass, i.e. doubling the mass produces twice the extension.
- (b) For extensions greater than about 2.0 mm, the wire extends more easily, and extension is no longer proportional to mass. (In fact, if the wire is extended more than 2.0 mm, it will be found that it does not return to its original length when the load is removed. This critical extension is called the elastic limit.) 



**Figure 4** Extension of a loaded copper wire.

- (c) The point plotted as extension = 1.5 mm, mass = 22.5 kg is anomalous — it is much farther from the line than any other point, in fact it is higher than the points from the next two larger masses; this point ought to be checked.
- (d) The extension that you would expect for a mass of 22.5 kg is 1.1 mm. In this case the straight line on the graph is used to *interpolate* between measured values.
- (e) Some of the other measurements correspond to points that lie a little off the line, but this is to be expected since the straight line *averages out* random experimental errors in individual measurements.



**Figure 4** Extension of a loaded copper wire.

All of the above statements could be made by examining the data in Table 4 — after all, the graph was plotted using only the information in the table. However, it would take a long time to arrive at points (a) to (e) from the tabulated information, whereas they can all be very rapidly deduced from the graph. This is the great advantage of graphs as visual aids. The form of the relationship between measured quantities becomes clear, and the presence of anomalous measurements is readily apparent. In addition, graphs allow straightforward averaging of random errors from experimental measurements, interpolation between measurements, and extrapolation beyond the measurements.

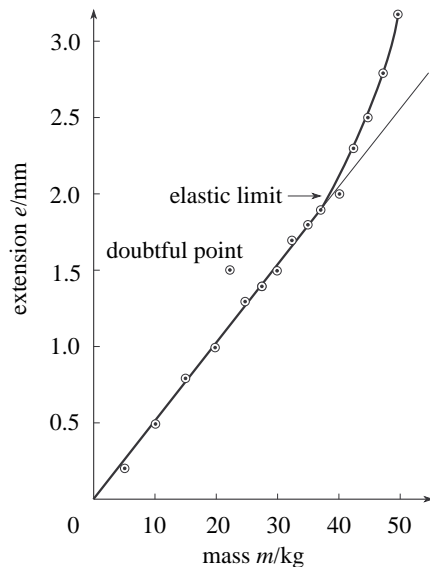
In some cases we can use the graph to determine the *equation* of the straight line that best fits the data, and, as we will see later in the module, we can sometimes do something similar when the data points lie on a curve.

### 3.3 Guidelines for plotting graphs

The last two subsections illustrate how graphs can be useful for exploring the relationship between two variables. This subsection summarizes some important guidelines for plotting graphs of experimental measurements.

**1 Plot the independent variable along the horizontal axis and the dependent variable along the vertical axis.** For example, to obtain the results shown in Figure 4, various masses were hung from the wire and the resulting extension was measured. The mass is the independent variable and is plotted horizontally, the extension is the dependent variable and is plotted vertically. Figure 4 shows how the extension of the wire depends on the mass hung from it.

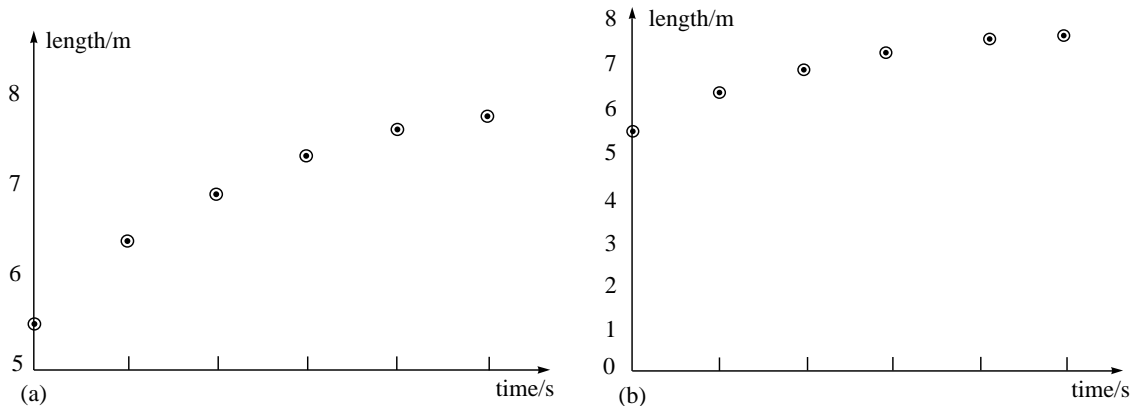
**2 Label both axes to show what quantity is plotted, and don't forget to include the units.** Since only pure numbers can be plotted, the quantity measured must be divided by its units before plotting, which means that the axis should be labelled quantity/units. Thus, in Figure 4, the axes are labelled extension  $e/\text{mm}$ , and mass  $m/\text{kg}$ .



**Figure 4** Extension of a loaded copper wire.

**3 Choose the scales on the axes to make plotting simple.** Generally this means letting 10 small divisions on the graph paper equal 1, 2, 5, or some multiple of 10 of these numbers. Don't make your life difficult by making 10 small divisions equal 3 or 7, for example; you would take much longer to plot the graph, and the chances of misplacing points (in a literal sense) would be very much higher.

**4 Choose the range of the scales on the axes so that the points are suitably spread out on the graph paper** and not all cramped into one corner. In some cases this may mean excluding zero from the axis: for example, if lengths measured in an experiment varied between 5.2 m and 7.7 m, it would be better to allow the scale to run from 5 m to 8 m as in Figure 5a rather than from zero to 8 m as in Figure 5b.

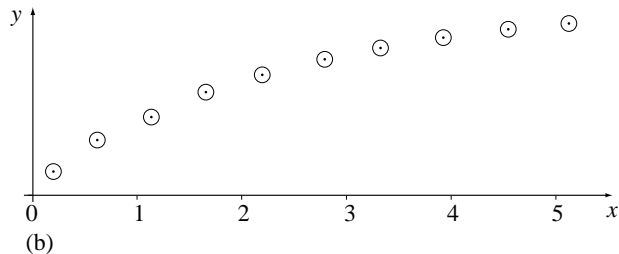
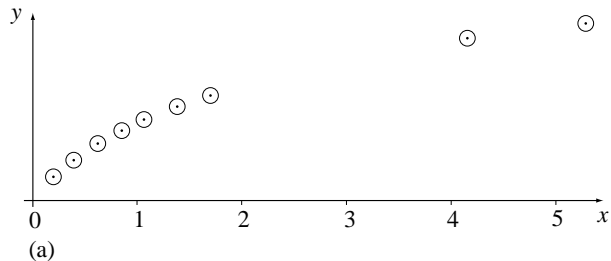


**Figure 5** (a) Sensible choice of scale. (b) Poor choice of scale.

**5 Plot results clearly.** Tiny dots may be confused with dirt on the graph paper, and big dots give loss of precision. Some authors indicate data points with crosses, but dots within circles are more common.




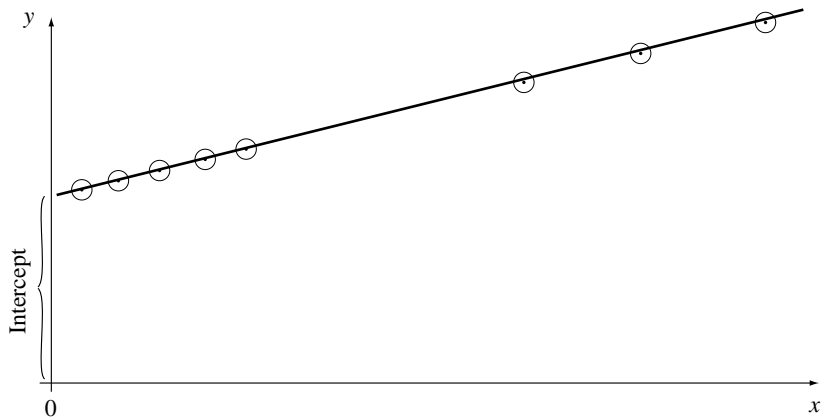
**6 When taking readings, generally spread them out evenly** over the range of values of the quantity measured. Figure 6a is poor, Figure 6b is better.



**Figure 6** (a) Readings unevenly spread. (b) Readings evenly spread.



Exceptions to this rule arise when you want to find where the line meets the  $y$ -axis — then it is desirable to have a few extra points close to the axis (as shown in Figure 7), or when you suspect that the graph is changing shape,  or near some other point of interest.



**Figure 7** Extra points near to the  $y$ -axis.


**7 Plot a graph as your experiment proceeds.** In this way you can check immediately if a point is so widely off that it needs to be measured again. You will also be able to check whether the points are reasonably evenly spaced or whether more points are needed.

**8 Draw a straight line or smooth curve through points plotted on the graph,** rather than joining up successive points by short straight lines. The graphs that you will draw generally represent some smooth variation of one quantity with another, so a smooth curve is usually appropriate.

## 4 Graphs of some common relationships

### 4.1 Linear relationships

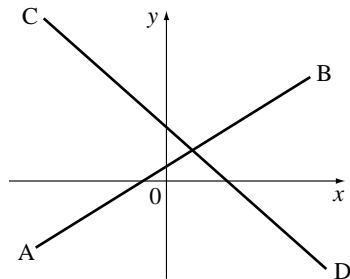
In this subsection, we will follow the convention that we introduced earlier — the *independent* variable will be denoted by  $x$  (on the horizontal axis) and the *dependent* variable by  $y$  (on the vertical axis).

As noted earlier, a [linear relationship](#) between  $y$  and  $x$  implies that the graph of  $y$  against  $x$  is a straight line. Any such straight line can be completely characterized by two constants; the *gradient* (or *slope*)  of the line, and the *intercept* (the value of  $y$  at which the line crosses the  $y$ -axis). We will examine these two characteristics in turn, starting with the gradient.

## The gradient of a straight line

You will have come across the terms **gradient** or **slope** many times before in such expressions as ‘the gradient of a hill’, or ‘the slope of a roof’. Look at Figure 8, which shows two straight lines with quite different gradients.

In moving from A to B we move from left to right across the page, so that the value of  $x$  increases, and we appear to move uphill because the value of  $y$  also *increases* as we move. Such a line is said to have a *positive gradient*.

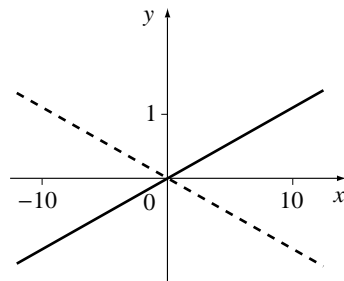


**Figure 8** An illustration of gradient.

In contrast, when we move from C to D, we still move from left to right across the page and  $x$  still increases, but now we appear to move downhill because the value of  $y$  *decreases* as we move. Such a line is said to have a *negative gradient*.

The value of the *gradient* tells us more about the line than whether it goes uphill or downhill; it also tells us *how steeply* the line rises or falls.

You have probably seen something like ‘gradient 1 in 10’, as a warning sign to motorists approaching steep hills. In mathematics a line with a gradient of 1 in 10 or  $1/10$  is a line such that for every 10 units crossed horizontally (from left to right) the line rises 1 unit; so, for every 10 units increase in the quantity plotted along the  $x$ -axis, there is a corresponding increase of 1 unit in the quantity plotted along the  $y$ -axis (see Figure 9). A line with a gradient of  $-1/10$  would be just as steep, but it slopes *downhill* (as illustrated by the dashed line in Figure 9).



**Figure 9** Positive and negative gradients.

In general, we use the term **rise** to describe the vertical separation between any two points on a line (measured in the upwards direction, the direction of increasing  $y$ ) and we use the term **run** to describe the corresponding horizontal separation between the same two points (measured from left to right, in the direction of increasing  $x$ ). Given the rise and run between two different points on a line, we can assign a numerical value to the gradient of that line by using the formula

$$\text{gradient} = \frac{\text{rise}}{\text{run}}$$

The gradient found above is dimensionless since the graph (in Figure 9) was dimensionless.

This is not usually so in examples from physics. For example, if we were to plot the position of a car (the dependent variable) with time then the gradient of the graph would probably be measured in  $\frac{\text{metres}}{\text{second}}$  with units  $\text{m s}^{-1}$ , it would *not* be a pure number. In physics examples you must remember the units, as indicated on the axes of the graph, when finding gradients.

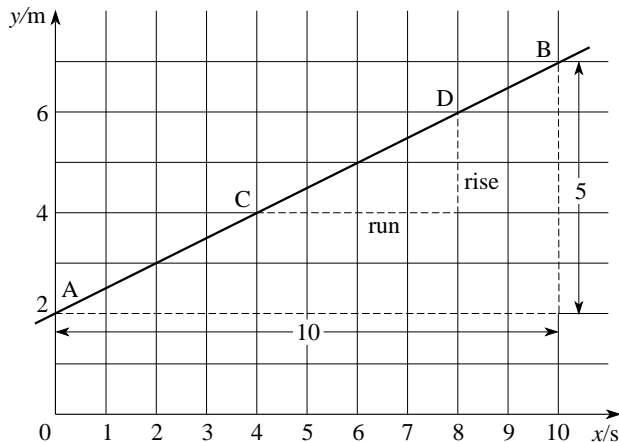
Thus, for the straight line shown in Figure 10, between the points A and B,

rise = 5 m and run = 10 s

$$\text{so, } \text{gradient} = \frac{\text{rise}}{\text{run}} = \frac{5 \text{ m}}{10 \text{ s}} = 0.5 \text{ m s}^{-1}$$

Similarly, for the same line, but using the points C and D

$$\text{rise} = 2 \text{ m and run} = 4 \text{ s} \quad \text{so,} \quad \text{gradient} = \frac{\text{rise}}{\text{run}} = \frac{2 \text{ m}}{4 \text{ s}} = 0.5 \text{ m s}^{-1}$$



**Figure 10** Gradient of a straight line.

Because the gradient of the line is constant, it does not depend on which pair of points on the line is used to calculate it. However, we can read off the coordinates of the points only with limited accuracy, and so, in practice, different pairs of points may produce slightly different values of the gradient. It is good practice to choose two points that are as widely separated as possible, since then the errors in reading the coordinates will be a much smaller fraction of the difference between the coordinates than if the points were close together.

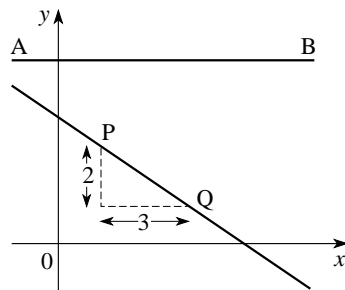
Using this rise/run definition of gradient, it is easy to see how the gradient of a line can be negative. In Figure 11, moving from P to Q, the run is just 3, but the ‘rise’ is in this case a ‘fall’, and we write:

$$\text{rise} = -2 \text{ and run} = 3$$

so,  $\text{gradient} = -2/3$

The gradient of a straight line can therefore be positive or negative, but what about lines with gradient 0?

From the definition,  $\text{gradient} = \frac{\text{rise}}{\text{run}} = 0$  implies that the rise is zero whatever run we choose, and the line must therefore be horizontal (like the line AB in Figure 11).



**Figure 11** A line with a negative gradient and a line with zero gradient.

Given any straight line, and two different points on the line, we can also determine the gradient from the coordinates of the points. In Figure 12 we suppose that the points are P and Q. We can use subscripts to indicate the coordinates  $(x_p, y_p)$  of P, and similarly  $(x_q, y_q)$  for the coordinates of point Q, where  $x_p, x_q, y_p$  and  $y_q$  simply represent numbers. The rise is the  $y$ -coordinate of Q minus the  $y$ -coordinate of P (expressed in the appropriate units):

$$\text{rise} = y_q - y_p$$

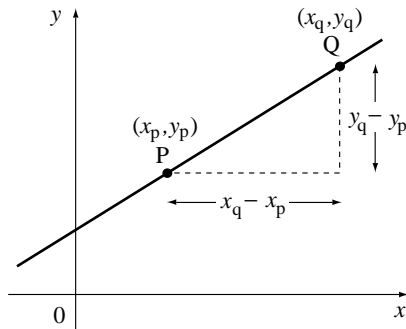
The run is the  $x$ -coordinate of Q minus the  $x$ -coordinate of P expressed in the appropriate units):

$$\text{run} = x_q - x_p$$

So the gradient is:

$$\frac{\text{rise}}{\text{run}} = \frac{y_q - y_p}{x_q - x_p}$$

Notice that if  $y$  decreases as  $x$  increases, then  $y_q$  will be less than  $y_p$  and the rise will be *negative*, giving a *negative* gradient.



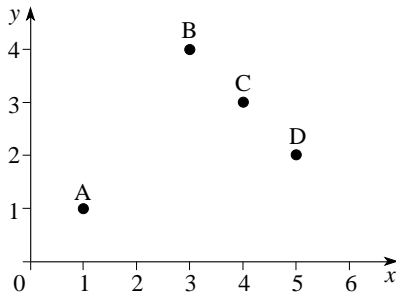
**Figure 12** Determining the gradient.



◆ Calculate the gradient of the line joining the points A(1, 1) and B(3, 4) in Figure 1.



◆ Calculate the gradient from B(3, 4) to C(4, 3), then the gradient from C to D(5, 2) in Figure 1.



**Figure 1** A simple graph.

### Question T3

Plot the graphs and find the gradients for the following equations:

(a)  $y = 15x$ , (b)  $y = 5x$ , (c)  $y = 5x + 10$ , (d)  $y = 5x + 30$ .

(Choose values of  $x$  from  $-4$  up to  $+4$ .) ☐



You will have found that the gradient of the straight line  $y = 15x$  is 15. The gradient of the second equation is 5, and the third and fourth lines also have gradients of 5. You will probably have noticed that the gradient is the same as the *coefficient* of  $x$  in the given equation. So, for example, the gradient of the line  $y = -3x + 1$  is  $-3$ ; and the gradient of the line  $y = x - 4$  is 1, because the coefficient of  $x$  is 1. This appears straightforward, but you still need to exercise a little care, as the following examples illustrate.

◆ Write down the gradients of the following lines:

(a)  $y = 3 + 5x$ , (b)  $y + 2x = 3$ , (c)  $5x - 2 = 1 - 2y$ , (d)  $3 = \frac{5x}{2y}$ .



## The intercept of a straight line

Figure 13 shows the graph of  $y = 2x + 1$ . The graph cuts the  $y$ -axis at the point  $y = 1$ . In other words, when  $x = 0$  the corresponding value of  $y$  is 1. The point where the graph cuts the  $y$ -axis is called the *y-intercept*, though this is usually abbreviated to [intercept](#).

◆ Find the  $y$ -intercepts of the lines from the previous question:

(a)  $y = 3 + 5x$ , (b)  $y + 2x = 3$ , (c)  $5x - 2 = 1 - 2y$ , (d)  $3 = \frac{5x}{2y}$ .

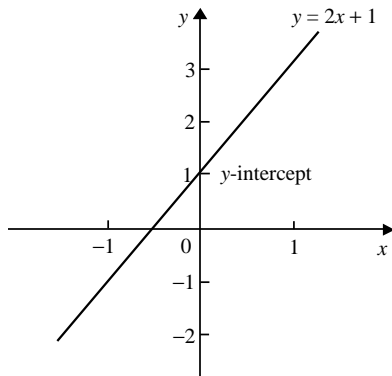


Figure 13 The graph of  $y = 2x + 1$ .

### Question T4

What are the  $y$ -intercepts of the graphs you plotted in Question T3?


[(a)  $y = 15x$ , (b)  $y = 5x$ , (c)  $y = 5x + 10$ , (d)  $y = 5x + 30$ .]

Do you notice any connection between these values and the equations? ☐




Questions T3 and T4 show that the equations can be written:

$$y = (\text{gradient})x + (\text{y-intercept})$$

So the gradient and the y-intercept tell us all the characteristics of a straight line, how steep it is and where it lies on the  $(x, y)$  plane in relation to the y-axis. 

We can now refer to a general **equation of a straight line**:

$$y = ax + b \quad \text{where } a \text{ and } b \text{ are constants} \quad (1)$$

in which the gradient is  $a$  and the y-intercept is  $b$ . 

❖ Find the gradients and intercepts of the straight lines that represent the following linear relationships:

(a)  $2y + 3 = 2x$ , (b)  $2y + x = 6 - y$ , (c)  $4y + 2(x - 1) = 6$ .



## 4.2 Determining the gradient and intercept of a linear graph

We now return to our rubber band experiment described in Subsection 3.1 and illustrated in Figure 3. Our  $y$ -variable here is the extension ( $e$ ), and our  $x$ -variable is the mass ( $m$ ). The line goes through the origin of the graph, so the  $y$ -intercept is equal to 0. The relationship implied by such a graph is therefore a simple equation relating length and mass:

$$\text{length} = \text{constant} \times \text{mass}$$

Using symbols to represent the quantities, the equation is more conveniently written as:

$$e = km$$

where  $k$  represents the constant, usually referred to as the [constant of proportionality](#) between  $e$  and  $m$ .

This is a special case of our general straight line equation,  $y = ax + b$  with  $y = e$ ,  $x = m$ ,  $a = k$  and  $b = 0$ .

It is certainly true that we could use the data provided in Table 3 in order to estimate the value of  $k$ , but, since we have already plotted this data and fitted a straight line to it in Figure 3, there is a far easier way. The constant  $k$  is just the gradient (i.e. the slope) of the line. To determine the gradient we choose two well separated points on the graph, such as the origin and the point (7 kg, 105 mm), and we use  $\text{gradient} = \text{rise/run}$ , thus

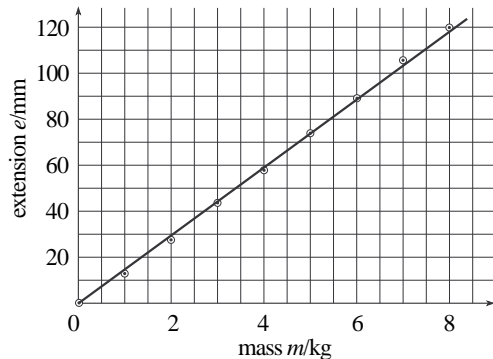
$$\text{gradient} = 105 \text{ mm}/7 \text{ kg} = 15 \text{ mm kg}^{-1}$$

Note that  $k$  is expressed in the correct units — the units of extension (mm) divided by the units of mass (kg).

Having evaluated the constant  $k$  we now have an expression relating the two columns of data in Table 3:

$$e \approx (15 \text{ mm kg}^{-1}) \times m$$

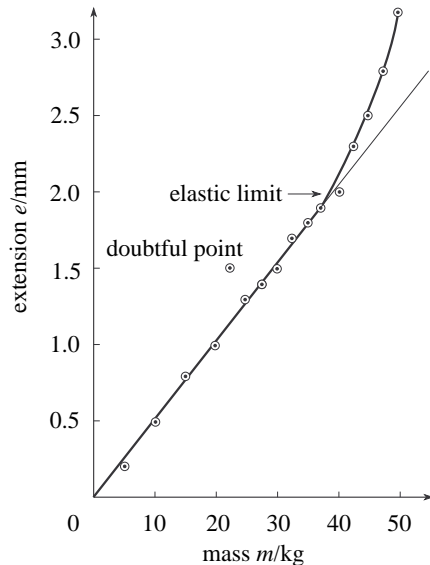
This is a much more convenient way to express the relationship than either the data in Table 3, or the graph in Figure 3.



**Figure 3** The best straight line drawn through the points.

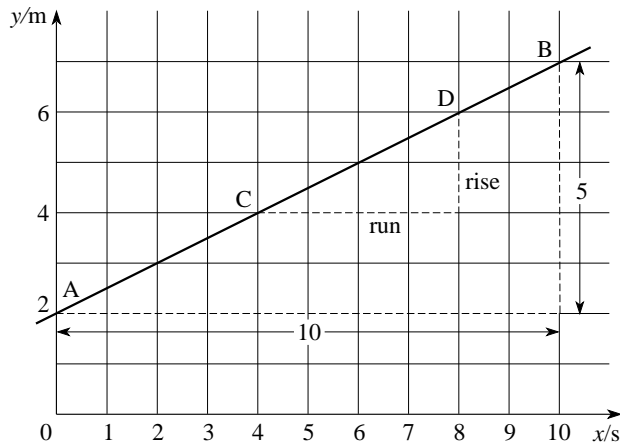
### Question T5

Calculate the gradient of the straight line in Figure 4 (derived from the extension up to the elastic limit in our second experiment, as described in Subsection 3.2) and write down the equation for the relationship between  $e$  and  $m$  where it is linear. □



**Figure 4** Extension of a loaded copper wire.

◆ Write down the equation of the straight line shown in Figure 10.



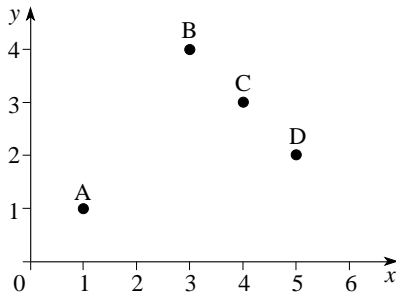
**Figure 10** Gradient of a straight line.



◆ Find the equation of the line joining B to D in Figure 1. Does the point E(37, -32) lie on this line?




◆ What is the gradient, and what is the intercept on the vertical axis, of the line  $3V + 2U = 5$ , when  $U$  and  $V$  are measured on the horizontal and vertical axes, respectively?




**Figure 1** A simple graph.


### 4.3 Non-linear graphs and their linearization

Equations such as  $y = 2x - 1$  are said to be **linear** because they correspond to straight-line graphs. More formally, we say in such a case that  $y$  is a **linear function** of  $x$ .  You will notice that such equations generally involve only the first power of  $x$  (i.e.  $x^1 = x$ ), for example

$$y = 3 - 2x, \quad 2x + 3y = 1, \quad \frac{x}{2} + \frac{y}{3} = 5 \quad \text{and} \quad \frac{1}{y} = \frac{2}{3x + 1}$$

are all linear because they can all be rearranged into the form  $y = ax + b$ , and they all give straight-line graphs when  $y$  is plotted against  $x$ . (Only straight lines parallel to the  $y$ -axis, such as  $x = 4$ , cannot be written in this form.)

However, many, if not most, relationships that you will encounter in physics will *not* be linear. Modern graphical calculators, and algebraic computing programs, certainly reduce the effort required to produce the graph of a given non-linear function, but they do not remove the need for some fundamental understanding of non-linear graphs. For example, if you try to produce the graph of  $y = 5000x^2$  on a calculator you may well have problems with the scale, and you may find that nothing appears on the screen. In order to produce a sensible output you need to have a rough idea of the shape of the graph before you try to plot it. 

In this subsection we consider some commonly encountered non-linear relationships, and the corresponding graphs. We start by considering the equations that give rise to two useful geometrical shapes — the *parabola* and the *hyperbola*. 

## Parabolas


The equation  $y = 3x^2 + 5$  includes a term in  $x^2$ , where the power of  $x$  is not 1 but 2, and the associated graph is not a straight line, but a special kind of curve known as a [parabola](#). Any equation of the form

$$y = ax^2 + bx + c$$

where  $a$ ,  $b$  and  $c$  are constants and  $a \neq 0$ , will provide a parabolic graph when  $y$  is plotted against  $x$ .

### Question T6

Construct a table of values for  $y$  corresponding to  $x$  ranging from  $-3$  to  $+3$  for both of the following equations, and plot their graphs:


(a)  $y = x^2 + 1$ ,   (b)  $y = x^2 - 1$ .   



## Hyperbolas

We now consider the equation  $y = 1/x$ . Table 5 shows values of  $y = 1/x$  corresponding to values of  $x$  for  $x = 1/5$  to 5.

For very large positive values of  $x$  the value of  $1/x$  is very small and positive. On the other hand, when  $x$  is very small and positive the value of  $1/x$  is very large and positive. But what happens when  $x = 0$ ?

Try calculating  $1/0$  on your calculator and it will probably produce an error message; this is because division by zero is not defined.  However, if you try evaluating  $1/0.001$  and then  $1/0.000\ 0001$  on your calculator, you will see that  $1/x$  can be made as large as you please by choosing  $x$  sufficiently close to zero.

The graph of any equation of the form


$$y = \frac{a}{x} + b$$

where  $a$  and  $b$  are constants and  $a \neq 0$  will be similar to the graph corresponding to Table 5; all such equations describe a kind of curve known as a [hyperbola](#).

**Table 5** The values of  $y = 1/x$ .

$x$	$y = 1/x$
1/5	5
1/3	3
1/2	2
1	1
2	1/2
3	1/3
5	1/5

## Question T7

Tabulate the values of the following equations for values of  $x$  between 0.2 and 5 (as in Table 5), and plot their graphs: (a)  $y = 2/x$  and (b)  $y = 2/x + 5$ . 



## Curve fitting and linearization

Our intuition helps us to see how to fit a straight line to a set of data, but it is not so easy to fit a parabola, or a hyperbola, to a set of data lying on a curve.

In general, non-linear relationships are much more difficult to deal with than linear relationships. For example, suppose that  $x$  and  $y$  are related in such a way that when we vary the value of  $x$  and measure the value of  $y$  we obtain the results given in Table 6.

If you were to plot these values on a graph you would find that the shape is very similar to that of [Figure 29](#); so you might suspect that a hyperbola, corresponding to an equation of the form  $y = \frac{a}{x} + b$ , would prove a good fit to the data. But how would you determine the appropriate values of  $a$  and  $b$ ?

**Table 6**

$x$	$y$
1.000	13
0.500	28
0.333	44
0.250	58
0.200	74
0.167	89
0.143	105
0.125	120

The task is certainly harder than that of fitting a straight line to a set of linearly related data. First we have the problem of drawing a neat hyperbola (for a straight line we just use a ruler and judge by eye) and then we need to actually determine  $a$  and  $b$  (for a straight line these are just the gradient and  $y$ -intercept, for the hyperbola  $a$  and  $b$  have no simple geometric interpretation.) Clearly determining  $a$  and  $b$  directly is a difficult undertaking, but fortunately there is a better way.

During our experiment it may have been convenient to measure the values of  $x$  and  $y$ , but *they need not be the variables that appear in our graph*. For example, if we suspect that an equation of the form  $y = \frac{a}{x} + b$  is a good fit to the above data list then if we replace  $1/x$  by  $X$  say, we would have an equation of the form  $y = aX + b$ , which is *the equation of a straight line* when  $y$  is plotted against  $X$ . Such a change of variable is said to *convert the original equation into linear form*, or to linearize the equation.

We now adapt our data list so that it shows the values of this new variable,  $X$ , as in Table 7.

You may well suspect that the original values were cunningly chosen (as indeed they were) since the  $X$  values here are precisely the same as the  $x$  values in Table 3, so we know from Figure 3 that the straight line  $y = 15X$  is a good fit to this data. In other words,  $a = 15$  and  $b = 0$ .

Now that we have determined the constants  $a$  and  $b$  from a linear graph we can return to our original variable  $x$ , knowing that the hyperbola

$$y = \frac{15}{x}$$

should be a good fit to the original data.

**Table 7**

$x$	$X = 1/x$	$y$
1.000	1.00	13
0.500	2.00	28
0.333	3.00	44
0.250	4.00	58
0.200	5.00	74
0.167	5.99	89
0.143	6.99	105
0.125	8.00	120

The measured values of  $x$  and  $y$ , together with the values of  $15/x$  are shown in Table 8 and, as you can see, the values of  $15/x$  agree quite well with the measured values of  $y$ .

❖ You suspect that an equation of the form  $y = ax^2 + b$  will provide a good fit to a set of data provided in a table of values of  $x$  and  $y$ . What change of variables would you choose in order to produce a straight-line graph?



❖ You complete an experiment in optics, varying the value of a length  $u$  (in metres) and measuring the value of a length  $v$  (in metres). You suspect that an equation of the form  $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$ , where  $f$  is a constant, will provide a good fit to the data, and you wish to find a value for  $f$ . What change of variables would you choose in order to produce a straight-line graph, and what will be the units on each axis? How would you use your graph to measure  $f$ ?



**Table 8**

$x$	$y$	$15/x$
1.000	13	15
0.500	28	30
0.333	44	45
0.250	58	60
0.200	74	75
0.167	89	90
0.143	105	105
0.125	120	120



## Question T8


You complete an experiment in which you vary the value of a length  $L$  (in metres) and measure the value of a time  $T$  (in seconds). You suspect that an equation of the form  $T = 2\pi\sqrt{L/g}$ , where  $g$  is a constant, will provide a good fit to the data, and you wish to find a value for  $g$ . What change of variables would you choose in order to produce a straight-line graph, and what will be the units on each axis? How would you use your graph to measure  $g$ ? ☐



.

## 4 Linearizing with logarithms

In this subsection we consider two important cases where equations can be converted into a linear form (linearized) using logarithms, and so it is worth spending a few moments recalling the general properties of logarithms:

- The *logarithm* of a number  $x$  to the base  $a$  is the power to which  $a$  must be raised in order to give the value  $x$ .  In other words, if  $y = \log_a x$  then  $x = a^y$ ; so that, for example,

$$\log_{10} 3 \approx 0.4771 \text{ because } 10^{0.4771} \approx 3$$

- $\log_a x + \log_a y = \log_a (xy)$ ; so that, for example

$$\log_{10} 2 + \log_{10} 3 = \log_{10} 6$$


- $\log_a (x^k) = k \log_a x$ ; so that, for example

$$\log_{10} \sqrt{x} = \log_{10} (x^{1/2}) = \frac{1}{2} \log_{10} x$$

- $\log_a (1/x) = -\log_a x$

- $\log_a 1 = 0$

- $\log_a a = 1$

In the work that follows it would be possible to use logarithms to any base, however in practice it is usually best to use either base 10 or base e, since these are provided on most calculators. 

◆ Solve the equation  $3^x = 7^{2x+1}$ .



## Power law relationships

Suppose we suspect that two variables  $x$  and  $y$  are related by a **power law**, that is, a relationship of the form

$$y = ax^n$$

(2)



where  $a$  and  $n$  are constants (but we do not, as yet, know their values). If we take logarithms to base 10 of both sides of the equation, we obtain:

$$\log_{10} y = \log_{10} (ax^n) = \log_{10} a + \log_{10} (x^n) = \log_{10} a + n \log_{10} x$$

If we now rearrange this equation

$$\log_{10} y = n(\log_{10} x) + \log_{10} a$$

we have once again something that looks like our standard straight-line equation. The dependent (vertical axis) variable will now be  $\log_{10} y$ , and the independent (horizontal axis) variable will be  $\log_{10} x$ . The gradient will be  $n$ , and the intercept on the vertical axis  $\log_{10} a$ .

◆ Linearize the equation  $s = \frac{1}{2} gt^2$  (where  $g$  is a constant) using logarithms.



## Question T9

Linearize the equation in Question T8 [  $T = 2\pi\sqrt{L/g}$  ] using logarithms.  $\square$



## Exponential relationships


The previous subsection concerned a relationship of the form

$$y = ax^n \quad (\text{Eqn 2})$$

where  $a$  and  $n$  are constants. In this subsection we discuss the situation in which two variables  $x$  and  $y$  are related by an **exponential law**, that is an equation of the form

$$y = ka^x \quad (3)$$

where  $k$  and  $a$  are constants.

A good example of an exponential law is the equation used to model the spontaneous decay of radioactive nuclei; the *radioactive decay equation* 

$$N = N_0(1/2)^{t/T} \quad (4) \quad \img alt="pointing hand icon" data-bbox="752 582 782 618"/>$$

where

- $N_0$  (a constant) is the number of radioactive nuclei present at time  $t = 0$ ;
- $N$  is the number of radioactive nuclei present after a time  $t$  has elapsed;
- $T$  (a constant) is the *half-life* of the radioactive material concerned, (i.e. the time after which half of the material has decayed).

After one half-life,  $t = T$ , and  $N = N_0/2$ ; after two half-lives,  $t = 2T$  and  $N = N_0(1/2)^2 = N_0/4$ , and so on.

Relationships of this kind are called *exponential*; they too can be easily linearized if we take logarithms. Once again we could use logarithms to any base, but, to add variety, this time we will use logarithms to base e.

$$\begin{aligned}\log_e N &= \log_e N_0 + \log_e [(1/2)^{t/T}] = \log_e N_0 + (t/T) \log_e (1/2) \\ &= \log_e N_0 + (t/T) (\log_e 1 - \log_e 2) = \log_e N_0 - (\log_e 2) (t/T) \\ &= -\frac{\log_e 2}{T} t + \log_e N_0\end{aligned}$$

If we put  $y = \log_e N$  and  $x = t$  we obtain the equation of a straight line with gradient  $-\frac{\log_e 2}{T}$  and y-intercept  $\log_e N_0$ .

◆ The amount of a radioactive material present at two different times in a sample obeying Equation 4

$$N = N_0(1/2)^{t/T} \quad (\text{Eqn 4})$$

is shown in Table 9. What is the half-life of the radioactive material, and what is the number of radioactive nuclei present in the sample at time  $t = 0$ ?



**Table 9**

Time $t$ /hours	Number of radioactive nuclei $N/10^8$
1.00	1500
2.00	1000



### Question T10

The measured amount of a radioactive material present at various times in a sample that obeys Equation 4

$$N = N_0(1/2)^{t/T} \quad (\text{Eqn 4})$$

is shown in Table 11. What is the half-life of the radioactive material? ☐



**Table 11** See Question T10.

Time $t$ /hours	Number of radioactive nuclei $N/10^8$
0.00	2150
1.00	1600
2.75	1020
4.00	700
5.50	490
6.50	390
7.25	300

## 5 Incorporation of data errors on a graph

### 5.1 Showing the likely data errors on the graph

When reporting experimental results, it is important to indicate the likely errors, or confidence limits, in any measured quantity. The most difficult part of this procedure is to assess how large the error in a measurement could be, and this is dealt with elsewhere in *FLAP*. However, once an error has been estimated, it is straightforward to represent it on a graph. Consider again the copper wire experiment that was described in Subsection 3.2. Table 4 shows that a 5.0 kg mass produces an extension of 0.2 mm, and 10.0 kg produces an extension of 0.5 mm.

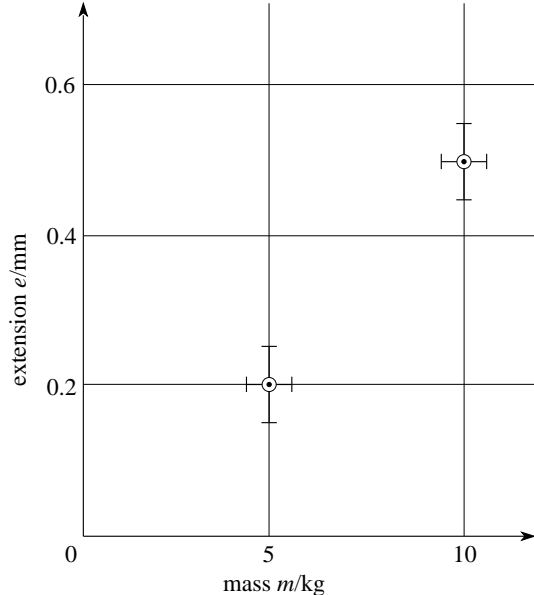
**Table 4** Experimental results for the extension of a loaded copper wire.

Mass $m/\text{kg}$	Extension $e/\text{mm}$	Mass $m/\text{kg}$	Extension $e/\text{mm}$
5.0	0.2	32.5	1.7
10.0	0.5	35.0	1.8
15.0	0.8	37.5	1.9
20.0	1.0	40.0	2.0
22.5	1.5	42.5	2.3
25.0	1.3	45.0	2.5
27.5	1.4	47.5	2.8
30.0	1.5	50.0	3.2

Suppose you have estimated that the uncertainty in measuring the extension ( $e$ ) is  $\pm 0.05$  mm and the uncertainty in the mass ( $m$ ) is  $\pm 0.6$  kg. (We may often write these errors as  $\Delta e = 0.05$  mm and  $\Delta m = 0.6$  kg.)

The correct way to plot these results is shown in Figure 14. The circled points indicate the measured values.

The uncertainty in each extension is indicated by a vertical **error bar** which extends 0.05 mm above and below the circled point, whereas the mass uncertainty is indicated by horizontal error bars which extend 0.6 kg on either side of the plotted mass value. Generally, both horizontal and vertical error bars should be shown, and they should only be omitted if the associated uncertainty is too small to show up on the graph. Plotting error bars is slightly more complicated if you are not plotting a measured quantity  $x$  directly, but are plotting  $x^2$ ,  $x^3$ ,  $\sin x$ , etc. because in such cases we need to estimate the error in  $x^2$  say, in terms of the known error in  $x$ .



**Figure 14** The first two points from the data of Table 4, plotted with vertical and horizontal error bars.

We will illustrate the procedure by taking as an example an experiment in which the speed  $v$  of a falling ball is measured after it has dropped through various distances  $s$ . The results of the experiment are tabulated in the first two columns of Table 12, and we will assume that the possible error in the measured speeds is  $\pm 0.5 \text{ m s}^{-1}$ , while the possible error in the measured distances is  $\pm 0.2 \text{ m}$ .

**Table 12** Data on a falling ball.

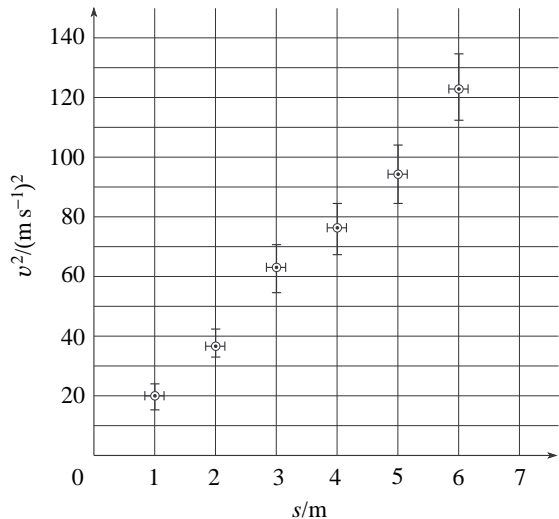
A	B	C	D	E	F	G
distance $s$ /m	speed $v$ /m s <sup>-1</sup>	$v^2$ /(m s <sup>-1</sup> ) <sup>2</sup>	$s + \Delta s$ /m	$s - \Delta s$ /m	$(v + \Delta v)^2$ /(m s <sup>-1</sup> ) <sup>2</sup>	$(v - \Delta v)^2$ /(m s <sup>-1</sup> ) <sup>2</sup>
1.0	4.4	19.4	1.2	0.8	24.0	15.2
2.0	6.0	36.0	2.2	1.8	42.3	30.3
3.0	7.9	62.4	3.2	2.8	70.6	54.8
4.0	8.7	75.7	4.2	3.8	84.6	67.2
5.0	9.7	94.1	5.2	4.8	104.0	84.6
6.0	11.1	123.2	6.2	5.8	134.6	112.4

Suppose now that we have reason to believe that  $v^2$  is proportional to  $s$  (perhaps from the physics), so that we choose to plot  $s$  on the horizontal axis and  $v^2$  on the vertical axis. Using the data in columns A and C of Table 12 we get the plot shown by the circled points in Figure 15.

Working out the error bars on  $s$  is straightforward — we simply work out  $s + \Delta s$  and  $s - \Delta s$  (columns D and E in Table 12), plot both against  $v^2$ , and draw the error bar between these extremes.

To produce the error bars on  $v^2$ , we work out  $(v + \Delta v)^2$  and  $(v - \Delta v)^2$  (as shown in columns F and G of Table 12), plot them against  $s$ , and draw the error bar between these extremes, as shown in Figure 15.

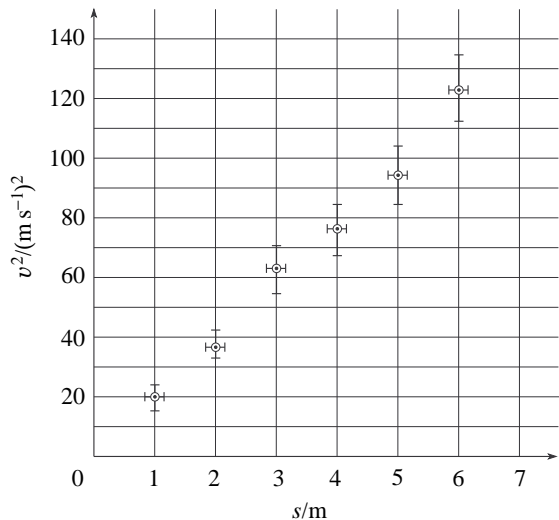
The important point to note here is that if we estimate that there is a possible error of  $\pm\Delta v$  in the measured value of  $v$ , then we mean that we are reasonably confident that  $v$  lies within the range  $v - \Delta v$  to  $v + \Delta v$ .



**Figure 15** A graph of the results in Table 12, showing the error bars for  $v^2$  and for  $s$ .

It therefore follows that the  $v^2$  should lie within the range  $(v - \Delta v)^2$  to  $(v + \Delta v)^2$ , and so we use these values to determine the error bars. Note that even though the size of the error in  $v$  is the same for all values of  $v$ , the errors in  $v^2$  become larger as  $v$  becomes larger.

◆ Show that the error in  $v^2$  is approximately  $2v\Delta v$ .

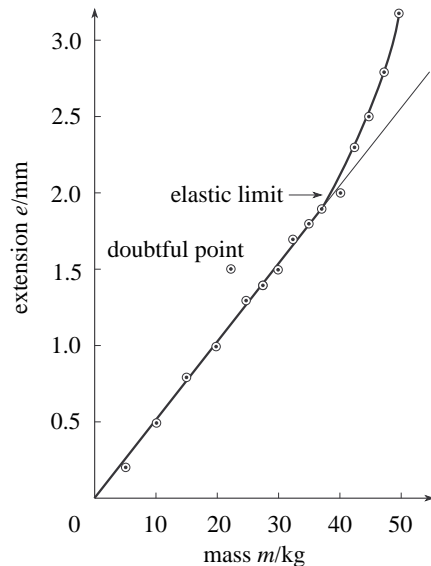


**Figure 15** A graph of the results in Table 12, showing the error bars for  $v^2$  and for  $s$ .

The presence of error bars on a graph serves a number of useful purposes, such as finding the best line choice, spotting false points or estimating the error on the slope or gradient, as we will see shortly.

### Overestimation and underestimation of errors

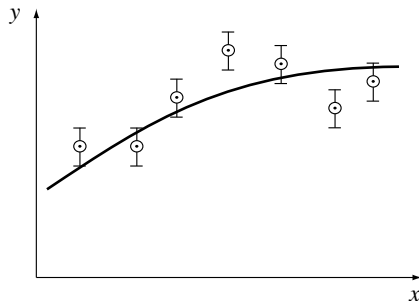
We may expect that the graph is a smooth curve, in which case we would also expect the data points to be scattered around that curve by amounts ranging up to the size of the error bar. For example, in the experiment with the copper wire (Subsection 3.2, Figure 4) we can deduce that the errors, or uncertainties, in the experimental measurements must be about  $\pm 0.05$  mm in the extension, and/or  $\pm 0.6$  kg in the mass.



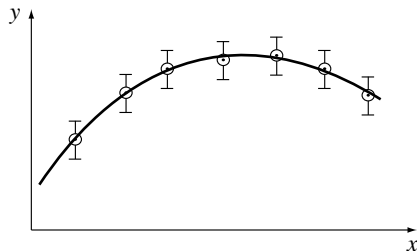
**Figure 4** Extension of a loaded copper wire.

Should the results deviate from a smooth curve by much more than the error bars, as shown in Figure 16, then, either we have underestimated the errors, or the assumption that a smooth curve should describe the results is not valid.

On the other hand, if all the results deviate from the expected curve by much less than the error bars, as in Figure 17, then we might well have overestimated the likely errors.



**Figure 16** Have the error bars been underestimated?



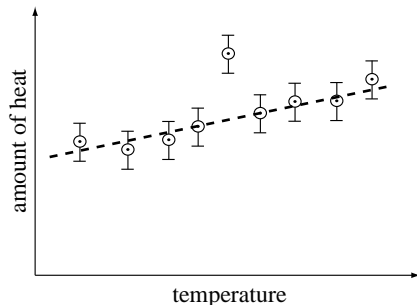
**Figure 17** Have the error bars been overestimated?



## Identifying errors

Error bars are also very helpful in identifying mistakes in measurements or the plotting of points. For example, Figure 18 shows measurements of the amount of [heat](#) needed to raise the [temperature](#) of 1 kg of a rare metal known as gadolinium by  $1^{\circ}\text{C}$  at various temperatures. All results lie on the broken line except for one, which deviates from the straight line by about three times the magnitude of the error bar. This immediately suggests a mistake or something interesting about gadolinium!

In cases like this, you should first check that you have plotted the points correctly; then check any calculations made to obtain the plotted values. If these do not show up a mistake, then the measurements that gave rise to the suspect point should be repeated. In some cases, of course, repeating the measurements is just not possible, and one is left with the difficult decision about whether to ignore the point or not. No fixed rule can be made about this, but you should always be aware that what appears to be an anomalous measurement may indicate a real (and possibly as yet undiscovered) effect.

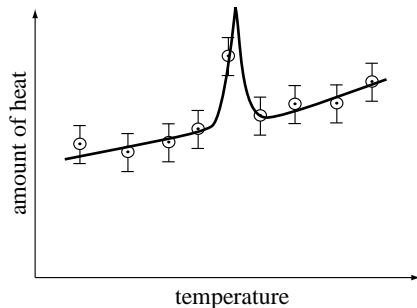


**Figure 18** An apparently odd point when gadolinium is heated.

In the case of gadolinium, more detailed measurements show that the real behaviour is as shown by the solid line in Figure 19. By ignoring the apparently anomalous result, we would have overlooked a real anomaly (which is, in fact, caused by the magnetic properties of gadolinium). It is an interesting property of gadolinium after all!

An important piece of advice follows from this: if at all possible, plot a graph while you are taking measurements. If you do this, you can immediately check for errors and investigate any anomalies that appear.

Plotting as you go along also helps to ensure that the results are reasonably distributed on the graphs, and this usually means having them equally spaced except where there are interesting features, such as in Figure 19.



**Figure 19** An anomaly in the specific heat of gadolinium, which could have been overlooked if one point had been discarded.

## 5.2 Estimating the errors in the gradient and the intercept

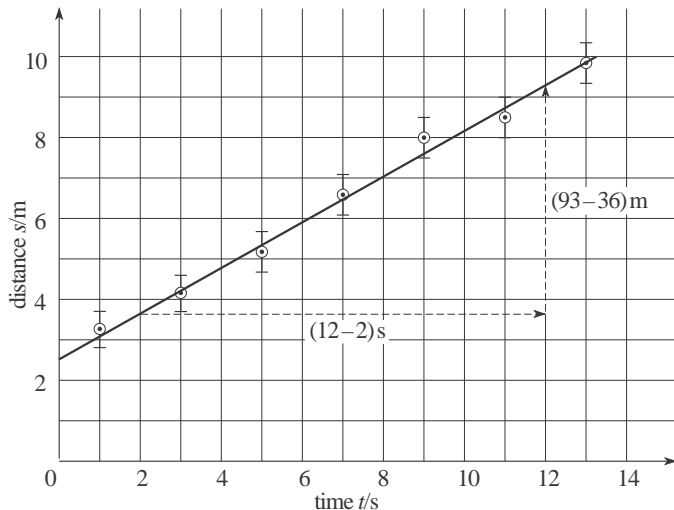
One particularly useful and important feature of plotting error bars on graphs is that it enables us to estimate the errors in the gradient and intercept. The next example shows how.

### The gradient

Figure 20 shows a graph of the distance travelled by an object over a period of 13 s. But how are we to estimate the uncertainty in the gradient of the straight line?


The gradient of the straight-line graph drawn in Figure 20 is given by

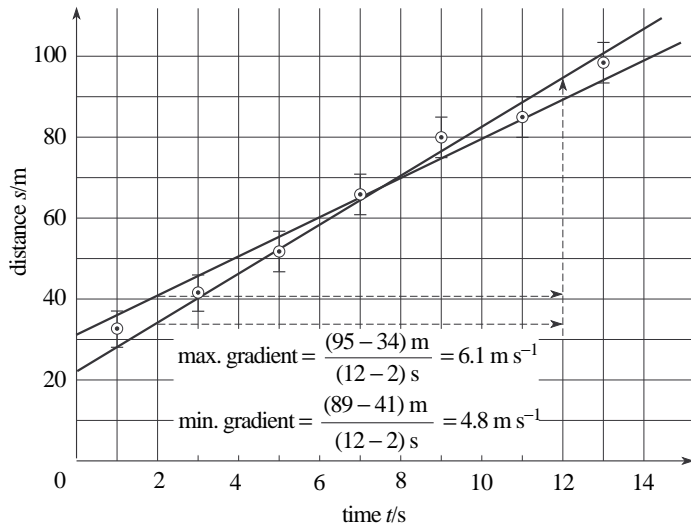
$$\begin{aligned}\text{gradient} &= \text{rise/run} \quad \text{👉} \\ &= \frac{(93 - 36) \text{ m}}{(12 - 2) \text{ s}} = 5.7 \text{ m s}^{-1}\end{aligned}$$



**Figure 20** How to measure the gradient of the ‘best’ straight line through the points.

The straight line drawn in Figure 20 is the one we think is best fitted to the results plotted. Some points above the line are balanced by other points below the line. However, the error bars indicate the uncertainty in the experimental results, and other lines with different gradients could be drawn to pass through all of the error bars.

So, in order to estimate the uncertainty in the gradient, we draw lines that pass through *all* of the error bars with the maximum and minimum possible gradients.  These lines are shown in Figure 21, and their gradients are  $6.1 \text{ m s}^{-1}$  and  $4.8 \text{ m s}^{-1}$ , respectively. These values differ from the gradient of the 'best' straight line by  $0.4 \text{ m s}^{-1}$  and by  $0.9 \text{ m s}^{-1}$  an average difference of  $(0.4 + 0.9) \text{ m s}^{-1} / 2 = 0.65 \text{ m s}^{-1} \approx 0.7 \text{ m s}^{-1}$ . We therefore quote the value of the gradient as  $(5.7 \pm 0.7) \text{ m s}^{-1}$ .



**Figure 21** How to estimate the uncertainty in the gradient.

## The intercept on the vertical axis

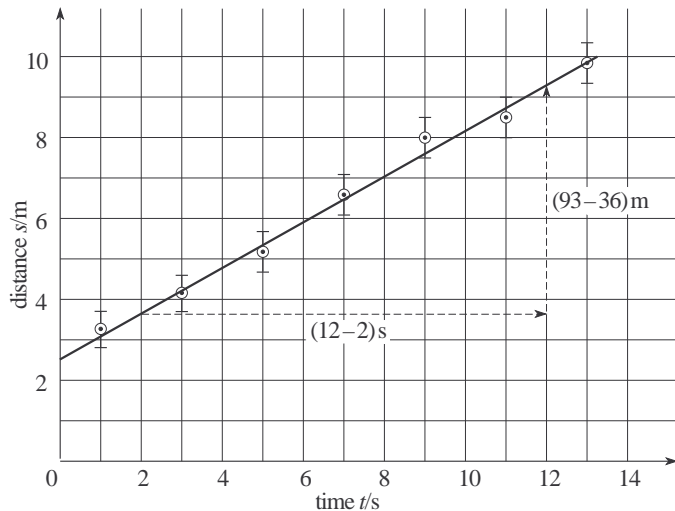
The other quantity that is normally used to specify the position of a straight line on a graph is its intercept with the vertical axis. We have seen earlier in this module that the equation of a straight line can be written in the form:

$$y = ax + b$$

where  $a$  is the gradient of the line and  $b$  is the intercept with the  $y$ -axis — that is, the value of  $y$  when  $x = 0$ .

In the case of the results shown in Figure 20, the intercept is at  $y = 25$  m and, since this is determined from the 'best' straight line, we regard it as the best estimate of the intercept.

(Note that again we must quote the appropriate units.)

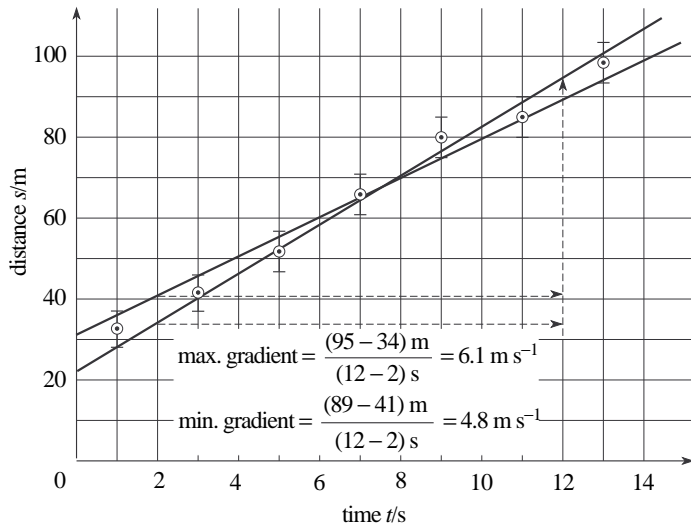


**Figure 20** How to measure the gradient of the 'best' straight line through the points.

The maximum and minimum likely values of the intercepts are found by drawing other lines through the error bars, in this case the lines with maximum and minimum slopes shown in Figure 21. The intercepts of these lines are 31 m and 22 m respectively, and the mean difference between these values and the best intercept is

$$\begin{aligned} & \frac{(31 - 25) \text{ m} + (25 - 22) \text{ m}}{2} \\ &= \frac{(6 + 3) \text{ m}}{2} \approx 5.0 \text{ m} \end{aligned}$$

We can, therefore, quote the experimentally determined intercept as  $(25 \pm 5) \text{ m}$ .



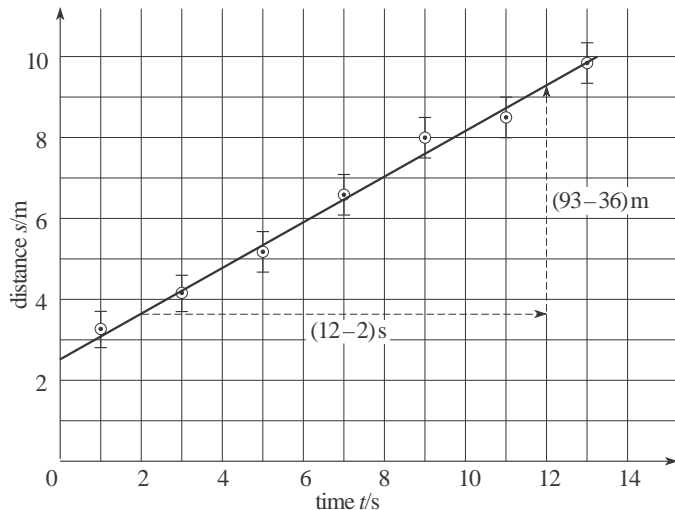
**Figure 21** How to estimate the uncertainty in the gradient.

Having now determined both gradient and intercept and the possible errors in each we can summarize the experimental results in Figure 20 very succinctly in the form of an equation, namely:

$$s = (5.7 \pm 0.7) \text{ m s}^{-1} \times t \\ + (25 \pm 5) \text{ m}$$

### Question T11

If the errors in the extensions measured in our rubber band experiment (Subsection 3.1) are each  $\pm 2 \text{ mm}$ , what is the error in the value of the proportionality constant  $k$ ? ☐



**Figure 20** How to measure the gradient of the 'best' straight line through the points.




## 6 Numerical analysis of gradients and intercepts

### 6.1 Choice of the ‘best’ line by numerical means — the method of least squares

Although it is often quite satisfactory to draw a ‘best line’ by eye, sometimes it is useful to have a numerical method of estimating the gradient and intercept of the ‘best’ line. Suppose we have a set of  $N$  readings

$(x_1, y_1), (x_2, y_2), (x_3, y_3)$  and so on, up to  $(x_N, y_N)$

We first find the mean of all the  $x$  values,  $\langle x \rangle$ , and then the mean of all the  $y$  values,  $\langle y \rangle$ . 

It can be shown that the best estimates of the gradient and intercept of the ‘best’ line are then given by:

$$a = \frac{\sum_{i=1}^N (x_i y_i) - N \langle x \rangle \langle y \rangle}{\sum_{i=1}^N (x_i^2) - N \langle x \rangle^2} \quad (5)$$



and  $b = \langle y \rangle - a \langle x \rangle \quad (6)$

This ‘best’ line will go through the point  $(\langle x \rangle, \langle y \rangle)$ —effectively the ‘centre of gravity’ of the points. It is drawn so as to minimize the sum of the squares of the deviations of the points from the line. For that reason, this method of estimating  $a$  and  $b$  is called the method of least squares.

◆ Find the ‘best’ straight line fit to the following data points

$(1.1, 2.5), (1.9, 3.2), (3.1, 3.9)$




## Question T12

Use the above method to calculate a line of best fit to the data in Table 14 (see the rubber band experiment in Subsection 3.1, i.e. Table 3)

(Hint: Start by completing Table 14 then use Equations 5 and 6

$$a = \frac{\sum_{i=1}^N (x_i y_i) - N \langle x \rangle \langle y \rangle}{\sum_{i=1}^N (x_i^2) - N \langle x \rangle^2} \quad (\text{Eqn 5})$$

and  $b = \langle y \rangle - a \langle x \rangle \quad (\text{Eqn 6})$

to find the equation of the line.) 



**Table 14** See Question T12.

$i$	Mass $x_i$ /kg	Extension $y_i$ /mm	$x_i y_i$	$x_i^2$
1	0	0		
2	1	13		
3	2	28		
4	3	44		
5	4	58		
6	5	74		
7	6	89		
8	7	105		
9	8	120		
$\sum_{i=1}^9 x_i =$		$\sum_{i=1}^9 y_i =$	$\sum_{i=1}^9 (x_i y_i) =$	$\sum_{i=1}^9 (x_i^2) =$
$\langle x \rangle =$		$\langle y \rangle =$		

## 6.2 A cautionary note

The above calculation may be tedious, but it is quite easy, and we are helped by the fact that many modern calculators have a program to do it. However, the method should always be used with caution, and definitely not used without having a good look at the data first!

There is no substitute for plotting the data to see what it looks like. It may be that a straight line is not appropriate; on the other hand, a straight line may give a good fit, but there may be one (or more) anomalous points that need to be looked at very carefully and perhaps not used in the calculation.

The formulae given in Subsection 6.1 make two very important assumptions: (i) the errors in the  $x$  (independent) variable are negligible; and (ii) the errors in  $y$  are the same size for each point. If these conditions do not apply, you must resort to more sophisticated ways of dealing with the problem that are beyond the scope of this module.

## 7 Closing items

### 7.1 Module summary

- 1 Graphs are drawn on two *axes*, the horizontal axis (often called the  $x$ -axis) for the *independent variable*; the vertical axis (often called the  $y$ -axis) for the *dependent variable*.
- 2 When plotting graphs:
  - suitable scales should be chosen, to take full advantage of the graph paper and to make plotting as easy as possible;
  - arrows should be used on the axes to indicate the directions in which the variables increase;
  - variables and units of measurements should be indicated on each axis;
  - data points should be indicated clearly (preferably with a dot within a circle);
  - points should generally be evenly spread.
- 3 Data points should be plotted as the experiment proceeds.
- 4 A straight line, or smooth curve, should be drawn through the data points in most cases.
- 5 The standard equation of a straight line is  $y = ax + b$  where  $a$  is the *gradient* of the line and  $b$  is the *intercept* on the vertical or  $y$ -axis.

- 6 To estimate the gradient of a linear (i.e. straight-line) graph, determine the coordinates of any two points on the graph  $(x_p, y_p)$  and  $(x_q, y_q)$ , and then use the formula

$$\text{gradient} = \frac{\text{rise}}{\text{run}} = \frac{y_q - y_p}{x_q - x_p}$$

- 7 The graph corresponding to the equation

$$y = ax^2 + bx + c \text{ is a } \text{parabola};$$

and the graph corresponding to

$$y = \frac{a}{x} + b \text{ is an } \text{hyperbola}.$$

- 8 A change of variables can sometimes be used to convert a non-linear relationship into linear form. This process is called linearization.
- 9 Power laws and exponential laws of the form  $y = ax^n$  and  $y = ka^x$  can be linearized by the use of logarithms.
- 10 The error in the gradient and the intercept of a straight-line graph can be estimated using error bars.

- 11 The *method of least squares* provides a numerical technique of estimating the gradient and intercept of the straight line that best fits a set of data points. According to this method

$$\text{gradient} = a = \frac{\sum_{i=1}^N (x_i y_i) - N \langle x \rangle \langle y \rangle}{\sum_{i=1}^N (x_i^2) - N \langle x \rangle^2} \quad (\text{Eqn 5})$$

and      intercept =  $b = \langle y \rangle - a \langle x \rangle$       (Eqn 6)

## 7.2 Achievements

Having completed this module, you should be able to:

- A1 Define the terms that are emboldened and flagged in the margins of the module.
- A2 Plot a graph from a set of data.
- A3 Draw the ‘best’ line through a set of data points, if this is appropriate, and *estimate* the gradient and intercept of this line.
- A4 Obtain the equation of a straight-line graph, and rearrange the equation of a line into the standard form  $y = ax + b$ .
- A5 Sketch the graphs of any linear function and the functions discussed in this module.
- A6 Convert non-linear equations into linear form by a change of variable in simple cases.
- A7 Use logarithms to linearize power laws and exponential laws.
- A8 Plot error bars on graphs, and find the error bars under a change of variables.
- A9 Estimate the error in the slope and intercept of a straight-line graph when the error bars have been plotted.
- A10 Calculate the ‘best’ estimates of intercept and slope using a numerical method (the method of least squares).



***Study comment*** You may now wish to take the [Exit test](#) for this module which tests these Achievements. If you prefer to study the module further before taking this test then return to the [Module contents](#) to review some of the topics.

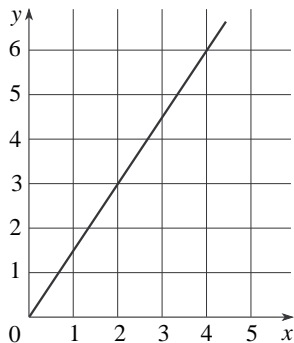
## 7.3 Exit test

*Study comment* Having completed this module, you should be able to answer the following questions, each of which tests one or more of the *Achievements*.

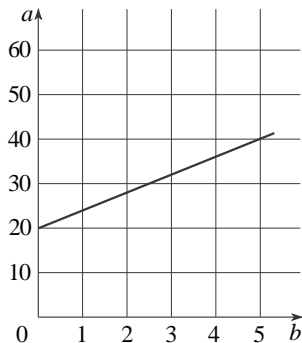
### Question E1

(A4) This question refers to Figure 22. (a) Find the gradient of each graph shown in (i) to (iv).

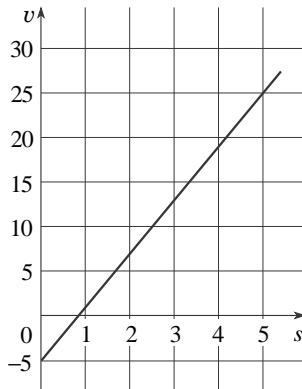
(b) Find the equation represented by each graph (i) to (iv).



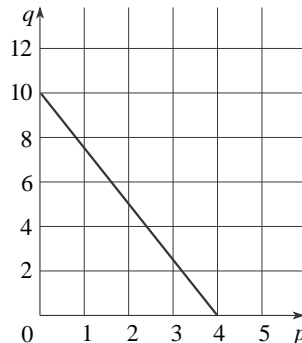
(i)



(ii)



(iii)



(iv)

**Figure 22** See Question E1.



## Question E2

(A5) Sketch graphs to represent the following equations:

(a)  $y = 4x$ , (b)  $e = 2f + 1$ , (c)  $g = (h/3) - 1$ , (d)  $k = -2m + 3$ .



## Question E3

(A2 and A5) Table 15 shows the percentage of men surviving until various ages (figures relate to data current in 1990). Plot the results in graphical form.



**Table 15** See Question E3.

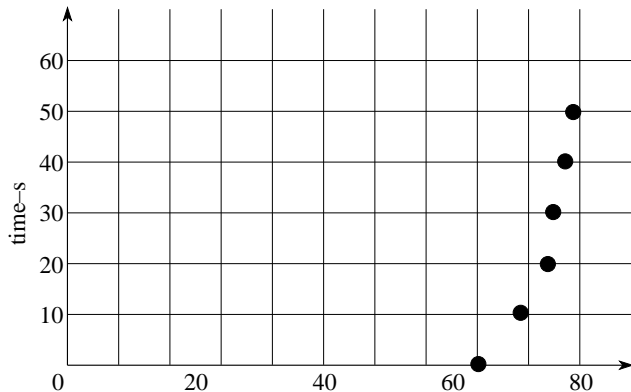
Age/years	Percentage surviving/%
51	89.3
57	83.5
63	73.1
69	58.2
75	40.3
81	20.6

## Question E4

(A2) Having noted the [speed](#) of his car at various times (Table 16), a naive graph-plotter presents his measurements in the way shown in Figure 23. His graph can be criticised on at least five counts. Point out the short-comings, and correct them by re-plotting the results given in Table 16.

**Table 16** See Question E4.

Time/s	Speed /miles h <sup>-1</sup>
0	64
10	69
20	73
30	76
40	78
50	79



**Figure 23** See Question E4.

### Question E5

(A2, A3, A8 and A9) A ball moves in a straight line with its [speed](#) increasing linearly with time (constant [acceleration](#)). Table 17 gives the speed of a ball at six different times (with the associated errors).

- (a) Plot the data given in Table 17 on a graph.
- (b) Use your graph to find the gradient of the graph (the magnitude of the acceleration of the ball), and its associated uncertainty. (The magnitude of the acceleration is given by the gradient of the graph.)



**Table 17** See Question E5.

Time/s	Speed/m s <sup>-1</sup>
$1.0 \pm 0.2$	$33 \pm 5$
$2.2 \pm 0.2$	$76 \pm 5$
$4.6 \pm 0.2$	$156 \pm 10$
$6.4 \pm 0.2$	$210 \pm 15$
$7.2 \pm 0.2$	$240 \pm 20$

### Question E6

(A2, A3 and A6) Table 18 shows results of a series of experiments in which the frequency  $f$  of vibration of a plastic ruler is measured for various values of the length  $l$  of the ruler overhanging the edge of a table. Plot a graph of  $\log_{10}(f/\text{Hz})$  against  $\log_{10}(l/\text{cm})$  and decide whether the frequency and length are related by a power law, of the form  $f = kl^\alpha$  for some constant  $k$ ; if they are, determine the value of  $\alpha$ .



**Table 18** See Question E6.

Length $l/\text{cm}$	Frequency $f/\text{Hz}$
5	79
10	19
15	9
20	4.7

### Question E7

(A10) Use the method of least squares to estimate the gradient and intercept of the straight line you produced in Question E6.



**Study comment** This is the final *Exit test* question. When you have completed the *Exit test* go back to Subsection 1.2 and try the [\*Fast track questions\*](#) if you have not already done so.

If you have completed **both** the *Fast track questions* and the *Exit test*, then you have finished the module and may leave it here.

