Extension 1: More about Numbers, Units and Measurements

"I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind…." Lord Kelvin, 1883

1. Introduction to Unit 1

It may be tempting to miss this unit out, but we present it at the beginning of the book for a good reason: you must be able to manipulate numbers in many of the units in this book. Your calculator is a vital part of your tool kit, but some calculators operate differently to others, so be ready to ask a friend if you get stuck, particularly when it comes to using logs and exponents. If you lack confidence with numbers, we recommend you take time to try out the exercises in Unit 1 before tackling the rest of the book.

Numbers are the way that scientists demonstrate the size of quantities and the uncertainties involved in their measurement. The range of the size of quantities is huge, from the incredibly small to the incredibly large. As an example of the incredibly small, consider the geometry of atoms and molecules, which is measured in picometres:

$$1 \text{ pm} = 10^{-12} \text{ m}$$

A molecule of hydrogen chloride, HCl, is 127 pm long. As you may easily verify with your calculator, a line of HCl molecules 1 cm (0.01 m) long contains 79 million HCl molecules.

An example of an incredibly large quantity is the number of carbon atoms in 12 g of carbon:

$$6\ 00\ 000\ 000\ 000\ 000\ 000\ 000\ 000 = 6 \times 10^{23}$$

But this unit is not just about numbers, it is also – like the rest of the book - about the language of science as expressed through definitions. The words accuracy, (absolute) error, precision and uncertainty have particular meanings in science and these words are often used differently in everyday conversation. For example, in everyday conversation, the words accuracy and precision are sometimes used interchangeably, whereas they have very distinct meanings in science. Accuracy means the closeness of a measurement to the true measurement. Precision means how close repeat measurements are to each other. To illustrate this, suppose that the true concentration of sodium ions in a solution is 10.0 parts per million by mass:

$$[Na^+] = 10.0 \text{ ppm}$$

and laboratory measurements on three samples of the solution give concentrations of 10.1, 10.0 and 9.9 ppm:

10.1 ppm 10.0 ppm 10.0 ppm

These repeat measurements are close enough to each other to justify stating that the measurements are precise. The mean concentration (= 30.1/3 = 10.03 ppm) is very close to 10.0 ppm – the true concentration – so the mean measurement is also described as accurate.

The Extension material for Unit 1 consists of three parts. First, an example of the range of time encountered in science, from the incredibly fast processes involving electrons to the age of the universe. This will allow more practise in the use of scientific notation. Second, an explanation of the mathematics of straight lines, using examples from the book. Third, following on from Unit 1 in the book, the use of standard deviation and relative uncertainty in analysing the results of measurements in analytical chemistry.

Part 1: Time Scales in Science

2. Time scales in science

A few calculations put big numbers into perspective. Check them with your calculator as we go along.

We start with the age of the Universe, which cosmologists take to be the time back to the 'Big Bang' when our Universe began (Fig. 1.1).

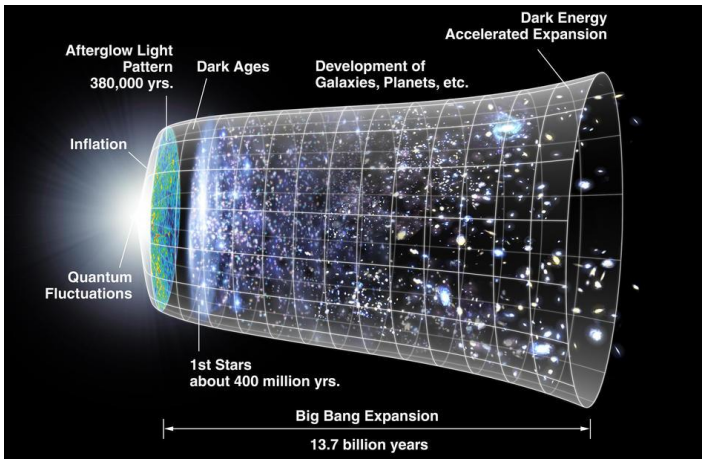


Fig. 1.1 Timescale of the Big Bang expansion of the Universe. Image courtesy of NASA.

The Universe is about 14 billion years old (1.4×10^{10} years). There are 60 seconds in a minute, 60 minutes in an hour, 24 hours in a day and 365.25 days in a year. Therefore, the Universe is:

$1.4 \times 10^{10} \times 365.5 \times 24 \times 60 \times 60 = 4.42 \times 10^{17}$ seconds old.

Supposing that a person lives for 70 years (or, $70 \times 365.5 \times 24 \times 60 \times 60 = 2.21 \times 10^{9}$ seconds). The ratio between these times is:

$$\frac{age\ of\ Universe}{human\ life\ span} = \frac{4.42 \times 10^{16}}{2.21 \times 10^{9}} = 2.00 \times 10^{8}$$

In other words, the Universe is the equivalent of 200 million human lives old.

The number of atoms in one mole of substance is even greater than the age of the Universe in seconds. This number is equal to Avogadro's Constant:

6.022×10^{23}

It has been estimated that there are about 10^{80} atoms in the universe. This is an incredibly large number – 1 followed by 80 zeros.

EXERCISE 1A

Suppose that the Universe consisted entirely of carbon. How much would the Universe weigh (in g)?

The span of timescales for processes is very great. Table 1 shows a few examples. The fastest process shown is the absorption of a photon of light by a molecule: 10^{-15} s is a femtosecond, such a short time that even light travels only 0.00001 cm during this time interval.

Table 1

SLOW

FAST

Process	*Timescale*
The expansion of the universe	Started 14 billion years ago and is continuing
Energy release by fusion in the Sun	Complete in about 10 billion years
Radioactive decay of Radium-226	A sample decays completely in about 16,000 years
Typical life expectancy of a human being	Birth to death in less than 100 years
Rusting of a nail outdoors	Typically, about 10 years
Rotation of the earth	Complete in one day
The denaturing of a protein (e.g. by boiling an egg)	3 minutes
Emission of light from a molecule	Variable, but over in 10^{-9} s to 100 s
Vibration of a bond within a molecule of HCl	Complete in 10^{-14} s
Absorption of a photon by a molecule	Complete in 10^{-15} s

EXERCISE 1B

Show that light only travels 0.00003 cm in 10^{-15} s.

Part 2: Graphs and the Equation of a Straight Line

3. Gradients and the equation of a straight line

Suppose that a quantity *y* is related to a quantity *x* by the equation:

$$y = 2x + 5 \qquad (1)$$

If x = 3 then y = 11. If x = 4, y = 13. If we plot y against values of x, we find that the graph is a straight line (see Fig. 1.2).

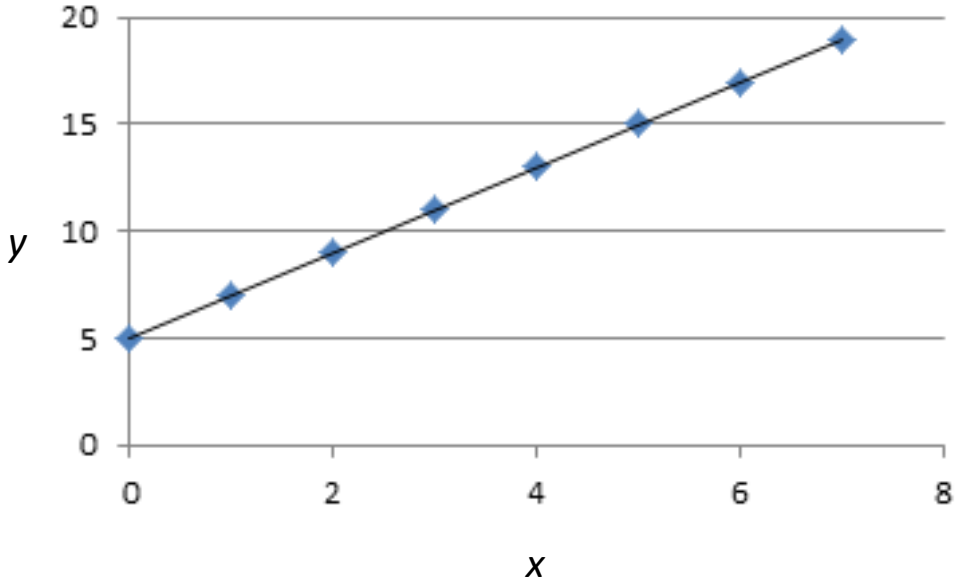
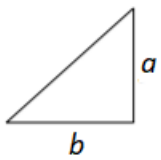


Fig. 1.2 Plot of $y=2x+5$ for the x values 0 to 7

Notice that the graph hits $x = 0$ at a y value of 5. Using the language of mathematics, we would say that the *intercept* of the graph occurs at $y = 5$. The symbol used for the intercept value of y is c. Here, $c = 5$.

The *gradient* (slope) of the graph shown in Fig 1.2 is calculated by dividing a perpendicular (upward) distance of the graph (a) by the corresponding horizontal ('across') distance (b). In general:



$$\text{gradient} = \frac{a}{b}$$

In fig 1.1, convenient values are $y = 5$ to $y = 19$ (so that distance $a = 19\text{-}5 = 14$) and $x = 0$ to $x = 7$ (so that distance $b = 7$) and:

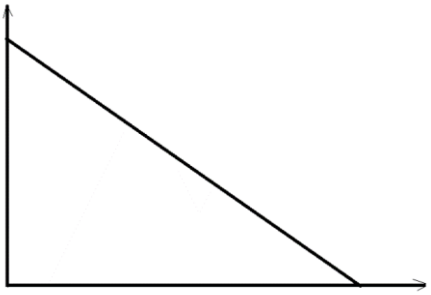
$$\text{gradient} = \frac{14}{7} = 2$$

However, it doesn't matter what values you choose because the gradient of a straight line is fixed.

It is no coincidence that the number 2 occurs in equation (1). For algebraic expressions which yield straight lines graphs, the gradient always appears in the equation in this way. The symbol for gradient is m, so that here $m = 2$.

The units of a gradient will be determined by the units of the axes in the graph. In our example, the axes are unitless (i.e. pure numbers) so that the *gradient* is unitless.

The gradient is also positive i.e. +2. By convention, a graph that slopes the other way (right to left):



is given a negative slope.

The gradient of a *curve* is **continuously changing**. An example is provided in chemical kinetics, where the rate of reactions are studied. For example, consider the reaction of hydrogen and iodine to make hydrogen iodide:

$$H_2(g) + I_2(g) \rightarrow 2HI(g)$$

As the reaction takes place, the concentration of HI increases, as shown in Fig 1.3.

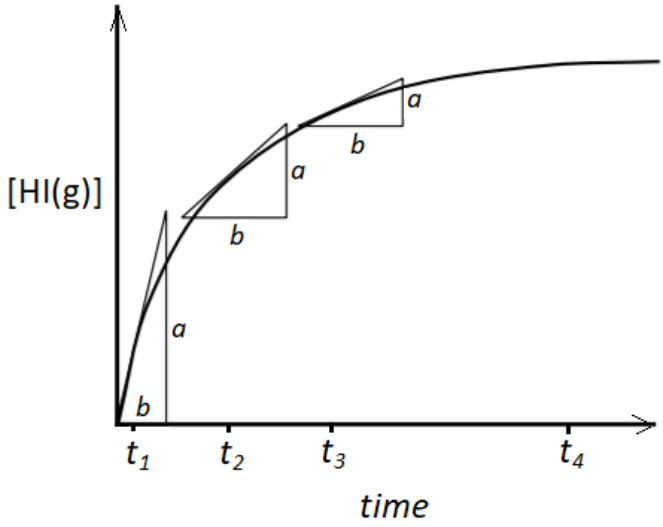


Fig. 1.3 Concentration of HI during the reaction of H_2 and I_2.

To illustrate the change in gradient in the curve, consider the three times shown in Fig 1.3. The gradient at time t_1, t_2, and t_3 is obtained by dividing distances *a* by *b* as shown. The relative size of a to b is greatest at time t_1. This means that the gradient at t_1 is greater than at t_2 or t_3 – corresponding to a higher reaction rate at earlier times. The maximum gradient would be at time zero[1]. At times beyond t_4 the curve has levelled out to a horizontal straight line and the gradient is zero. This means that the concentration of HI (the product) does not change at further times: the reaction is over and no more HI is being made.

What are the units of the gradient in Fig. 1.3? Since the units on the y axis is concentration (mol dm^{-3}) and the unit of the x-axis is time (in seconds), the units of the gradient is:

[1] It might seem odd to say that the initial rate (at $t = 0$) is the maximum rate, since if the reaction has not begun the rate will be zero! However, when scientists refer to a gradient at time zero, they mean the gradient a tiny fraction of a second after $t = 0$. For example, if reaction rate changes appreciably over 1000 s (about 17 mins), then the gradient for the concentration change 0 to 5 s will approximate to the initial rate.

$$\frac{\text{mol dm}^{-3}}{\text{s}} = \text{mol dm}^{-3}\,\text{s}^{-1}$$

This is read as 'moles per dm³ per second'.

4. The equation of a straight line

The general equation of a straight line (not a curve) is:

$$y = mx + c \qquad (2)$$

where *m* and *c* is the gradient and intercept of the line. In our example:

$$y = 2x + 5$$

in which $mx = 2x$ and *c* = 5. Note that in the equation:

$$y = 6x \qquad (3)$$

c = 0, meaning that the plot goes through the coordinate *y* = 0 and *x* = 0. This is more obvious if we write equation (3) as:

$$y = 6x + 0$$

It is often very useful to arrange a formula in the form of equation (2) and examples are shown in Table 2.

Table 2 Examples of the use of *y* = *mx* + *c* to be found in the book. In these examples, *c* (the intercept) is zero.

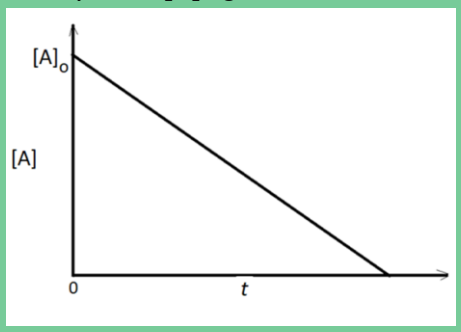
Equation	*What corresponds to the quantities in y = mx + c?*	*Book page number*
Concentration of dissolved gas, *c*, is related to the partial pressure of the dissolving gas by: $\boldsymbol{c = K_H \times p}$ Where K_H is a constant at that temperature. A plot of *c* versus *p* is a straight line.	*y is c, the concentration of the dissolved gas* *m is* K_H *c is zero*	186
The variation of the hydronium ion concentration $[H_3O^+(aq)]$ with the concentrations of acid (C_A) and salt (C_S). $K_{a(T)}$ is a constant at that temperature: $\boldsymbol{[H_3O^+(aq)] \approx \frac{C_A \times K_{a(T)}}{C_S}}$ (The ≈ symbol means 'approximately'). A plot of measured $[H_3O^+(aq)]$ against values of C_A and C_S is a straight line.	*y* is the measured $[H_3O^+(aq)]$ *m* is $K_{a(T)}$ *c* is zero	310
The Beer-Lambert Law: $\boldsymbol{A_\lambda = \varepsilon_\lambda \times c \times b}$ where *b* is the sample cell thickness and ε_λ is a constant at that wavelength. A plot of sample absorbance A_λ (at wavelength λ) against absorber concentration (*c*) gives a straight line.	*y* is the sample absorbance, A_λ *m* is $\varepsilon_\lambda b$. Since you know *b*, the cell thickness, the gradient allows ε_λ to be calculated. *c* is zero.	406

EXERCISE 1C

The fall in the concentration of alcohol in blood in an individual follows the equation:

$$[A] = [A]_o - kt \qquad (4)$$

where [A] is the concentration of alcohol at time t, $[A]_o$ is the concentration of alcohol initially (i.e. at $t = 0$) and k is a constant. A plot of [A] against t is shown below:



In terms of equation (4), what is the value of (i) the gradient of the plot; and (ii) the intercept?

Part 3: Standard Deviation, Precision and Relative Uncertainty

5. Standard deviation and precision

The following sections deal with the ideas of precision and accuracy in greater detail and describe the use of standard deviation in evaluating analytical data.

The greater the degree of scatter of a set of measurements the lower is the precision of those measurements. Precision is often measured by the standard deviation of the set. The standard deviation s of a set of n repeat measurements is defined as

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

where x is a single measurement and $\bar{x}$ is the mean (average) measurement. The symbol $\sum$ means 'sum of'.

The lower a standard deviation of a set of n repeat measurements, the better the precision of those measurements.

The relative precision of two or more methods of measurement (e.g. of two analytical methods) is compared by calculating their percent relative standard deviation. The **percent relative standard deviation** (%RSD) is calculated from the standard deviation s and mean measurement $\bar{x}$ using the equation:

$$\%\text{RSD} = \frac{100 \times s}{\bar{x}}$$

The standard deviation enables us to estimate the spread of measurements on either side of the mean measurement. If the only errors in the measurements are random ones, there is a 90% chance that the mean value of a measured quantity will be within the confidence limits

$$\bar{x} \pm 2\frac{s}{\sqrt{n}}$$

However, to rely on such a calculation we need to be sure of our value of s: if the spread of measurements is considerable, we may have to make 10 or more repeat measurements.

EXAMPLE 1

Repeat measurements of the percentage by volume of benzene in a fuel gave the following results: 5.5, 5.0, 5.3, 5.0, 5.2 and 5.5. **(i)** Calculate the standard deviation of the results. **(ii)** Estimate the confidence limits of the mean concentration. **(iii)** Calculate the %RSD of the measurement.

Answer

(i) Standard deviation

mean = $x = \frac{(5.5+5.0+5.3+5.0+5.2+5.5)}{6}$ = 5.3%

x	$(x-\bar{x})$	$(x-\bar{x})^2$
5.5	0.2	0.04
5.0	-0.3	0.09
5.3	0.0	0.00
5.0	-0.3	0.09
5.2	-0.1	0.01
5.5	0.2	0.04
	sum =	0.27

$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{0.27}{n-1}} = \sqrt{\frac{0.27}{5}}$ = 0.23%

Note that the units of standard deviation are the same as that of the quantity being measured – here, the percentage.

(ii) Range of the mean concentration

range = $\bar{x} \pm 2\frac{s}{\sqrt{n}}$ = 5.3 ± (2 × 0.23/(2.45)) ≈ 5.3 ± 0.2%

We conclude that provided the systematic errors are absent and that we are confident we have taken sufficient measurements, there is a 90% chance that the mean benzene concentration is in the range 5.5-5.1%.

(iii) %RSD = $\frac{100 \times s}{\bar{x}}$ = 100 × 0.23/5.3 = 4.3%

6. The relative uncertainty of a measurement

In s.5 in Unit 1 we briefly introduced the idea of *error* and also *measurement uncertainty*. The measurement uncertainty is 'the ± part' of the reported measurement. The definition of the term absolute error is:

absolute error = experimental value – true value

Uncertainty and absolute error are closely connected. The uncertainty is a measurement of the range of the error that is likely to apply to the measurement. For example, if the concentration of sulfuric acid in a solution is measured and the mean measurement reported as $[H_2SO_4(aq)]$ = 1.000 ± 0.002 mol dm^{-1}, then the uncertainty is the 'plus and minus bit':

uncertainty = ± 0.002 mol dm^{-1}

The uncertainty is the same as the range of the absolute error, which is ± 0.002 mol dm^{-3}. In other words, the maximum absolute error is expected to be + 0.002 mol dm^{-1} and the minimum absolute error is expected to be -0.002 mol dm^{-1}. Put another way, the absolute error is between 0.002 mol dm^{-1} *above* the measurement and 0.002 mol dm^{-1} *below* the mean measurement.

The Percentage Relatively Uncertainty of a measurement is often quoted. The Percentage Relative Uncertainty, symbolized % s_r, is defined as:

$$Percentage\ Relative\ Uncertainty = 100 \text{ x } \frac{\text{uncertainty}}{\text{reported measurement}}$$

s_r increases as the uncertainty increases, or as the reported measurement decreases.

Applying this to the measurement of the sulfuric acid concentration:

$$\%\ s_r = 100 \text{ x } \frac{0.002}{1.000} = 0.2\ \%$$

The advantage of calculating % s_r is that it indicates the scale of the doubt about a measurement. An uncertainty of 0.002 is not huge when the reported mean measurement is 1.000, but if the concentration was reported as 0.001 mol dm^{-3} (and the same uncertainty was estimated) then:

$$\%\ s_r = 100 \text{ x } \frac{0.002}{0.0010} = 200\ \%$$

Depending on the circumstances, that might be a worryingly high % s_r.

7. Using a calculator or a spreadsheet to calculate standard deviations

It is very useful to be able to calculate standard deviations using your pocket calculator: study the calculator instructions – or as a knowledgeable friend!

A spreadsheet, such as *Microsoft Excel*, may also be used to carry out the calculations. To illustrate this, the following *Microsoft Excel* spreadsheet shows calculations of the mean, standard deviation and %RSD of the five numbers in column B:

	A	B
1		1.211
2		1.435
3		1.303
4		1.456
5		1.843
6	Mean	1.4496
7	Standard deviation	0.241578
8	%RSD	16.66513

For simplicity, only rows 1-8 and columns A and B are shown. The commands entered in cells B6, B7 and B8 were:

=AVERAGE(B1:B5) in cell B6
=STDEV(B1:B5) in cell B7
=(B7/B6)*100 in cell B8

If the following commands are used:

=ROUND(AVERAGE(B1:B5), 3)
=ROUND(STDEV(B1:B5), 3)
=ROUND((B7/B6)*100, 3)

the answers are displayed to fewer significant places.

EXERCISE 1D

The concentration of chromium in a powder was measured six times. The results (in units of $\mu g\,dm^{-3}$) were: 1.2214, 1.2219, 1.2215, 1.2217, 1.2210 and 1.2208. Calculate the percent relevant standard deviation and the confidence limits of the mean of the concentration.

EXAMPLE 2

The concentration of magnesium in a tank of water was found using two different analytical methods, (i) atomic absorption spectroscopy (ASS) and (ii) by titration with EDTA. Twenty 10 cm³ samples were withdrawn from the tank, and five of these were used for five repeat measurements (after appropriate dilution) using AAS and the remaining five used for repeat measurements using the EDTA titration. The results were:

Sample	AAS	EDTA Titration
		$[Mg^{2+}]$/mg dm^{-3}
1	2.52	3.82
2	2.43	3.74
3	2.31	3.75
4	2.50	3.79
5	2.61	3.77

The true concentration of magnesium in the tank water is 2.50 $mg\,dm^{-3}$.

(i) Calculate the mean concentration determined by each method and the standard deviation and %RSD of the results for each method.

(ii) Comment upon the accuracy and precision of each method.

EXAMPLE 2 *(continued)*

Answer

(i) For AAS, $[Mg^{2+}]_{mean} = 2.47\ mg\ dm^{-3}, s = 0.11$ and $\%RDS = \frac{0.11}{2.47} = 4.5\%$.
For titration, $[Mg^{2+}]_{mean} = 3.78\ mg\ dm^{-3}, s = 0.032$ and $\%RDS = \frac{0.32}{3.78} = 0.85\%$.

(ii) The best way to assess the precision of the methods is to compare their %RSDs. The results from the titration have the lower %RSD, so they are more precise (less scattered) than those obtained by AAS. On the other hand, the AAS results are more accurate, since the mean AAS concentration ($2.47\ mg\ dm^{-3}$) is closer to the true concentration and the error is only 2.50 – 2.47 = 0.03 $mg\ dm^{-3}$. The poor accuracy of the results obtained by the EDTA titration suggests that the titration is subject to chemical interference. (For example, this might be because other metal ions present in the mixture are combining with the EDTA.)

General revision questions for Unit 1

1. The rate ('speed') of the reaction between bromopropane and sodium hydroxide is given by the equation
rate of reaction = $k\ [C_3H_7Br]\ [NaOH]$

where $[C_3H_7Br]$ is the concentration of bromopropane (in $mol\ dm^{-3}$), $[NaOH]$ is the concentration of sodium hydroxide (also in $mol\ dm^{-3}$) and k is a constant called the rate constant. The units of rate of reaction are $mol\ dm^{-3}\ s^{-1}$. What are the units of k?

2. The concentration of a hydrogen ion in a solution is 0.0545 $mol\ dm^{-3}$. Calculate the pH of the solution, expressing your answer to the correct number of significant figures.

3. A standard solution containing exactly 30.0 parts per million (ppm) of chromium(III) ions is sent to a laboratory for analysis. The lab reported that $[Cr^{3+}(aq)]_{mean} = 35.0 \pm 5\ ppm$.

(i) What is the uncertainty of the mean measurement?
(ii) What is the absolute error of the mean measurement?
(iii) What is the percentage relevant uncertainty of the measurement?

Revision questions for Extension 1

1. The ozone concentration in cm^3 of ozone per million cm^3 of air (i.e. in parts per million by volume, symbolised ppm(v)), was measured at a traffic junction. Repeat measurements over a five-minute period gave the following concentrations: 0.21, 0.22, 0.21, 0.23, 0.22, 0.22, 0.24, 0.19, 0.23 and 0.24. Calculate **(i)** the mean ozone concentration to two significant figures and the standard deviation to one significant figure, and **(ii)** the confidence limits of the mean concentration of ozone.

2. The concentration of benzene in a sealed flask was determined by two analytical methods, A and B. The results of four repeat measurements taken at the same time were:

Measurement	Method A	Method B
	Hydrocarbon concentration/ppb	
1	2.4	3.5
2	2.6	3.3
3	2.5	2.9
4	2.4	3.2

The true concentration of benzene in the flask is 3.100 ppb. Calculate the mean concentration of benzene determined by two methods, the percent relative deviation of both methods and the error of both methods. Comment on the accuracy and precision of both methods.

Answers

Exercises

Exercise 1A

Number of moles of C = $10^{80}/6.022 \times 10^{23} = 1.066 \times 10^{56}$

One mol of C has a mass of 12 g. Mass of Universe = $1.066 \times 10^{56} \times 12 = 1.99 \times 10^{57}$ g.

Exercise 1B

$c = 3 \times 10^8$ m s^{-1}. In 10^{-15} s, light travels $3 \times 10^8 \times 10^{-15} = 3 \times 10^{-7}$ m or 3×10^{-5} (= 0.00003) cm.

Exercise 1C

(i) Gradient is $-k$, the minus reflecting the fact that [A] falls with time, t.

(ii) Intercept is the value of the y-axis (here [A]) when the value of the x-axis (here, t) is zero. From the plot and also by inspection of equation (3), we see that at $t = 0$, $[A] = [A]_o$.

Exercise 1D

mean conc = 1.22138 $\mu g\ dm^{-3}$

Std dev = 0.00417 $\mu g\ dm^{-3}$

%RSD = $100 \times \frac{1.22138}{0.00417}$ = 0.0341%

Confidence limits = $1.22138 \pm \frac{2 \times 0.000417}{\sqrt{6}}$ = 1.22138 ± 0.00034.

Therefore, in the absence of systematic errors there is a 90% chance that the concentration of Cr lies between the limits 1.2217 – 1.2210 $\mu g\ dm^{-3}$

Revision questions for Unit 1

1. $k = \frac{rate\ of\ reaction}{[C_3H_7Br][NaOH]}$

So that the units of k = $\frac{mol\ dm^{-3} s^1}{(mol\ dm^{-3})^2} = mol^{-1} dm^{+3} s^{-1}$

2. pH = $-\log[H^+(aq)]$ = –log (0.545) = –(–) 0.2636 = 0.2636 = 0.264.

3. (i) uncertainty = ± 5ppm

(ii) Absolute error = 30.0 – 35.0 = -5 ppm

(iii) % s_r = $100 \times \frac{5}{35}$ = 14%.

Revision questions for Extension 1

1. Mean concentration = 0.22 ppm(v), s ≈ 0.02 ppm(v).

Confidence limits = $\bar{x} \pm 2 \times \frac{s}{\sqrt{n}}$ = 0.22 ± 0.013 ≈ 0.23 – 0.21 ppm(v).

2.

	A	B
Mean	2.475	3.225
STD Dev	0.0957	0.250
%RSD	3.89	7.75

The errors are: (3.100 – 2.475) = 0.625 ppb (for Method A) and: (3.225 – 3.100) = 0.125 ppb, so that Method B is more accurate than Method A. However, the %RSD of Method A is only 3.89%, so that it is the more precise of the two methods.