THE CHEMIST'S TOOLKITS

CONTENTS BY THEME

THE CHEMIST'S TOOLKIT 1 Quantities and units

The result of a measurement is a **physical quantity** that is reported as a numerical multiple of a unit:

physical quantity = numerical value × unit

It follows that units may be treated like algebraic quantities and may be multiplied, divided, and cancelled. Thus, the expression (physical quantity)/unit is the numerical value (a dimensionless quantity) of the measurement in the specified units. For instance, the mass m of an object could be reported as $m = 2.5$ kg or $m/$kg $= 2.5$. In this instance the unit of mass is 1 kg, but it is common to refer to the unit simply as kg (and likewise for other units). See Table A.1 in the *Resource section* of the text for a list of units.

Although it is good practice to use only SI units, there will be occasions where accepted practice is so deeply rooted that physical quantities are expressed using other, non-SI units. By international convention, all physical quantities are represented by oblique (sloping) letters (for instance, m for mass); units are given in roman (upright) letters (for instance m for metre).

Units may be modified by a prefix that denotes a factor of a power of 10. Among the most common SI prefixes are those listed in Table A.2 in the *Resource section*. Examples of the use of these prefixes are:

1 nm $= 10^{-9}$ m $\qquad$ 1 ps $= 10^{-12}$ s $\qquad$ 1 µmol $= 10^{-6}$ mol

Powers of units apply to the prefix as well as the unit they modify. For example, 1 cm^3 $= 1$ (cm)^3, and $(10^{-2}$ m)^3 $= 10^{-6}$ m^3. Note that 1 cm^3 does not mean 1 c(m^3). When carrying out numerical calculations, it is usually safest to write out the numerical value of an observable in scientific notation (as $n.nnn \times 10^n$).

There are seven SI base units, which are listed in Table A.3 in the *Resource section*. All other physical quantities may be expressed as combinations of these base units. *Molar concentration* (more formally, but very rarely, *amount of substance concentration*) for example, which is an amount of substance divided by the volume it occupies, can be expressed using the derived units of mol dm^{-3} as a combination of the base units for amount of substance and length. A number of these derived combinations of units have special names and symbols. For example, force is reported in the derived unit newton, 1 N $= 1$ kg m s^{-2} (see Table A.4 in the *Resource section* of the text).

THE CHEMIST'S TOOLKIT 2 Properties of bulk matter

The state of a bulk sample of matter is defined by specifying the values of various properties. Among them are:

> The **mass**, m, a measure of the quantity of matter present (unit: kilogram, kg).
>
> The **volume**, V, a measure of the quantity of space the sample occupies (unit: cubic metre, m^3).
>
> The **amount of substance**, n, a measure of the number of specified entities (atoms, molecules, or formula units) present (unit: mole, mol).

The amount of substance, n (colloquially, 'the number of moles'), is a measure of the number of specified entities present in the sample. 'Amount of substance' is the official name of the quantity; it is commonly simplified to 'chemical amount' or simply 'amount'. A mole is currently defined as the number of carbon atoms in exactly 12 g of carbon-12. (In 2011 the decision was taken to replace this definition, but the change has not yet, in 2018, been implemented.) The number of entities per mole is called **Avogadro's constant**, N_A; the currently accepted value is $6.022 \times 10^{23}\,mol^{-1}$ (note that N_A is a constant with units, not a pure number).

The **molar mass of a substance**, M (units: formally $kg\,mol^{-1}$ but commonly $g\,mol^{-1}$) is the mass per mole of its atoms, its molecules, or its formula units. The amount of substance of specified entities in a sample can readily be calculated from its mass, by noting that

$$n = \frac{m}{M} \qquad \text{Amount of substance} \quad (2.1)$$

A note on good practice Be careful to distinguish atomic or molecular mass (the mass of a single atom or molecule; unit: kg) from molar mass (the mass per mole of atoms or molecules; units: $kg\,mol^{-1}$). *Relative* molecular masses of atoms and molecules, $M_r = m/m_u$, where m is the mass of the atom or molecule and m_u is the atomic mass constant (see inside front cover of the text), are still widely called 'atomic weights' and 'molecular weights' even though they are dimensionless quantities and not weights ('weight' is the gravitational force exerted on an object).

A sample of matter may be subjected to a **pressure**, p (unit: pascal, Pa; $1\,Pa = 1\,kg\,m^{-1}\,s^{-2}$), which is defined as the force, F, it is subjected to, divided by the area, A, to which that force is applied. Although the pascal is the SI unit of pressure, it is also common to express pressure in bar ($1\,bar = 10^5\,Pa$) or atmospheres ($1\,atm = 101\,325\,Pa$ exactly), both of which correspond to typical atmospheric pressure. Because many physical properties depend on the pressure acting on a sample, it is appropriate to select a certain value of the pressure to report their values. The **standard pressure** for reporting physical quantities is currently defined as $p^{\ominus} = 1\,bar$ exactly.

To specify the state of a sample fully it is also necessary to give its **temperature**, T. The temperature is formally a property that determines in which direction energy will flow as heat when two samples are placed in contact through thermally conducting walls: energy flows from the sample with the higher temperature to the sample with the lower temperature. The symbol T is used to denote the **thermodynamic temperature** which is an absolute scale with $T = 0$ as the lowest point. Temperatures above $T = 0$ are then most commonly expressed by using the **Kelvin scale**, in which the gradations of temperature are expressed as multiples of the kelvin (K). The Kelvin scale is currently defined by setting the triple point of water (the temperature at which ice, liquid water, and water vapour are in mutual equilibrium) at exactly 273.16 K (as for certain other units, a decision has been taken to revise this definition, but it has not yet, in 2018, been implemented). The freezing point of water (the melting point of ice) at 1 atm is then found experimentally to lie 0.01 K below the triple point, so the freezing point of water is 273.15 K.

Suppose a sample is divided into smaller samples. If a property of the original sample has a value that is equal to the sum of its values in all the smaller samples (as mass would), then it is said to be **extensive**. Mass and volume are extensive properties. If a property retains the same value as in the original sample for all the smaller samples (as temperature would), then it is said to be **intensive**. In other words, the sum of the values of the smaller samples would not be equal to the value of the property of the initial sample. Temperature and pressure are intensive properties. Mass density, $\rho = m/V$, is also intensive because it would have the same value for all the smaller samples and the original sample. All molar properties, $X_m = X/n$, are intensive, whereas X and n are both extensive.

THE CHEMIST'S TOOLKIT 3 Momentum and force

The **speed**, v, of a body is defined as the rate of change of position. The **velocity**, $\boldsymbol{v}$, defines the direction of travel as well as the rate of motion, and particles travelling at the same speed but in different directions have different velocities. As shown in Sketch 3.1, the velocity can be depicted as an arrow in the direction of travel, its length being the speed v and its components v_x, v_y, and v_z along three perpendicular axes. These components have a sign: $v_x = +5\,\mathrm{m\,s^{-1}}$, for instance, indicates that a body is moving in the positive x-direction, whereas $v_x = -5\,\mathrm{m\,s^{-1}}$ indicates that it is moving in the opposite direction. The length of the arrow (the speed) is related to the components by Pythagoras' theorem: $v^2 = v_x^2 + v_y^2 + v_z^2$.
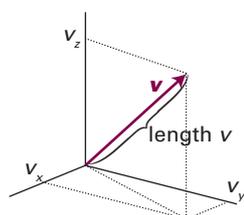
Sketch 3.1

The concepts of classical mechanics are commonly expressed in terms of the linear momentum, $\boldsymbol{p}$, which is defined as

$$\boldsymbol{p} = m\boldsymbol{v} \qquad \text{Linear momentum [definition]} \qquad (3.1)$$

Momentum also mirrors velocity in having a sense of direction; bodies of the same mass and moving at the same speed but in different directions have different linear momenta.

Acceleration, $\boldsymbol{a}$, is the rate of change of velocity. A body accelerates if its speed changes. A body also accelerates if its speed remains unchanged but its direction of motion changes. According to Newton's **second law of motion**, the acceleration of a body of mass m is proportional to the force, $\boldsymbol{F}$, acting on it:

$$\boldsymbol{F} = m\boldsymbol{a} \qquad \text{Force} \qquad (3.2)$$

Because $m\boldsymbol{v}$ is the linear momentum and $\boldsymbol{a}$ is the rate of change of velocity, $m\boldsymbol{a}$ is the rate of change of momentum. Therefore, an alternative statement of Newton's second law is that the force is equal to the rate of change of momentum. Newton's law indicates that the acceleration occurs in the same direction as the force acts. If, for an isolated system, no external force acts, then there is no acceleration. This statement is the **law of conservation of momentum**: that the momentum of a body is constant in the absence of a force acting on the body.

THE CHEMIST'S TOOLKIT 4 Integration

Integration is concerned with the areas under curves. The **integral** of a function $f(x)$, which is denoted $\int f(x)\mathrm{d}x$ (the symbol $\int$ is an elongated S denoting a sum), between the two values $x = a$ and $x = b$ is defined by imagining the x-axis as divided into strips of width δx and evaluating the following sum:

$$\int_a^b f(x)\mathrm{d}x = \lim_{\delta x \to 0} \sum_i f(x_i)\delta x \qquad \text{Integration [definition]} \quad (4.1)$$

As can be appreciated from Sketch 4.1, the integral is the area under the curve between the limits a and b. The function to be integrated is called the **integrand**. It is an astonishing mathematical fact that the integral of a function is the inverse of the differential of that function. In other words, if differentiation of f is followed by integration of the resulting function, the result is the original function f (to within a constant).

The integral in the preceding equation with the limits specified is called a **definite integral**. If it is written without the limits specified, it is called an **indefinite integral**. If the result of carrying out an indefinite integration is $g(x) + C$, where C is a constant, the following procedure is used to evaluate the corresponding definite integral:

$$I = \int_a^b f(x)\mathrm{d}x = \{g(x)+C\}\Big|_a^b = \{g(b)+C\}-\{g(a)+C\}$$

$$= g(b)-g(a) \qquad \text{Definite integral} \quad (4.2)$$

Note that the constant of integration disappears. The definite and indefinite integrals encountered in this text are listed in the *Resource section*. They may also be calculated by using mathematical software.
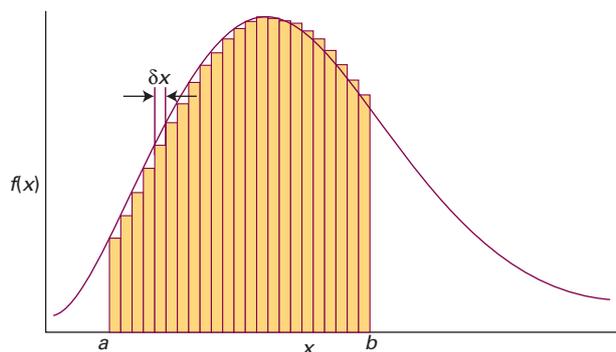
Sketch 4.1

Further information

When an indefinite integral is not in the form of one of those listed in the *Resource section* it is sometimes possible to transform it into one of these forms by using integration techniques such as:

Integration by parts. See *The chemist's toolkit* 15.

Substitution. Introduce a variable u related to the independent variable x (for example, an algebraic relation such as $u = x^2 - 1$ or a trigonometric relation such as $u = \sin x$). Express the differential $\mathrm{d}x$ in terms of $\mathrm{d}u$ (for these substitutions, $\mathrm{d}u = 2x\,\mathrm{d}x$ and $\mathrm{d}u = \cos x\,\mathrm{d}x$, respectively). Then transform the original integral written in terms of x into an integral in terms of u for which, in some cases, a standard form such as one of those listed in the *Resource section* can be used.

Brief illustration 4.1: Integration by substitution

To evaluate the indefinite integral $\int \cos^2 x \sin x\,\mathrm{d}x$ make the substitution $u = \cos x$. It follows that $\mathrm{d}u/\mathrm{d}x = -\sin x$, and therefore that $\sin x\,\mathrm{d}x = -\mathrm{d}u$. The integral is therefore

$$\int \cos^2 x \sin x\,\mathrm{d}x = -\int u^2\,\mathrm{d}u = -\tfrac{1}{3}u^3 + C = -\tfrac{1}{3}\cos^3 x + C$$

To evaluate the corresponding definite integral, convert the limits on x into limits on u. Thus, if the limits are $x = 0$ and $x = \pi$, the limits become $u = \cos 0 = 1$ and $u = \cos \pi = -1$:

$$\int_0^\pi \cos^2 x \sin x\,\mathrm{d}x = -\int_1^{-1} u^2\mathrm{d}u = \left\{-\tfrac{1}{3}u^3 + C\right\}\Big|_1^{-1} = \tfrac{2}{3}$$

A function may depend on more than one variable, in which case it may be necessary to integrate over all the variables, as in:

$$I = \int_a^b \int_c^d f(x,y)\mathrm{d}x\mathrm{d}y$$

We (but not everyone) adopt the convention that a and b are the limits of the variable x and c and d are the limits for y (as depicted by the colours in this instance). This procedure is simple if the function is a product of functions of each variable and of the form $f(x,y) = X(x)Y(y)$. In this case, the double integral is just a product of each integral:

$$I = \int_a^b \int_c^d X(x)Y(y)\mathrm{d}x\mathrm{d}y = \int_a^b X(x)\mathrm{d}x \int_c^d Y(y)\mathrm{d}y$$

Brief illustration 4.2: A double integral

Double integrals of the form

$$I = \int_0^{L_1} \int_0^{L_2} \sin^2(\pi x/L_1)\sin^2(\pi y/L_2)\mathrm{d}x\mathrm{d}y$$

occur in the discussion of the translational motion of a particle in two dimensions, where L_1 and L_2 are the maximum extents of travel along the x- and y-axes, respectively. To evaluate I write

$$I = \overbrace{\int_0^{L_1} \sin^2(\pi x/L_1)\mathrm{d}x}^{\text{Integral T.2}} \overbrace{\int_0^{L_2} \sin^2(\pi y/L_2)\mathrm{d}y}^{\text{Integral T.2}}$$

$$= \left\{\tfrac{1}{2}x - \frac{\sin(2\pi x/L_1)}{4\pi/L_1} + C\right\}\Big|_0^{L_1} \left\{\tfrac{1}{2}y - \frac{\sin(2\pi y/L_2)}{4\pi/L_2} + C\right\}\Big|_0^{L_2}$$

$$= \tfrac{1}{4}L_1 L_2$$

THE CHEMIST'S TOOLKIT 5 Differentiation

Differentiation is concerned with the slopes of functions, such as the rate of change of a variable with time. The formal definition of the **derivative**, df/dx, of a function $f(x)$ is

$$\frac{df}{dx}=\lim_{\delta x\to 0}\frac{f(x+\delta x)-f(x)}{\delta x} \qquad \text{First derivative [definition]} \qquad (5.1)$$

As shown in Sketch 5.1, the derivative can be interpreted as the slope of the tangent to the graph of $f(x)$ at a given value of x. A positive first derivative indicates that the function slopes upwards (as x increases), and a negative first derivative indicates the opposite. It is sometimes convenient to denote the first derivative as $f'(x)$. The **second derivative**, d^2f/dx^2, of a function is the derivative of the first derivative (here denoted f'):

$$\frac{d^2f}{dx^2}=\lim_{\delta x\to 0}\frac{f'(x+\delta x)-f'(x)}{\delta x} \qquad \text{Second derivative [definition]} \qquad (5.2)$$

It is sometimes convenient to denote the second derivative f''. As shown in Sketch 5.2, the second derivative of a function can be interpreted as an indication of the sharpness of the curvature of the function. A positive second derivative indicates that the function is $\cup$ shaped, and a negative second derivative indicates that it is $\cap$ shaped. The second derivative is zero at a **point of inflection**, where the first derivative passes through zero but does not change sign.

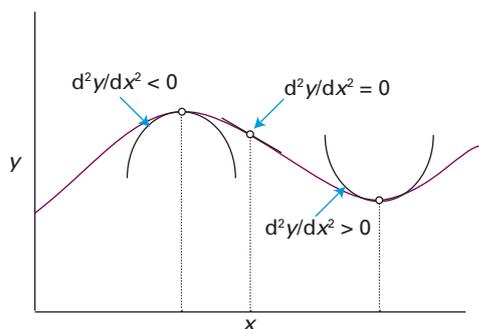The derivatives of some common functions are as follows:

$$\frac{d}{dx}x^n=nx^{n-1}$$

$$\frac{d}{dx}e^{ax}=ae^{ax}$$

$$\frac{d}{dx}\sin ax=a\cos ax \qquad \frac{d}{dx}\cos ax=-a\sin ax$$
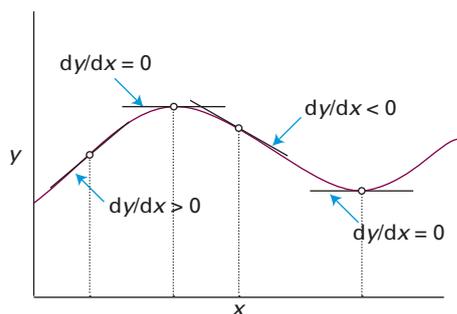
Sketch 5.1

Sketch 5.2

$$\frac{d}{dx}\ln ax=\frac{1}{x}$$

It follows from the definition of the derivative that a variety of combinations of functions can be differentiated by using the following rules:

$$\frac{d}{dx}(u+v)=\frac{du}{dx}+\frac{dv}{dx}$$

$$\frac{d}{dx}uv=u\frac{dv}{dx}+v\frac{du}{dx}$$

$$\frac{d}{dx}\frac{u}{v}=\frac{1}{v}\frac{du}{dx}-\frac{u}{v^2}\frac{dv}{dx}$$

Brief illustration 5.1: Derivatives of a product of functions

To differentiate the function $f=\sin^2 ax/x^2$ write

$$\frac{d}{dx}\frac{\sin^2 ax}{x^2}=\frac{d}{dx}\left(\frac{\sin ax}{x}\right)\left(\frac{\sin ax}{x}\right)=2\left(\frac{\sin ax}{x}\right)\frac{d}{dx}\left(\frac{\sin ax}{x}\right)$$

$$=2\left(\frac{\sin ax}{x}\right)\left\{\frac{1}{x}\frac{d}{dx}\sin ax+\sin ax\frac{d}{dx}\frac{1}{x}\right\}$$

$$=2\left\{\frac{a}{x^2}\sin ax\cos ax-\frac{\sin^2 ax}{x^3}\right\}$$

The function and this first derivative are plotted in Sketch 5.3.
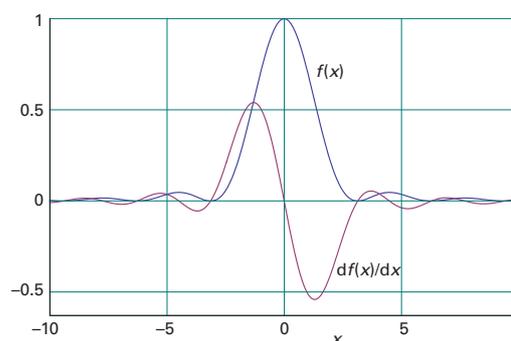
Sketch 5.3

It is sometimes convenient to differentiate with respect to a function of x, rather than x itself.

Brief illustration 5.2: Differentiation with respect to a function

Suppose that

$$f(x)=a+\frac{b}{x}+\frac{c}{x^2}$$

where a, b, and c are constants and you need to evaluate $df/d(1/x)$, rather than df/dx. To begin, let $y=1/x$. Then $f(y)=a+by+cy^2$ and

$$\frac{df}{dy}=b+2cy$$

Because $y=1/x$, it follows that

$$\frac{df}{d(1/x)}=b+\frac{2c}{x}$$

THE CHEMIST'S TOOLKIT 6 Work and energy

Work, w, is done when a body is moved against an opposing force. For an infinitesimal displacement through ds (a vector), the work done *on the body* is

$$dw_{body} = -F \cdot ds \qquad \text{Work done on body [definition]} \qquad (6.1)$$

where $F \cdot ds$ is the 'scalar product' of the vectors F and ds:

$$F \cdot ds = F_x dx + F_y dy + F_z dz \qquad \text{Scalar product [definition]} \qquad (6.2)$$

The energy lost as work by the system, dw, is the negative of the work done on the body, so

$$dw = F \cdot ds \qquad \text{Work done on system [definition]} \qquad (6.3)$$

For motion in one dimension, $dw = F_x dx$, with $F_x < 0$ (so $F_x = -|F_x|$) if it opposes the motion. The total work done along a path is the integral of this expression, allowing for the possibility that F changes in direction and magnitude at each point of the path. With force in newtons (N) and distance in metres, the units of work are joules (J), with

$$1\,J = 1\,N\,m = 1\,kg\,m^2\,s^{-2}$$

Energy is the capacity to do work. The SI unit of energy is the same as that of work, namely the joule. The rate of supply of energy is called the **power** (P), and is expressed in watts (W):

$$1\,W = 1\,J\,s^{-1}$$

A particle may possess two kinds of energy, kinetic energy and potential energy. The **kinetic energy**, E_k, of a body is the energy the body possesses as a result of its motion. For a body of mass m travelling at a speed v,

$$E_k = \tfrac{1}{2}mv^2 \qquad \text{Kinetic energy [definition]} \qquad (6.4)$$

Because $p = mv$ (*The chemist's toolkit* 3), where p is the magnitude of the linear momentum, it follows that

$$E_k = \frac{p^2}{2m} \qquad \text{Kinetic energy [definition]} \qquad (6.5)$$

The **potential energy**, E_p, (and commonly V, but do not confuse that with the volume!) of a body is the energy it possesses as a result of its position. In the absence of losses, the potential energy of a stationary particle is equal to the work that had to be done on the body to bring it to its current location. Because $dw_{body} = -F_x dx$, it follows that $dE_p = -F_x dx$ and therefore

$$F_x = -\frac{dE_p}{dx} \qquad \text{Potential energy [relation to force]} \qquad (6.6)$$

If E_p increases as x increases, then F_x is negative (directed towards negative x, Sketch 6.1). Thus, the steeper the gradient (the more strongly the potential energy depends on position), the greater is the force.
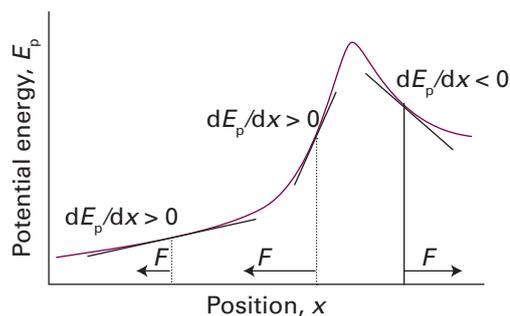
Sketch 6.1

No universal expression for the potential energy can be given because it depends on the type of force the body experiences. For a particle of mass m at an altitude h close to the surface of the Earth, the gravitational potential energy is

$$E_p(h) = E_p(0) + mgh \qquad \text{Gravitational potential energy} \qquad (6.7)$$

where g is the **acceleration of free fall** (g depends on location, but its 'standard value' is close to $9.81\,m\,s^{-2}$). The zero of potential energy is arbitrary. For a particle close to the surface of the Earth, it is common to set $E_p(0) = 0$.

The **Coulomb potential energy** of two electric charges, Q_1 and Q_2, separated by a distance r is

$$E_p = \frac{Q_1 Q_2}{4\pi\varepsilon r} \qquad \text{Coulomb potential energy} \qquad (6.8)$$

The quantity ε (epsilon) is the **permittivity**; its value depends upon the nature of the medium between the charges. If the charges are separated by a vacuum, then the constant is known as the **vacuum permittivity**, ε_0 (epsilon zero), or the **electric constant**, which has the value $8.854 \times 10^{-12}\,J^{-1}\,C^2\,m^{-1}$. The permittivity is greater for other media, such as air, water, or oil. It is commonly expressed as a multiple of the vacuum permittivity:

$$\varepsilon = \varepsilon_r \varepsilon_0 \qquad \text{Permittivity [definition]} \qquad (6.9)$$

with ε_r the dimensionless **relative permittivity** (formerly, the *dielectric constant*).

The **total energy** of a particle is the sum of its kinetic and potential energies:

$$E = E_k + E_p \qquad \text{Total energy [definition]} \qquad (6.10)$$

Provided no external forces are acting on the body, its total energy is constant. This central statement of physics is known as the **law of the conservation of energy**. Potential and kinetic energy may be freely interchanged, but their sum remains constant in the absence of external influences.

THE CHEMIST'S TOOLKIT 7 The equipartition theorem

The Boltzmann distribution (see the *Prologue* in the text) can be used to calculate the average energy associated with each mode of motion of an atom or molecule in a sample at a given temperature. However, when the temperature is so high that many energy levels are occupied, there is a much simpler way to find the average energy, through the **equipartition theorem**:

> For a sample at thermal equilibrium the average value of each quadratic contribution to the energy is $\frac{1}{2}kT$.

A 'quadratic contribution' is a term that is proportional to the square of the momentum (as in the expression for the kinetic energy, $E_k = p^2/2m$; *The chemist's toolkit* 6) or the displacement from an equilibrium position (as for the potential energy of a harmonic oscillator, $E_p = \frac{1}{2}k_f x^2$). The theorem is a conclusion from classical mechanics and for quantized systems is applicable only when the separation between the energy levels is so small compared to kT that many states are populated. Under normal conditions the equipartition theorem gives good estimates for the average energies associated with translation and rotation. However, the separation between vibrational and electronic states is typically much greater than for rotation or translation, and for these types of motion the equipartition theorem is unlikely to apply.

THE CHEMIST'S TOOLKIT 8 Electrical charge, current, power, and energy

Electrical charge, Q, is measured in *coulombs*, C. The elementary charge, e, the magnitude of charge carried by a single electron or proton, is approximately 1.6×10^{-19} C. The motion of charge gives rise to an **electric current**, I, measured in coulombs per second, or *amperes*, A, where $1\,A = 1\,C\,s^{-1}$. If the electric charge is that of electrons (as it is for the current in a metal), then a current of 1 A represents the flow of 6×10^{18} electrons ($10\,\mu mol\,e^-$) per second.

When a current I flows through a potential difference $\Delta\phi$ (measured in volts, V, with $1\,V = 1\,J\,C^{-1}$), the power, P, is

$$P = I\Delta\phi \tag{8.1}$$

It follows that if a constant current flows for a period t the energy supplied is

$$E = Pt = It\Delta\phi \tag{8.2}$$

Because $1\,A\,V\,s = 1\,(C\,s^{-1})\,V\,s = 1\,C\,V = 1\,J$, the energy is obtained in joules with the current in amperes, the potential difference in volts, and the time in seconds. That energy may be supplied as either work (to drive a motor) or as heat (through a 'heater'). In the latter case

$$q = It\Delta\phi \tag{8.3}$$

THE CHEMIST'S TOOLKIT 9 Partial derivatives

A **partial derivative** of a function of more than one variable, such as $f(x,y)$, is the slope of the function with respect to one of the variables, all the other variables being held constant (Sketch 9.1). Although a partial derivative shows how a function changes when one variable changes, it may be used to determine how the function changes when more than one variable changes by an infinitesimal amount. Thus, if f is a function of x and y, then when x and y change by dx and dy, respectively, f changes by

$$df = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy \tag{9.1}$$

where the symbol ∂ ('curly d') is used (instead of d) to denote a partial derivative and the subscript on the parentheses indicates which variable is being held constant.
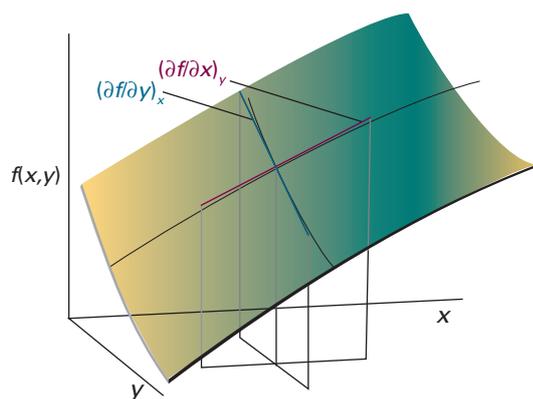
Sketch 9.1

The quantity df is also called the **differential** of f. Successive partial derivatives may be taken in any order:

$$\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)_y\right)_x = \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)_x\right)_y \tag{9.2}$$

Brief illustration 9.1: partial derivatives

Suppose that $f(x,y) = ax^3y + by^2$ (the function plotted in Sketch 9.1) then

$$\left(\frac{\partial f}{\partial x}\right)_y = 3ax^2y \qquad \left(\frac{\partial f}{\partial y}\right)_x = ax^3 + 2by$$

When x and y undergo infinitesimal changes, f changes by

$$df = 3ax^2y\,dx + (ax^3 + 2by)\,dy$$

To verify that the order of taking the second partial derivative is irrelevant, form

$$\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)_y\right)_x = \left(\frac{\partial(3ax^2y)}{\partial y}\right)_x = 3ax^2$$

$$\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)_x\right)_y = \left(\frac{\partial(ax^3+2by)}{\partial x}\right)_y = 3ax^2$$

Now suppose that z is a variable on which x and y depend (for example, x, y, and z might correspond to p, V, and T). The following relations then apply:

Relation 1. When x is changed at constant z:

$$\left(\frac{\partial f}{\partial x}\right)_z = \left(\frac{\partial f}{\partial x}\right)_y + \left(\frac{\partial f}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_z \tag{9.3}$$

Relation 2

$$\left(\frac{\partial y}{\partial x}\right)_z = \frac{1}{(\partial x/\partial y)_z} \tag{9.4}$$

Relation 3

$$\left(\frac{\partial x}{\partial y}\right)_z = -\left(\frac{\partial x}{\partial z}\right)_y\left(\frac{\partial z}{\partial y}\right)_x \tag{9.5}$$

Combining Relations 2 and 3 results in the **Euler chain relation**:

$$\left(\frac{\partial y}{\partial x}\right)_z\left(\frac{\partial x}{\partial z}\right)_y\left(\frac{\partial z}{\partial y}\right)_x = -1 \qquad \text{Euler chain relation} \tag{9.6}$$

THE CHEMIST'S TOOLKIT 10 Exact differentials

Suppose that $\mathrm{d}f$ can be expressed in the following way:

$$\mathrm{d}f = g(x,y)\mathrm{d}x + h(x,y)\mathrm{d}y \tag{10.1}$$

Is $\mathrm{d}f$ is an exact differential? If it is exact, then it can be expressed in the form

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial f}{\partial y}\right)_x \mathrm{d}y \tag{10.2}$$

Comparing these two expressions gives

$$\left(\frac{\partial f}{\partial x}\right)_y = g(x,y) \quad \left(\frac{\partial f}{\partial y}\right)_x = h(x,y) \tag{10.3}$$

It is a property of partial derivatives that successive derivatives may be taken in any order:

$$\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)_y\right)_x = \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)_x\right)_y \tag{10.4}$$

Taking the partial derivative with respect to x of the first equation, and with respect to y of the second gives

$$\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)_y\right)_x = \left(\frac{\partial g(x,y)}{\partial y}\right)_x$$
$$\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)_x\right)_y = \left(\frac{\partial h(x,y)}{\partial x}\right)_y \tag{10.5}$$

By the property of partial derivatives these two successive derivatives of f with respect to x and y must be the same, hence

$$\left(\frac{\partial g(x,y)}{\partial y}\right)_x = \left(\frac{\partial h(x,y)}{\partial x}\right)_y \tag{10.6}$$

If this equality is satisfied, then $\mathrm{d}f = g(x,y)\mathrm{d}x + h(x,y)\mathrm{d}y$ is an exact differential. Conversely, if it is known from other arguments that $\mathrm{d}f$ is exact, then this relation between the partial derivatives follows.

Brief illustration 10.1: Exact differentials

Suppose

$$\mathrm{d}f = \overbrace{3ax^2 y}^{g(x,y)}\mathrm{d}x + \overbrace{(ax^3 + 2by)}^{h(x,y)}\mathrm{d}y$$

To test whether $\mathrm{d}f$ is exact, form

$$\left(\frac{\partial g}{\partial f}\right)_x = \left(\frac{\partial(3ax^2 y)}{\partial y}\right)_x = 3ax^2$$

$$\left(\frac{\partial h}{\partial x}\right)_y = \left(\frac{\partial\left(ax^3 + 2by\right)}{\partial x}\right)_v = 3ax$$

The two second derivatives are the same, so $\mathrm{d}f$ is an exact differential and the function $f(x,y)$ can be constructed (see below).

Brief illustration 10.2: Inexact differentials

Suppose the following expression is encountered:

$$\mathrm{d}f = \overbrace{3ax^2 y}^{g(x,y)}\mathrm{d}x + \overbrace{(ax^2 + 2by)}^{h(x,y)}\mathrm{d}y$$

(Note the presence of ax^2 rather than the ax^3 in the preceding *Brief illustration*.) To test whether this is an exact differential, form

$$\left(\frac{\partial g}{\partial y}\right)_x = \left(\frac{\partial(3ax^2 y)}{\partial y}\right)_x = 3ax^2$$

$$\left(\frac{\partial h}{\partial x}\right)_y = \left(\frac{\partial(ax^2 + 2by)}{\partial x}\right)_y = 2ax$$

These two expressions are not equal, so this form of $\mathrm{d}f$ is not an exact differential and there is not a corresponding integrated function of the form $f(x,y)$.

Further information

If $\mathrm{d}f$ is exact, then

- From a knowledge of the functions g and h the function f can be constructed.
- It then follows that the integral of $\mathrm{d}f$ between specified limits is independent of the path between those limits.

The first conclusion is best demonstrated with a specific example.

Brief illustration 10.3: The reconstruction of an equation

Consider the differential $\mathrm{d}f = 3ax^2 y\mathrm{d}x + (ax^3 + 2by)\mathrm{d}y$, which is known to be exact. Because $(\partial f/\partial x)_y = 3ax^2 y$, it can be integrated with respect to x with y held constant, to obtain

$$f = \int \mathrm{d}f = \int 3ax^2 y\,\mathrm{d}x = 3ay\int x^2 \mathrm{d}x = ax^3 y + k$$

where the 'constant' of integration k may depend on y (which has been treated as a constant in the integration), but not on x. To find $k(y)$, note that $(\partial f/\partial y)_x = ax^3 + 2by$, and therefore

$$\left(\frac{\partial f}{\partial y}\right)_x = \left(\frac{\partial(ax^3 y + k)}{\partial y}\right)_x = ax^3 + \frac{\mathrm{d}k}{\mathrm{d}y} = ax^3 + 2by$$

Therefore

$$\frac{\mathrm{d}k}{\mathrm{d}y} = 2by$$

from which it follows that $k = by^2 +$ constant. It follows that

$$f(x,y) = ax^3 y + by^2 + \text{constant}$$

The value of the constant is pinned down by stating the boundary conditions; thus, if it is known that $f(0,0) = 0$, then the constant is zero.

To demonstrate that the integral of df is independent of the path is now straightforward. Because df is a differential, its integral between the limits a and b is

$$\int_a^b df = f(b) - f(a) \qquad (10.7)$$

The value of the integral depends only on the values at the end points and is independent of the path between them. If df is not an exact differential, the function f does not exist, and this argument no longer holds. In such cases, the integral of df does depend on the path.

Brief illustration 10.4: Path-dependent integration

Consider the inexact differential (the expression with ax^2 in place of ax^3 inside the second parentheses):

$$df = 3ax^2y\,dx + (ax^2 + 2by)dy$$

Suppose df is integrated from (0,0) to (2,2) along the two paths shown in Sketch 10.1.

(0,2)   y = 2   (2,2)

x = 0

x = 2

y

Path 2   Path 1

y = 0   (2,0)

x
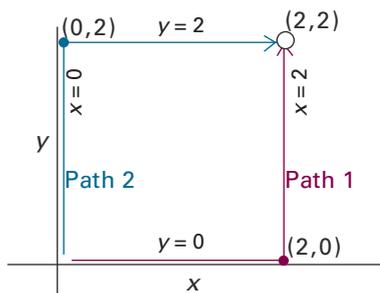
Sketch 10.1

Along Path 1,

$$\int_{\text{Path 1}} df = \int_{0,0}^{2,0} 3ax^2y\,dx + \int_{2,0}^{2,2}(ax^2 + 2by)dy$$

$$= 0 + 4a\int_0^2 dy + 2b\int_0^2 y\,dy = 8a + 4b$$

whereas along Path 2,

$$\int_{\text{Path 2}} df = \int_{0,2}^{2,2} 3ax^2y\,dx + \int_{0,0}^{0,2}(ax^2 + 2by)dy$$

$$= 6a\int_0^2 x^2\,dx + 0 + 2b\int_0^2 y\,dy = 16a + 4b$$

The two integrals are not the same.

An inexact differential may sometimes be converted into an exact differential by multiplication by a factor known as an *integrating factor*. A physical example is the integrating factor $1/T$ that converts the inexact differential dq_{rev} into the exact differential dS in thermodynamics (Topic 3B of the text).

Brief illustration 10.5: An integrating factor

The differential $df = 3ax^2y\,dx + (ax^2 + 2by)dy$ is inexact; the same is true when $b = 0$ and so for simplicity consider $df = 3ax^2y\,dx + ax^2\,dy$ instead. Multiplication of this df by x^my^n and writing $x^my^n\,df = df'$ gives

$$df' = \overbrace{3ax^{m+2}y^{n+1}}^{g(x,y)}\,dx + \overbrace{ax^{m+2}y^n}^{h(x,y)}\,dy$$

Now

$$\left(\frac{\partial g}{\partial y}\right)_x = \left(\frac{\partial(3ax^{m+2}y^{n+1})}{\partial y}\right)_x = 3a(n+1)x^{m+2}y^n$$

$$\left(\frac{\partial h}{\partial x}\right)_y = \left(\frac{\partial(ax^{m+2}y^n)}{\partial x}\right)_y = a(m+2)x^{m+1}y^n$$

For the new differential to be exact, these two partial derivatives must be equal, so write

$$3a(n + 1)x^{m+2}y^n = a(m + 2)x^{m+1}y^n$$

which simplifies to

$$3(n + 1)x = m + 2$$

The only solution that is independent of x is $n = -1$ and $m = -2$. It follows that

$$df' = 3a\,dx + (a/y)dy$$

is an exact differential. By the procedure already illustrated, its integrated form is $f'(x,y) = 3ax + a\ln y + \text{constant}$.

THE CHEMIST'S TOOLKIT 11 Measures of concentration

Let A be the solvent and B the solute. The **molar concentration** (informally: 'molarity'), c_B or [B], is the amount of solute molecules (in moles) divided by the volume, V, of the solution:

$$c_B = \frac{n_B}{V} \tag{11.1}$$

It is commonly reported in moles per cubic decimetre (mol dm^{-3}) or, equivalently, in moles per litre (mol L^{-1}). It is convenient to define its 'standard' value as $c^\ominus = 1 \text{ mol dm}^{-3}$.

The **molality**, b_B, of a solute is the amount of solute species (in moles) in a solution divided by the total mass of the solvent (in kilograms), m_A:

$$b_B = \frac{n_B}{m_A} \tag{11.2}$$

Both the molality and mole fraction are independent of temperature; in contrast, the molar concentration is not. It is convenient to define the 'standard' value of the molality as $b^\ominus = 1 \text{ mol kg}^{-1}$.

1. The relation between molality and mole fraction

Consider a solution with one solute and having a total amount n of molecules. If the mole fraction of the solute is x_B, the amount of solute molecules is $n_B = x_B n$. The mole fraction of solvent molecules is $x_A = 1 - x_B$, so the amount of solvent molecules is $n_A = x_A n = (1 - x_B)n$. The mass of solvent, of molar mass M_A, present is $m_A = n_A M_A = (1-x_B)n M_A$. The molality of the solute is therefore

$$b_B = \frac{n_B}{m_A} = \frac{x_B n}{(1-x_B)n M_A} = \frac{x_B}{(1-x_B)M_A} \tag{11.3a}$$

The inverse of this relation, the mole fraction in terms of the molality, is

$$x_B = \frac{b_B M_A}{1 + b_B M_A} \tag{11.3b}$$

2. The relation between molality and molar concentration

The total mass of a volume V of *solution* (not solvent) of mass density ρ is $m = \rho V$. The amount of solute molecules in this volume is $n_B = c_B V$. The mass of solute present is $m_B = n_B M_B = c_B V M_B$. The mass of solvent present is therefore $m_A = m - m_B = \rho V - c_B V M_B = (\rho - c_B M_B)V$. The molality is therefore

$$b_B = \frac{n_B}{m_A} = \frac{c_B V}{(\rho - c_B M_B)V} = \frac{c_B}{\rho - c_B M_B} \tag{11.4a}$$

The inverse of this relation, the molar concentration in terms of the molality, is

$$c_B = \frac{b_B \rho}{1 + b_B M_B} \tag{11.4b}$$

3. The relation between molar concentration and mole fraction

By inserting the expression for b_B in terms of x_B into the expression for c_B, the molar concentration of B in terms of its mole fraction is

$$c_B = \frac{x_B \rho}{x_A M_A + x_B M_B} \tag{11.5}$$

with $x_A = 1 - x_B$. For a dilute solution in the sense that $x_B M_B \ll x_A M_A$,

$$c_B \approx \left(\frac{\rho}{x_A M_A}\right) x_B \tag{11.6}$$

If, moreover, $x_B \ll 1$, so $x_A \approx 1$, then

$$c_B \approx \left(\frac{\rho}{M_A}\right) x_B \tag{11.7}$$

THE CHEMIST'S TOOLKIT 12 Series expansions

A function $f(x)$ can be expressed in terms of its value in the vicinity of $x = a$ by using the **Taylor series**

$$f(x) = f(a) + \left(\frac{df}{dx}\right)_a (x-a) + \frac{1}{2!}\left(\frac{d^2 f}{dx^2}\right)_a (x-a)^2 + \cdots$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n f}{dx^n}\right)_a (x-a)^n \qquad \text{Taylor series} \quad (12.1)$$

where the notation $(\ldots)_a$ means that the derivative is evaluated at $x = a$ and $n!$ denotes a **factorial** defined as

$$n! = n(n-1)(n-2)\ldots 1, \quad 0! \equiv 1 \qquad \text{Factorial} \quad (12.2)$$

The **Maclaurin series** for a function is a special case of the Taylor series in which $a = 0$. The following Maclaurin series are used at various stages in the text:

$$(1+x)^{-1} = 1 - x + x^2 - \cdots = \sum_{n=0}^{\infty} (-1)^n x^n \qquad (12.3a)$$

$$e^x = 1 + x + \tfrac{1}{2}x^2 + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad (12.3b)$$

$$\ln(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \cdots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \qquad (12.3c)$$

Series expansions are used to simplify calculations, because when $|x| \ll 1$ it is possible, to a good approximation, to terminate the series after one or two terms. Thus, provided $|x| \ll 1$,

$$(1+x)^{-1} \approx 1 - x \qquad (12.4a)$$

$$e^x \approx 1 + x \qquad (12.4b)$$

$$\ln(1+x) \approx x \qquad (12.4c)$$

A series is said to **converge** if the sum approaches a finite, definite value as n approaches infinity. If the sum does not approach a finite, definite value, then the series is said to **diverge**. Thus, the series expansion of $(1+x)^{-1}$ converges for $|x| < 1$ and diverges for $|x| \geq 1$. Tests for convergence are explained in mathematical texts.

THE CHEMIST'S TOOLKIT 13 Electromagnetic radiation

Electromagnetic radiation consists of oscillating electric and magnetic disturbances that propagate as waves. The two components of an electromagnetic wave are mutually perpendicular and are also perpendicular to the direction of propagation (Sketch 13.1). Electromagnetic waves travel through a vacuum at a constant speed called the **speed of light**, c, which has the defined value of exactly $2.99792458 \times 10^8\,\mathrm{m\,s^{-1}}$.
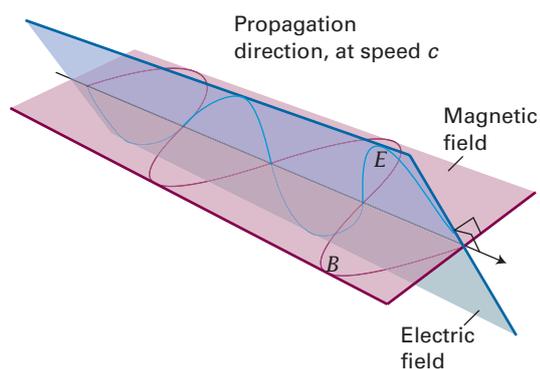
Sketch 13.1

A wave is characterized by its **wavelength**, λ (lambda), the distance between consecutive peaks of the wave (Sketch 13.2). The classification of electromagnetic radiation according to its wavelength is shown in Sketch 13.3. Light, which is electromagnetic radiation that is visible to the human eye, has a wavelength in the range 420 nm (violet light) to 700 nm (red light). The properties of a wave may also be expressed in terms of its **frequency**, ν (nu), the number of oscillations in a time interval divided by the duration of the interval. Frequency is reported in hertz, Hz, with $1\,\mathrm{Hz} = 1\,\mathrm{s^{-1}}$ (i.e. 1 cycle per second). Light spans the frequency range from 710 THz (violet light) to 430 THz (red light).
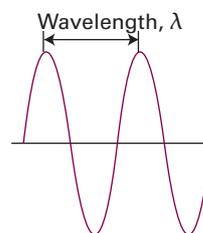
Sketch 13.2

The wavelength and frequency of an electromagnetic wave are related by:

$$c = \lambda \nu \qquad \text{(13.1)}$$

The relation between wavelength and frequency in a vacuum

It is also common to describe a wave in terms of its **wavenumber**, $\tilde{\nu}$ (nu tilde), which is defined as

$$\tilde{\nu} = \frac{1}{\lambda} \text{ or equivalently } \tilde{\nu} = \frac{\nu}{c} \qquad \text{(13.2)}$$

Wavenumber [definition]

Thus, wavenumber is the reciprocal of the wavelength and can be interpreted as the number of wavelengths in a given distance. In spectroscopy, for historical reasons, wavenumber is usually reported in units of reciprocal centimetres ($\mathrm{cm^{-1}}$). Visible light therefore corresponds to electromagnetic radiation with a wavenumber of $14\,000\,\mathrm{cm^{-1}}$ (red light) to $24\,000\,\mathrm{cm^{-1}}$ (violet light).

Electromagnetic radiation that consists of a single frequency (and therefore a single wavelength) is **monochromatic**, because it corresponds to a single colour. *White light* consists of electromagnetic waves with a continuous, but not uniform, spread of frequencies throughout the visible region of the spectrum.
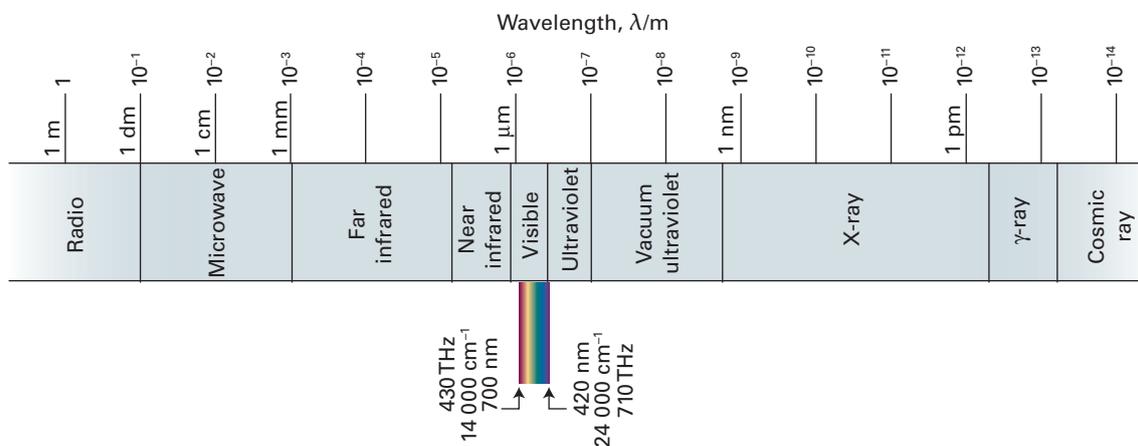
Sketch 13.3

A characteristic property of waves is that they interfere with one another, which means that they result in a greater amplitude where their displacements add and a smaller amplitude where their displacements subtract (Sketch 13.4). The former is called 'constructive interference' and the latter 'destructive interference'. The regions of constructive and destructive interference show up as regions of enhanced and diminished intensity. The phenomenon of **diffraction** is the interference caused by an object in the path of waves and occurs when the dimensions of the object are comparable to the wavelength of the radiation. Light waves, with wavelengths of the order of 500 nm, are diffracted by narrow slits.

Constructive interference

Destructive interference
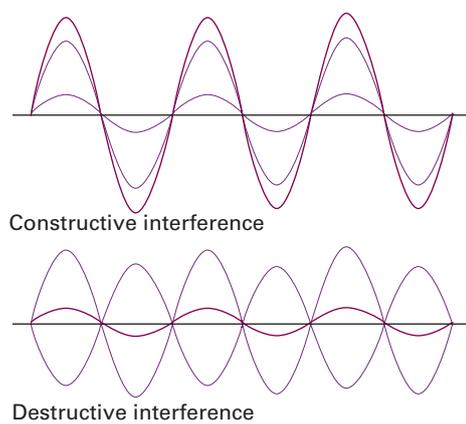
Sketch 13.4

THE CHEMIST'S TOOLKIT 14 Complex numbers

Complex numbers have the general form

$$z = x + iy \tag{14.1a}$$

where $i = \sqrt{-1}$. The real number x is the 'real part of z', denoted $\mathrm{Re}(z)$; likewise, the real number y is 'the imaginary part of z', $\mathrm{Im}(z)$. The **complex conjugate** of z, denoted z^*, is formed by replacing i by $-i$:

$$z^* = x - iy \tag{14.1b}$$

Brief illustration 14.1: Operations with complex numbers

Consider the complex numbers $z_1 = 6 + 2i$ and $z_2 = -4 - 3i$. Then

$$z_1 + z_2 = (6 - 4) + (2 - 3)i = 2 - i$$
$$z_1 - z_2 = 10 + 5i$$
$$z_1 z_2 = \{6(-4) - 2(-3)\} + \{6(-3) + 2(-4)\}i = -18 - 26i$$

The product of z^* and z is denoted $|z|^2$ and is called the **square modulus** of z. From the definition of z and z^* and $i^2 = -1$ it follows that

$$|z|^2 = z^*z = (x + iy)(x - iy) = x^2 + y^2 \tag{14.2}$$

The square modulus is a real, non-negative number. The **absolute value** or **modulus** is denoted $|z|$ and is given by:

$$|z| = (z^*z)^{1/2} = (x^2 + y^2)^{1/2} \tag{14.3}$$

The inverse of z, denoted z^{-1}, is such that $zz^{-1} = 1$, which is satisfied if

$$z^{-1} = \frac{z^*}{|z|^2} \tag{14.4}$$

This construction is used in the division of complex numbers:

$$\frac{z_1}{z_2} = z_1 z_2^{-1} \tag{14.5}$$

Brief illustration 14.2: Inverse

Consider the complex number $z = 8 - 3i$. Its square modulus is

$$|z|^2 = z^*z = (8 - 3i)^*(8 - 3i) = (8 + 3i)(8 - 3i) = 64 + 9 = 73$$

The modulus is therefore $|z| = 73^{1/2}$. The inverse of z is

$$z^{-1} = \frac{8 + 3i}{73} = \frac{8}{73} + \frac{3}{73}i$$

Then

$$\frac{6 + 2i}{8 - 3i} = (6 + 2i)(8 - 3i)^{-1} = (6 + 2i)\left(\frac{8}{73} + \frac{3}{73}i\right) = -\frac{42}{73} + \frac{34}{73}i$$

For further information about complex numbers, see *The chemist's toolkit* 16 .

THE CHEMIST'S TOOLKIT 15 Integration by parts

Many integrals in quantum mechanics have the form $\int f(x)h(x)dx$, where $f(x)$ and $h(x)$ are two different functions. Such integrals can often be evaluated by regarding $h(x)$ as the derivative of another function, $g(x)$, such that $h(x) = dg(x)/dx$. For instance, if $h(x) = x$, then $g(x) = \frac{1}{2}x^2$. The integral is then found using **integration by parts**:

$$\int f \frac{dg}{dx}dx = fg - \int g \frac{df}{dx}dx \qquad (15.1a)$$

The procedure is successful only if the integral on the right turns out to be one that can be evaluated more easily than the one on the left. The procedure is often summarized by expressing this relation as

$$\int f\,dg = fg - \int g\,df \qquad (15.1b)$$

Brief illustration 15.1: Integration by parts

Integrals over xe^{-ax} and their analogues occur commonly in the discussion of atomic structure and spectra. They may be integrated by parts, as in the following. Consider integration of xe^{-ax}. In this case, $f(x) = x$, so $df(x)/dx = 1$ and $dg(x)/dx = e^{-ax}$, so $g(x) = -(1/a)e^{-ax}$. Then

$$\int \overset{f}{x}\,\overset{dg/dx}{e^{-ax}}\,dx = \overset{f}{x}\,\overset{g}{\frac{-e^{-ax}}{a}} - \int \overset{g}{\frac{-e^{-ax}}{a}}\,\overset{df/dx}{1}\,dx$$

$$= -\frac{xe^{-ax}}{a} + \frac{1}{a}\int e^{-ax}\,dx = -\frac{xe^{-ax}}{a} - \frac{e^{-ax}}{a^2} + \text{constant}$$

If the integral is definite, then apply the limits to the final step above and write

$$\int_0^\infty \overset{f}{x}\,\overset{dg/dx}{e^{-ax}}\,dx = \overset{f}{x}\,\overset{g}{\frac{-e^{-ax}}{a}}\Big|_0^\infty - \int_0^\infty \overset{g}{\frac{-e^{-ax}}{a}}\,\overset{df/dx}{1}\,dx$$

$$= -\frac{xe^{-ax}}{a}\Big|_0^\infty + \frac{1}{a}\int_0^\infty e^{-ax}\,dx = 0 - \frac{e^{-ax}}{a^2}\Big|_0^\infty = \frac{1}{a^2}$$

A complex number $z = x + iy$ can be represented as a point in a plane, the **complex plane**, with $\text{Re}(z)$ along the x-axis and $\text{Im}(z)$ along the y-axis (Sketch 16.1). The position of the point can also be specified in terms of a distance r and an angle ϕ (the polar coordinates). Then $x = r\cos\phi$ and $y = r\sin\phi$, so it follows that

$$z = r(\cos\phi + i\sin\phi) \tag{16.1}$$

The angle ϕ, called the **argument** of z, is the angle that r makes with the x-axis. Because $y/x = \tan\phi$, it follows that

$$r = (x^2 + y^2)^{1/2} = |z| \qquad \phi = \arctan\frac{y}{x} \tag{16.2}$$
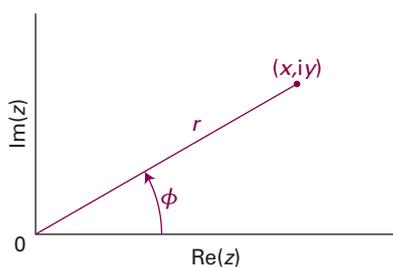
Sketch 16.1

One of the most useful relations involving complex numbers is **Euler's formula**:

$$e^{i\phi} = \cos\phi + i\sin\phi \tag{16.3}$$

from which it follows that $z = r(\cos\phi + i\sin\phi)$ can be written

$$z = re^{i\phi} \tag{16.4}$$
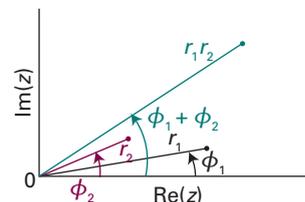
Two more useful relations arise by noting that $e^{-i\phi} = \cos(-\phi) + i\sin(-\phi) = \cos\phi - i\sin\phi$; it then follows that

$$\cos\phi = \tfrac{1}{2}(e^{i\phi} + e^{-i\phi}) \qquad \sin\phi = -\tfrac{1}{2}i(e^{i\phi} - e^{-i\phi}) \tag{16.5}$$

The polar form of a complex number is commonly used to perform arithmetical operations. For instance, the product of two complex numbers in polar form is

$$z_1 z_2 = (r_1 e^{i\phi_1})(r_2 e^{i\phi_2}) = r_1 r_2 e^{i(\phi_1 + \phi_2)} \tag{16.6}$$

This construction is illustrated in Sketch 16.2.

Sketch 16.2

Brief illustration 16.1: Polar representation

Consider the complex number $z = 8 - 3i$. From *Brief illustration* 14.1, $r = |z| = 73^{1/2}$. The argument of z is

$$\phi = \arctan\left(\frac{-3}{8}\right) = -0.359 \text{ rad, or } -20.6°$$

The polar form of the number is therefore

$$z = 73^{1/2}\,e^{-0.359i}$$

Brief illustration 16.2: Roots

To determine the 5th root of $z = 8 - 3i$, note that from *Brief illustration* 16.1 its polar form is

$$z = 73^{1/2}\,e^{-0.359i} = 8.544e^{-0.359i}$$

The 5th root is therefore

$$z^{1/5} = (8.544e^{-0.359i})^{1/5} = 8.544^{1/5}\,e^{-0.359i/5} = 1.536e^{-0.0718i}$$

It follows that $x = 1.536\cos(-0.0718) = 1.532$ and $y = 1.536\sin(-0.0718) = -0.110$ (note that the ϕ are in radians), so

$$(8 - 3i)^{1/5} = 1.532 - 0.110i$$

THE CHEMIST'S TOOLKIT 17 Vectors

A vector is a quantity with both magnitude and direction. The vector $\boldsymbol{v}$ shown in Sketch 17.1 has components on the x, y, and z axes with values v_x, v_y, and v_z, respectively, which may be positive or negative. For example, if $v_x = -1.0$, the x-component of the vector $\boldsymbol{v}$ has a magnitude of 1.0 and points in the $-x$ direction. The magnitude of a vector is denoted v or $|\boldsymbol{v}|$ and is given by

$$v = (v_x^2 + v_y^2 + v_z^2)^{1/2} \tag{17.1}$$

Thus, a vector with components $v_x = -1.0$, $v_y = +2.5$, and $v_z = +1.1$ has magnitude 2.9 and would be represented by an arrow of length 2.9 units and the appropriate orientation (as in the inset in the Sketch). Velocity and momentum are vectors; the magnitude of a velocity vector is called the speed. Force, too, is a vector. Electric and magnetic fields are two more examples of vectors.
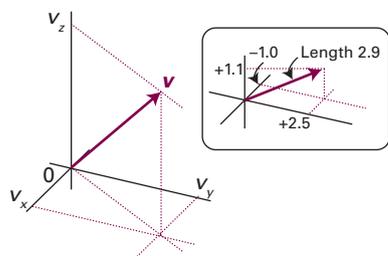
Sketch 17.1

Further information

If the polar coordinates of the vector $\boldsymbol{v}$ are θ and ϕ (the colatitude and azimuth, respectively), then

$$\begin{aligned} v_x &= v \sin\theta \cos\phi \\ v_y &= v \sin\theta \sin\phi \\ v_z &= v \cos\theta \end{aligned} \qquad \text{Orientation} \tag{17.2}$$

and therefore that

$$\theta = \arccos(v_z/v) \qquad \phi = \arctan(v_y/v_x) \tag{17.3}$$

Brief illustration 17.1: Vector orientation

The vector $\boldsymbol{v} = 2\boldsymbol{i} + 3\boldsymbol{j} - \boldsymbol{k}$ has magnitude

$$v = \{2^2 + 3^2 + (-1)^2\}^{1/2} = 14^{1/2} = 3.74$$

Its direction is given by

$$\theta = \arccos(-1/14^{1/2}) = 105.5° \qquad \phi = \arctan(3/2) = 56.3°$$

The operations involving vectors (addition, multiplication, etc.) needed for the text are described in *The chemist's toolkit* 22.

THE CHEMIST'S TOOLKIT 18 The classical harmonic oscillator

A **harmonic oscillator** consists of a particle of mass m that experiences a 'Hooke's law' restoring force, one that is proportional to the displacement of the particle from equilibrium. An example is a particle of mass m attached to a spring or an atom attached to another by a chemical bond. For a one-dimensional system,

$$F_x = -k_f x \tag{18.1}$$

where the constant of proportionality is called the **force constant**. From Newton's second law of motion ($F = ma = m(\mathrm{d}^2 x/\mathrm{d}t^2)$; see *The chemist's toolkit* 3),

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -k_f x \tag{18.2}$$

If $x = 0$ at $t = 0$, a solution (as may be verified by substitution) is

$$x(t) = A \sin 2\pi\nu t \qquad \nu = \frac{1}{2\pi}\left(\frac{k_f}{m}\right)^{1/2} \tag{18.3}$$

This solution shows that the position of the particle oscillates *harmonically* (i.e. as a sine function) with frequency ν (units: Hz) and that the frequency of oscillation is high for light particles (m small) attached to stiff springs (k_f large). It is useful to define the *angular frequency* as $\omega = 2\pi\nu$ (units: radians per second). It follows that the angular frequency of a classical harmonic oscillator is $\omega = (k_f/m)^{1/2}$.

The negative sign in the expression for the force implies that it is negative (directed toward negative x) if the displacement is positive, and vice versa. Potential energy V is related to force by $F = -\mathrm{d}V/\mathrm{d}x$ (*The chemist's toolkit* 6), so the potential energy corresponding to a Hooke's law restoring force is

$$V(x) = \tfrac{1}{2}k_f x^2 \tag{18.4}$$

Such a potential energy is called a 'harmonic potential energy' or a 'parabolic potential energy'.

As the particle moves away from the equilibrium position its potential energy increases and so its kinetic energy, and hence its speed, decreases. At some point all the energy is potential and the particle comes to rest at a turning point. The particle then accelerates back towards and through the equilibrium position. The greatest probability of finding the particle is where it is moving most slowly, which is close to the turning points.

The turning point, x_{tp}, of a classical oscillator occurs when its potential energy $\tfrac{1}{2}k_f x^2$ is equal to its total energy, so

$$x_{tp} = \pm\left(\frac{2E}{k_f}\right)^{1/2}$$

The turning point increases with the total energy: in classical terms, the amplitude of the swing of a pendulum or the displacement of a mass on a spring increases.

THE CHEMIST'S TOOLKIT 19 Cylindrical coordinates

For systems with cylindrical symmetry it is best to work in **cylindrical coordinates** r, ϕ, and z (Sketch 19.1), with

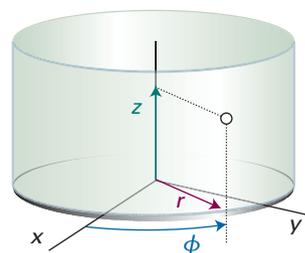$$x = r\cos\phi \qquad y = r\sin\phi \qquad (19.1)$$

and where

$$0 \le r \le \infty \qquad 0 \le \phi \le 2\pi \qquad -\infty \le z \le +\infty \qquad (19.2)$$

The volume element is

$$d\tau = r\,dr\,d\phi\,dz \qquad (19.3a)$$

For motion in a plane, $z = 0$ and the volume element is

$$d\tau = r\,dr\,d\phi \qquad (19.3b)$$

Sketch 19.1

THE CHEMIST'S TOOLKIT 20 Angular momentum

Angular velocity, ω (omega), is the rate of change of angular position; it is reported in radians per second (rad s^{-1}). There are 2π radians in a circle, so 1 cycle per second is the same as 2π radians per second. For convenience, the 'rad' is often dropped, and the units of angular velocity are denoted s^{-1}.

Expressions for other angular properties follow by analogy with the corresponding equations for linear motion (*The chemist's toolkit* 3). Thus, the magnitude, J, of the **angular momentum**, **J**, is defined, by analogy with the magnitude of the linear momentum ($p = mv$):

$$J = I\omega \tag{20.1}$$

The quantity I is the **moment of inertia** of the object. It represents the resistance of the object to a change in the state of rotation in the same way that mass represents the resistance of the object to a change in the state of translation. In the case of a rotating molecule the moment of inertia is defined as

$$I = \sum_i m_i r_i^2 \tag{20.2}$$

where m_i is the mass of atom i and r_i is its perpendicular distance from the axis of rotation (Sketch 20.1). For a point particle of mass m moving in a circle of radius r, the moment of inertia about the axis of rotation is

$$I = mr^2 \tag{20.3}$$

The SI units of moment of inertia are therefore kilogram metre^2 (kg m^2), and those of angular momentum are kilogram metre^2 per second (kg m^2 s^{-1}).

$$m_A \quad r_A = 0$$
$$I = m_B r_B^2 + m_C r_C^2 + m_D r_D^2$$
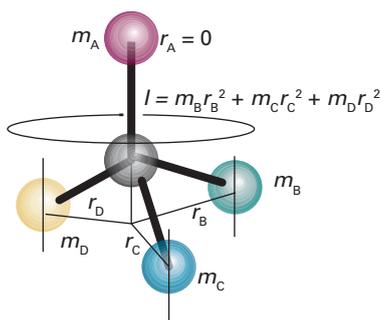
Sketch 20.1

The angular momentum is a vector, a quantity with both magnitude and direction (*The chemist's toolkit* 17). For rotation in three dimensions, the angular momentum has three components: J_x, J_y, and J_z. For a particle travelling on a circular path of radius r about the z-axis, and therefore confined to the xy-plane, the angular momentum vector points in the z-direction only (Sketch 20.2), and its only component is

$$J_z = \pm pr \tag{20.4}$$

where p is the magnitude of the linear momentum in the xy-plane at any instant. When $J_z > 0$, the particle travels in a clockwise direction as viewed from below; when $J_z < 0$, the motion is anticlockwise. A particle that is travelling at high speed in a circle has a higher angular momentum than a particle of the same mass travelling more slowly. An object with a high angular momentum (like a flywheel) requires a strong braking force (more precisely, a strong 'torque') to bring it to a standstill.
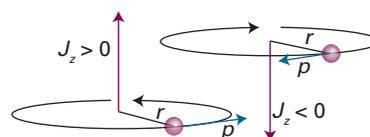
Sketch 20.2

The components of the angular momentum vector **J** when it lies in a general orientation are

$$J_x = yp_z - zp_y \qquad J_y = zp_x - xp_z \qquad J_z = xp_y - yp_x \tag{20.5}$$

where p_x is the component of the linear momentum in the x-direction at any instant, and likewise p_y and p_z in the other directions. The square of the magnitude of the vector is given by

$$J^2 = J_x^2 + J_y^2 + J_z^2 \tag{20.6}$$

By analogy with the expression for linear motion ($E_k = \frac{1}{2}mv^2 = p^2/2m$), the kinetic energy of a rotating object is

$$E_k = \tfrac{1}{2}I\omega^2 = \frac{J^2}{2I} \tag{20.7}$$

For a given moment of inertia, high angular momentum corresponds to high kinetic energy. As may be verified, the units of rotational energy are joules (J).

The analogous roles of m and I, of v and ω, and of p and J in the translational and rotational cases respectively provide a ready way of constructing and recalling equations. These analogies are summarized below:

| Translation		Rotation	
Property	**Significance**	**Property**	**Significance**
Mass, m	Resistance to the effect of a force	Moment of inertia, I	Resistance to the effect of a twisting force (torque)
Speed, v	Rate of change of position	Angular velocity, ω	Rate of change of angle
Magnitude of linear momentum, p	$p = mv$	Magnitude of angular momentum, J	$J = I\omega$
Translational kinetic energy, E_k	$E_k = \frac{1}{2}mv^2 = p^2/2m$	Rotational kinetic energy, E_k	$E_k = \frac{1}{2}I\omega^2 = J^2/2I$

THE CHEMIST'S TOOLKIT 21 Spherical polar coordinates

The mathematics of systems with spherical symmetry (such as atoms) is often greatly simplified by using **spherical polar coordinates** (Sketch 21.1): r, the distance from the origin (the radius), θ, the colatitude, and ϕ, the azimuth. The ranges of these coordinates are (with angles in radians, Sketch 21.2):

$$0 \leq r \leq +\infty, \; 0 \leq \theta \leq \pi, \; 0 \leq \phi \leq 2\pi$$
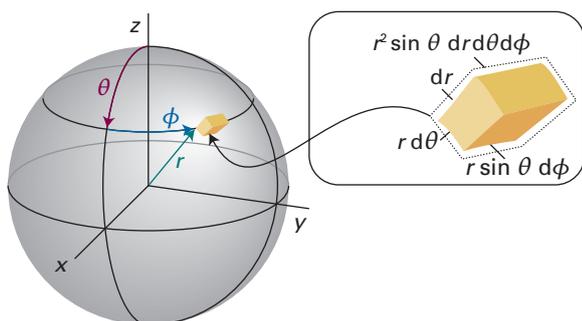
Sketch 21.1
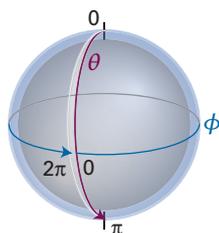
Sketch 21.2

An angle in radians is defined as the ratio of the length of an arc, s, to the radius r of a circle, so $\theta = s/r$. For a complete circle, the arc length is the circumference, $2\pi r$, so the angle subtended in radians for a complete revolution is $2\pi r/r = 2\pi$. That is, 360° corresponds to 2π radians, and consequently 180° corresponds to π radians.

Cartesian and polar coordinates are related by

$$x = r\sin\theta\cos\phi \quad y = r\sin\theta\sin\phi \quad z = r\cos\theta \tag{21.1}$$

The volume element in Cartesian coordinates is $d\tau = dxdydz$, and in spherical polar coordinates it becomes

$$d\tau = r^2 \sin\theta \, drd\theta d\phi \tag{21.2}$$

An integral of a function $f(r,\theta,\phi)$ over all space in polar coordinates therefore has the form

$$\int f \, d\tau = \int_0^\infty \int_0^\pi \int_0^{2\pi} f(r,\theta,\phi) r^2 \sin\theta \, drd\theta d\phi \tag{21.3}$$

where the limits on the integrations are for r, θ, and ϕ, respectively.

THE CHEMIST'S TOOLKIT 22 The manipulation of vectors

In three dimensions, the vectors $\boldsymbol{u}$ (with components u_x, u_y, and u_z) and $\boldsymbol{v}$ (with components v_x, v_y, and v_z) have the general form:

$$\boldsymbol{u} = u_x\boldsymbol{i} + u_y\boldsymbol{j} + u_z\boldsymbol{k} \qquad \boldsymbol{v} = v_x\boldsymbol{i} + v_y\boldsymbol{j} + v_z\boldsymbol{k} \qquad (22.1)$$

where $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ are **unit vectors**, vectors of magnitude 1, pointing along the positive directions on the x, y, and z axes. The operations of addition, subtraction, and multiplication are as follows:

1. *Addition*:

$$\boldsymbol{v} + \boldsymbol{u} = (v_x + u_x)\boldsymbol{i} + (v_y + u_y)\boldsymbol{j} + (v_z + u_z)\boldsymbol{k} \qquad (22.2)$$

2. *Subtraction*:

$$\boldsymbol{v} - \boldsymbol{u} = (v_x - u_x)\boldsymbol{i} + (v_y - u_y)\boldsymbol{j} + (v_z - u_z)\boldsymbol{k} \qquad (22.3)$$

Brief illustration 22.1: Addition and subtraction

Consider the vectors $\boldsymbol{u} = \boldsymbol{i} - 4\boldsymbol{j} + \boldsymbol{k}$ (of magnitude 4.24) and $\boldsymbol{v} = -4\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}$ (of magnitude 5.39) Their sum is

$$\boldsymbol{u} + \boldsymbol{v} = (1-4)\boldsymbol{i} + (-4+2)\boldsymbol{j} + (1+3)\boldsymbol{k} = -3\boldsymbol{i} - 2\boldsymbol{j} + 4\boldsymbol{k}$$

The magnitude of the resultant vector is $29^{1/2} = 5.39$. The difference of the two vectors is

$$\boldsymbol{u} - \boldsymbol{v} = (1+4)\boldsymbol{i} + (-4-2)\boldsymbol{j} + (1-3)\boldsymbol{k} = 5\boldsymbol{i} - 6\boldsymbol{j} - 2\boldsymbol{k}$$

The magnitude of this resultant is 8.06. Note that in this case the difference is longer than either individual vector.

---

3. *Multiplication*:

    (a) The **scalar product**, or *dot product*, of the two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_x v_x + u_y v_y + u_z v_z \qquad (22.4)$$

The scalar product of a vector with itself gives the square magnitude of the vector:

$$\boldsymbol{u} \cdot \boldsymbol{u} = u_x^2 + u_y^2 + u_z^2 = u^2 \qquad (22.5)$$

    (b) The **vector product,** or *cross product,* of two vectors is

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix}$$

$$= (u_y v_z - u_z v_y)\boldsymbol{i} - (u_x v_z - u_z v_x)\boldsymbol{j} + (u_x v_y - u_y v_x)\boldsymbol{k} \qquad (22.6)$$

(Determinants are discussed in *The chemist's toolkit* 23.) If the two vectors lie in the plane defined by the unit vectors $\boldsymbol{i}$ and $\boldsymbol{j}$, their vector product lies parallel to the unit vector $\boldsymbol{k}$.

Brief illustration 22.2: Scalar and vector products

The scalar and vector products of the two vectors in *Brief ilustration* 22.1, $\boldsymbol{u} = \boldsymbol{i} - 4\boldsymbol{j} + \boldsymbol{k}$ (of magnitude 4.24) and $\boldsymbol{v} = -4\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}$ (of magnitude 5.39) are

$$\boldsymbol{u} \cdot \boldsymbol{v} = \{1 \times (-4)\} + \{(-4) \times 2\} + \{1 \times 3\} = -9$$

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & -4 & 1 \\ -4 & 2 & 3 \end{vmatrix}$$

$$= \{(-4)(3) - (1)(2)\}\boldsymbol{i} - \{(1)(3) - (1)(-4)\}\boldsymbol{j} + \{(1)(2) - (-4)(-4)\}\boldsymbol{k}$$

$$= -14\boldsymbol{i} - 7\boldsymbol{j} - 14\boldsymbol{k}$$

The vector product is a vector of magnitude 21.00 pointing in a direction perpendicular to the plane defined by the two individual vectors.

Further information

The manipulation of vectors is commonly represented graphically. Consider two vectors $\boldsymbol{v}$ and $\boldsymbol{u}$ making an angle θ (Sketch 22.1a). The first step in the addition of $\boldsymbol{u}$ to $\boldsymbol{v}$ consists of joining the tip (the 'head') of $\boldsymbol{u}$ to the starting point (the 'tail') of $\boldsymbol{v}$ (Sketch 22.1b). In the second step, draw a vector $\boldsymbol{v}_{res}$, the **resultant vector**, originating from the tail of $\boldsymbol{u}$ to the head of $\boldsymbol{v}$ (Sketch 22.1c). Reversing the order of addition leads to the same result; that is, the same $\boldsymbol{v}_{res}$ is obtained whether $\boldsymbol{u}$ is added to $\boldsymbol{v}$ or $\boldsymbol{v}$ to $\boldsymbol{u}$. To calculate the magnitude of $\boldsymbol{v}_{res}$, note that

$$\begin{aligned} v_{res}^2 &= (\boldsymbol{u} + \boldsymbol{v}) \cdot (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{u} + \boldsymbol{v} \cdot \boldsymbol{v} + 2\boldsymbol{u} \cdot \boldsymbol{v} \\ &= u^2 + v^2 + 2uv \cos \theta' \end{aligned} \qquad (22.7a)$$
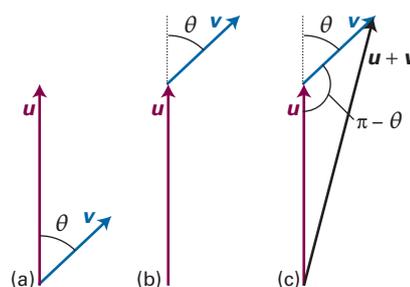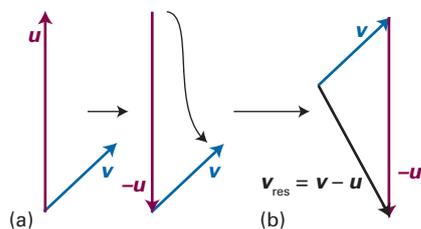

Sketch 22.1

where θ' is the angle between u and v. In terms of the angle $\theta = \pi - \theta'$ shown in the Sketch, and $\cos(\pi - \theta) = -\cos\theta$,

$$v_{res}^2 = u^2 + v^2 - 2uv\cos\theta \qquad \text{Law of cosines} \qquad (22.7b)$$

which is the **law of cosines** for the relation between the lengths of the sides of a triangle.

Sketch 22.2

Subtraction of u from v amounts to addition of $-u$ to v. It follows that in the first step of subtraction $-u$ is drawn by reversing the direction of u (Sketch 22.2a). Then, the second step consists of adding $-u$ to v by using the strategy shown in the Sketch; a resultant vector v_{res} is drawn by joining the tail of $-u$ to the head of v (Sketch 22.2b).
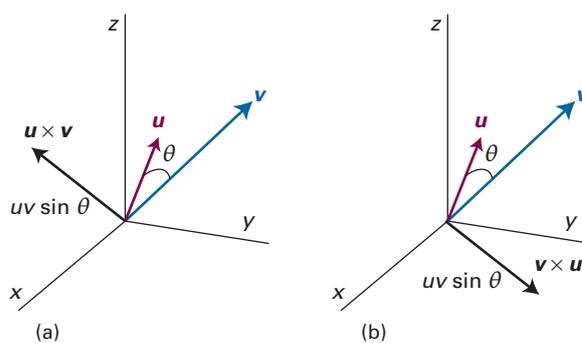
Sketch 22.3

Vector multiplication is represented graphically by drawing a vector (using the right–hand rule) perpendicular to the plane defined by the vectors u and v, as shown in Sketch 22.3. Its length is equal to $uv\sin\theta$, where θ is the angle between u and v. Note that $u \times v$ (Sketch 22.3a) is opposite in direction to $v \times u$ (Sketch 22.3b).

THE CHEMIST'S TOOLKIT 23 Determinants

A 2×2 determinant is the entity

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

2×2 Determinant (23.1a)

A 3×3 determinant is evaluated by expanding it as a sum of 2×2 determinants:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= a(ei - fh) - b(di - fg) + c(dh - eg)$$

3×3 Determinant (23.1a)

Note the sign change in alternate columns (b occurs with a negative sign in the expansion). An important property of a determinant is that if any two rows or any two columns are interchanged, then the determinant changes sign:

Exchange columns:

$$\begin{vmatrix} b & a \\ d & c \end{vmatrix} = bc - ad = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Exchange rows:

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - da = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

An implication is that if any two columns or rows are identical, then the determinant is zero.

THE CHEMIST'S TOOLKIT 24 Matrices

A **matrix** is an array of numbers arranged in a certain number of rows and a certain number of columns; the numbers of rows and columns may be different. The rows and columns are numbered 1, 2, … so that the number at each position in the matrix, called the **matrix element**, has a unique row and column index. The element of a matrix M at row r and column c is denoted M_{rc}. For instance, a 3×3 matrix is

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}$$

The **trace** of a matrix, Tr M, is the sum of the diagonal elements.

$$\text{Tr } M = \sum_n M_{nn} \tag{24.1}$$

In this case

$$\text{Tr } M = M_{11} + M_{22} + M_{33}$$

A **unit** matrix has diagonal elements equal to 1 and all other elements zero. A 3×3 unit matrix is therefore

$$1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Matrices are added by adding the corresponding matrix elements. Thus, to add the matrices A and B to give the sum $S = A + B$, each element of S is given by

$$S_{rc} = A_{rc} + B_{rc} \tag{24.2}$$

Only matrices of the same dimensions can be added together.

Matrices are multiplied to obtain the product $P = AB$; each element of P is given by

$$P_{rc} = \sum_n A_{rn} B_{nc} \tag{24.3}$$

Matrices can be multiplied only if the number of columns in A is equal to the number of rows in B. Square matrices (those with the same number of rows and columns) can therefore be multiplied only if both matrices have the same dimension (that is, both are $n \times n$). The products AB and BA are not necessarily the same, so matrix multiplication is in general 'non-commutative'.

Brief illustration 24.1: Matrix addition and multiplication

Consider the matrices

$$M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ and } N = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

Their sum is

$$S = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$$

and their product is

$$P = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{pmatrix}$$

$$= \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

An $n \times 1$ matrix (with n elements in one column) is called a **column vector**. It may be multiplied by a square $n \times n$ matrix to generate a new column vector, as in

$$\begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \times \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix}$$

The elements of the two column vectors need only one index to indicate their row. Each element of P is given by

$$P_r = \sum_n A_{rn} B_n \tag{24.4}$$

A $1 \times n$ matrix (a single row with n elements) is called a **row vector**. It may be multiplied by a square $n \times n$ matrix to generate a new row vector, as in

$$(P_1 \, P_2 \, P_3) = (B_1 \, B_2 \, B_3) \times \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

In general the elements of P are

$$P_c = \sum_n B_n A_{nc} \tag{24.5}$$

Note that a column vector is multiplied 'from the left' by the square matrix and a row vector is multiplied 'from the right'. The **inverse** of a matrix A, denoted A^{-1}, has the property that $AA^{-1} = A^{-1}A = 1$, where 1 is a unit matrix with the same dimensions as A.

Brief illustration 24.2: Inversion

Mathematical software gives the following inversion of a matrix A:

	Matrix	Inverse
	A	A^{-1}

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

In matrix form, an **eigenvalue equation** is

$$Mx = \lambda x \qquad \text{Eigenvalue equation} \qquad (25.1a)$$

where M is a square matrix with n rows and n columns, λ is a constant, the **eigenvalue**, and x is the **eigenvector**, an $n \times 1$ (column) matrix that satisfies the conditions of the eigenvalue equation and has the form:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

In general, there are n eigenvalues $\lambda^{(i)}$, $i = 1, 2, \ldots, n$, and n corresponding eigenvectors $x^{(i)}$. Equation 25.1a can be rewritten as

$$(M - \lambda 1)x = 0 \qquad (25.1b)$$

where 1 is an $n \times n$ unit matrix, and where the property $1x = x$ has been used. This equation has a solution only if the determinant $|M - \lambda 1|$ of the matrix $M - \lambda 1$ is zero. It follows that the n eigenvalues may be found from the solution of the **secular equation**:

$$|M - \lambda 1| = 0 \qquad (25.2)$$

Brief illustration 25.1: Simultaneous equations

Consider the matrix equation

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

rearranged into $\begin{pmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$

From the rules of matrix multiplication, the latter form expands into

$$\begin{pmatrix} (1-\lambda)x_1 + 2x_2 \\ 3x_1 + (4-\lambda)x_2 \end{pmatrix} = 0$$

which is simply a statement of the two simultaneous equations

$$(1-\lambda)x_1 + 2x_2 = 0 \text{ and } 3x_1 + (4-\lambda)x_2 = 0$$

The condition for these two equations to have solutions is

$$|M - \lambda 1| = \begin{vmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) - 6 = 0$$

This condition corresponds to the quadratic equation

$$\lambda^2 - 5\lambda - 2 = 0$$

with solutions $\lambda = +5.372$ and $\lambda = -0.372$, the two eigenvalues of the original equation.

The n eigenvalues found by solving the secular equations are used to find the corresponding eigenvectors. To do so, begin by considering an $n \times n$ matrix X the columns of which are formed from the eigenvectors corresponding to all the eigenvalues. Thus, if the eigenvalues are $\lambda_1, \lambda_2, \ldots$, and the corresponding eigenvectors are

$$x^{(1)} = \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \quad x^{(2)} = \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix} \cdots x^{(n)} = \begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_n^{(n)} \end{pmatrix} \qquad (25.3a)$$

then the matrix X is

$$X = (x^{(1)} \; x^{(2)} \cdots x^{(n)}) = \begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \cdots & x_1^{(n)} \\ x_2^{(1)} & x_2^{(2)} & \cdots & x_2^{(n)} \\ \vdots & \vdots & & \vdots \\ x_n^{(1)} & x_n^{(2)} & \cdots & x_n^{(n)} \end{pmatrix} \qquad (25.3b)$$

Similarly, form an $n \times n$ matrix Λ with the eigenvalues λ along the diagonal and zeroes elsewhere:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \qquad (25.4)$$

Now all the eigenvalue equations $Mx^{(i)} = \lambda_i x^{(i)}$ may be combined into the single matrix equation

$$MX = X\Lambda \qquad (25.5)$$

Brief illustration 25.2: Eigenvalue equations

In *Brief illustration* 25.1 it is established that if $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ then $\lambda_1 = +5.372$ and $\lambda_2 = -0.372$. Then, with eigenvectors

$$x^{(1)} = \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} \text{ and } x^{(2)} = \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \end{pmatrix} \text{ form}$$

$$X = \begin{pmatrix} x_1^{(1)} & x_1^{(2)} \\ x_2^{(1)} & x_2^{(2)} \end{pmatrix} \quad \Lambda = \begin{pmatrix} 5.372 & 0 \\ 0 & -0.372 \end{pmatrix}$$

The expression $MX = X\Lambda$ becomes

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1^{(1)} & x_1^{(2100)} \\ x_2^{(1)} & x_2^{(2)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)} & x_1^{(2)} \\ x_2^{(1)} & x_2^{(2)} \end{pmatrix} \begin{pmatrix} 5.372 & 0 \\ 0 & -0.372 \end{pmatrix}$$

which expands to

$$
\begin{pmatrix} x_1^{(1)}+2x_2^{(1)} & x_1^{(2)}+2x_2^{(2)} \\ 3x_1^{(1)}+4x_2^{(1)} & 3x_1^{(2)}+4x_2^{(2)} \end{pmatrix} = \begin{pmatrix} 5.372x_1^{(1)} & -0.372x_1^{(2)} \\ 5.372x_2^{(1)} & -0.372x_2^{(2)} \end{pmatrix}
$$

This is a compact way of writing the four equations

$$
x_1^{(1)}+2x_2^{(1)}=5.372x_1^{(1)} \quad x_1^{(2)}+2x_2^{(2)}=-0.372x_1^{(2)}
$$
$$
3x_1^{(1)}+4x_2^{(1)}=5.372x_2^{(1)} \quad 3x_1^{(2)}+4x_2^{(2)}= -0.372x_2^{(2)}
$$

corresponding to the two original simultaneous equations and their two roots.

Finally, form X^{-1} from X and multiply eqn 25.5 by it from the left:

$$
X^{-1}MX = X^{-1}X\Lambda = \Lambda \tag{25.6}
$$

A structure of the form $X^{-1}MX$ is called a **similarity transformation**. In this case the similarity transformation $X^{-1}MX$ makes M diagonal (because Λ is diagonal). It follows that if the matrix X that causes $X^{-1}MX$ to be diagonal is known, then the problem is solved: the diagonal matrix so produced has the eigenvalues as its only nonzero elements, and the matrix X used to bring about the transformation has the corresponding eigenvectors as its columns. In practice, the eigenvalues and eigenvectors are obtained by using mathematical software.

Brief illustration 25.3: Similarity transformations

To apply the similarity transformation to the matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ from *Brief illustration* 25.1 it is best to use mathematical software to find the form of X and X^{-1}. The result is

$$
X = \begin{pmatrix} 0.416 & 0.825 \\ 0.909 & -0.566 \end{pmatrix} \quad X^{-1} = \begin{pmatrix} 0.574 & 0.837 \\ 0.922 & -0.422 \end{pmatrix}
$$

This result can be verified by carrying out the multiplication

$$
X^{-1}MX = \begin{pmatrix} 0.574 & 0.837 \\ 0.922 & -0.422 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0.416 & 0.825 \\ 0.909 & -0.566 \end{pmatrix}
$$
$$
= \begin{pmatrix} 5.372 & 0 \\ 0 & -0.372 \end{pmatrix}
$$

The result is indeed the diagonal matrix Λ calculated in *Brief illustration* 25.2. It follows that the eigenvectors $x^{(1)}$ and $x^{(2)}$ are
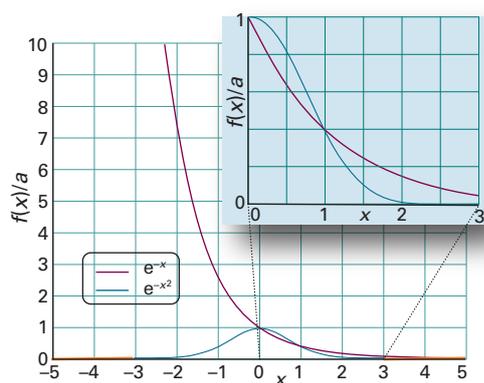
$$
x^{(1)} = \begin{pmatrix} 0.416 \\ 0.909 \end{pmatrix} \quad x^{(2)} = \begin{pmatrix} 0.825 \\ -0.566 \end{pmatrix}
$$

THE CHEMIST'S TOOLKIT 26 Exponential and Gaussian functions

An **exponential function** is a function of the form

$$f(x) = ae^{-bx}$$

Exponential function (26.1)

This function has the value a at $x=0$ and decays toward zero as $x \to \infty$. This decay is faster when b is large than when it is small. The function rises rapidly to infinity as $x \to -\infty$. See Sketch 26.1.

The general form of a **Gaussian function** is

$$f(x) = ae^{-(x-b)^2/2\sigma^2}$$

Gaussian function (26.2)

The graph of this function is a symmetrical bell-shaped curve centred on $x = b$; the function has its maximum values of a at its centre. The width of the function, measured at half its height, is $\delta x = 2\sigma(2\ln 2)^{1/2}$; the greater σ, the greater is the width at half-height. Sketch 26.1 also shows a Gaussian function with $b = 0$ and $2\sigma^2 = 1$.

Sketch 26.1

THE CHEMIST'S TOOLKIT 27 Dipolar magnetic fields

Standard electromagnetic theory gives the magnetic field at a point r from a point magnetic dipole μ as

$$\mathcal{B} = -\frac{\mu_0}{4\pi r^3}\left(\mu - \frac{3(\mu \cdot r)r}{r^2}\right) \tag{27.1}$$

where μ_0 is the vacuum permeability (a fundamental constant with the defined value $4\pi \times 10^{-7}\,\text{T}^2\,\text{J}^{-1}\,\text{m}^3$). The component of magnetic field in the z-direction is

$$\mathcal{B}_z = -\frac{\mu_0}{4\pi r^3}\left(\mu_z - \frac{3(\mu \cdot r)z}{r^2}\right) \tag{27.2}$$

with $z = r\cos\theta$, the z-component of the distance vector r. If the magnetic dipole is also parallel to the z-direction, it follows that

$$\mathcal{B}_z = -\frac{\mu_0}{4\pi r^3}\left(\overset{\mu_z}{\overbrace{\mu}} - \frac{3\overset{\mu \cdot r}{\overbrace{(\mu r\cos\theta)}}\overset{z}{\overbrace{(r\cos\theta)}}}{r^2}\right) = -\frac{\mu\mu_0}{4\pi r^3}(1 - 3\cos^2\theta) \tag{27.3}$$

THE CHEMIST'S TOOLKIT 28 The Fourier transform
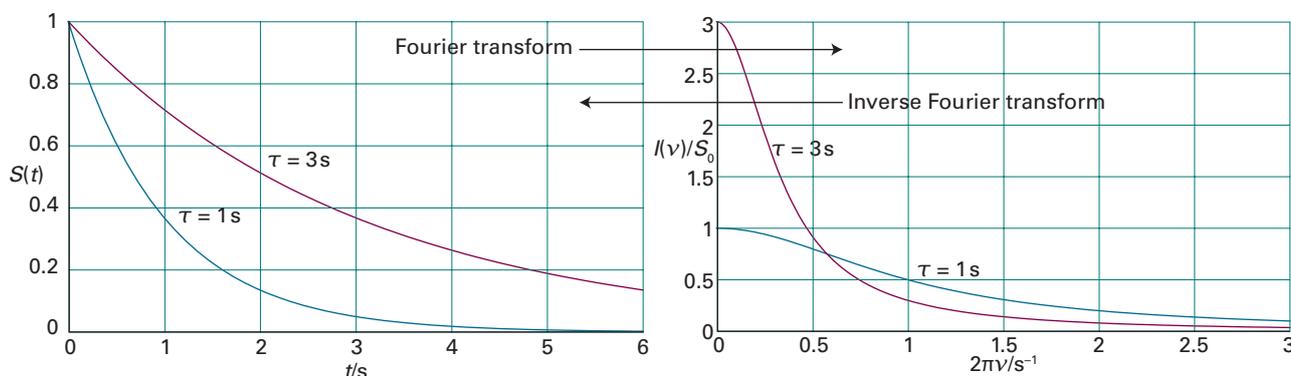
Sketch 28.1

A **Fourier transform** expresses any waveform as a superposition of harmonic (sine and cosine) waves. If the waveform is $S(t)$, then the contribution $I(v)$ of the oscillating function $\cos(2\pi vt)$ is given by the integral

$$I(v) = \int_0^\infty S(t)\cos(2\pi vt)\,dt \tag{28.1}$$

If the signal varies slowly, then the greatest contribution comes from low-frequency waves; rapidly changing features in the signal are reproduced by high-frequency contributions. If the signal is a simple exponential decay of the form $S(t) = S_0 e^{-t/\tau}$, the contribution of the wave of frequency v is

$$I(v) = S_0 \int_0^\infty e^{-t/\tau}\cos(2\pi vt)\,dt = \frac{S_0\tau}{1+(2\pi v\tau)^2} \tag{28.2}$$

Sketch 28.1 shows a fast and slow decay and the corresponding frequency contributions: note that a slow decay has predominantly low-frequency contributions and a fast decay has many high-frequency contributions.

If an experimental procedure results in the function $I(v)$ itself, then the corresponding signal can be reconstructed by forming the **inverse Fourier transform**:

$$S(t) = \frac{2}{\pi}\int_0^\infty I(v)\cos(2\pi vt)\,dv \tag{28.3}$$

There are complex versions of these cosine transforms, as described below.

Fourier transforms are applicable to spatial functions too. Their interpretation is similar but it is more appropriate to think in terms of the wavelengths of the contributing waves. Thus, if the function varies only slowly with distance, then its Fourier transform has mainly long-wavelength contributions. If the features vary quickly with distance (as in the electron density in a crystal), then short-wavelength contributions feature.

Further information

Some insight into the physical significance of taking a Fourier transform can be obtained by considering the process for analysing a wave of general form, like that at the top of each part of Sketch 28.2. According to eqn 28.1, the procedure involves forming the product of the waveform and a cosine wave with frequency v, and then determining the area under the product.
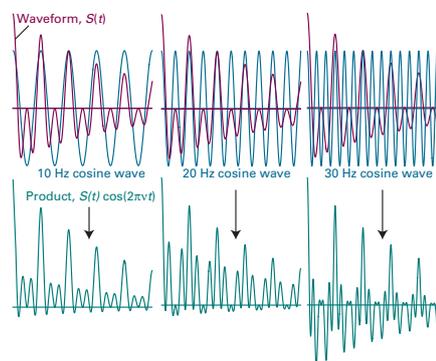
Sketch 28.2

When $S(t)$ is multiplied by a cosine wave with frequency 10 Hz, the oscillations in the two functions largely coincide, with the result that the product $S(t)\cos(2\pi vt)$ has more positive peaks than negative peaks, and therefore a non-zero area. The wave of frequency 10 Hz, therefore makes a significant contribution. When the procedure is repeated with a cosine function oscillating at 20 Hz, the product also results in a non-zero area, so a cosine function oscillating at this frequency also makes a significant contribution to the original waveform. However, if the frequency of the cosine function is 30 Hz, the product has as many positive and negative peaks and the area is essentially zero. A cosine function at this frequency makes a negligible contribution to the waveform.

The most general formulation of a Fourier transform is to express the function $f(t)$ as a superposition of cosine and sine functions, not just cosine functions. The two types of functions can be handled simultaneously by using de Moivre's relation $e^{ix} = \cos x + i \sin x$ and writing

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(v) e^{2\pi i v} \, dv \qquad (28.4a)$$

where $\tilde{f}(v)$ is the Fourier transform of $f(t)$, and can be interpreted as the amplitude of the contribution of the cosine and sine waves in the superposition that recreates the function $f(t)$. The inverse relation is

$$\tilde{f}(v) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i v t} \, dt \qquad (28.4b)$$

The cosine contribution is given by the real part of $\tilde{f}(v)$ and the sine contribution is given by the imaginary part.

Brief illustration 28.1: The Fourier transform of an oscillating, exponentially decaying wave

The introductory part of this Toolkit illustrated the result of a cosine Fourier transformation of an exponential decay. It is instructive to consider the complex version of that calculation and to generalize it to a function that oscillates with a frequency v_0 as it decays. To avoid the repetition of many factors of 2π, henceforth, write $2\pi v = \omega$. Then the (complex) function in the time domain is

$$f(t) = \begin{cases} f_0 \overbrace{e^{i\omega_0 t}}^{\text{Oscillating component}} \overbrace{e^{-t/\tau}}^{\text{Decaying exponential}} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases}$$

The observed signal is the real part of $f(t)$, bearing in mind that $\mathrm{Re}\, e^{ix} = \cos x$. See Sketch 28.3 (which also displays the Fourier transform, as explained below). The Fourier transform of this function is

$$\tilde{f}(v) = \int_0^{\infty} f_0 e^{i\omega_0 t} e^{-t/\tau} e^{-i\omega t} \, dt = f_0 \int_0^{\infty} e^{\{i(\omega_0 - \omega) - 1/\tau\} t} \, dt \qquad (28.5)$$

$$= f_0 \frac{e^{\{i(\omega_0 - \omega) - 1/\tau\} t}}{i(\omega_0 - \omega) - 1/\tau} \Bigg|_0^{\infty} = f_0 \frac{-1}{i(\omega_0 - \omega) - 1/\tau} \qquad (28.6)$$

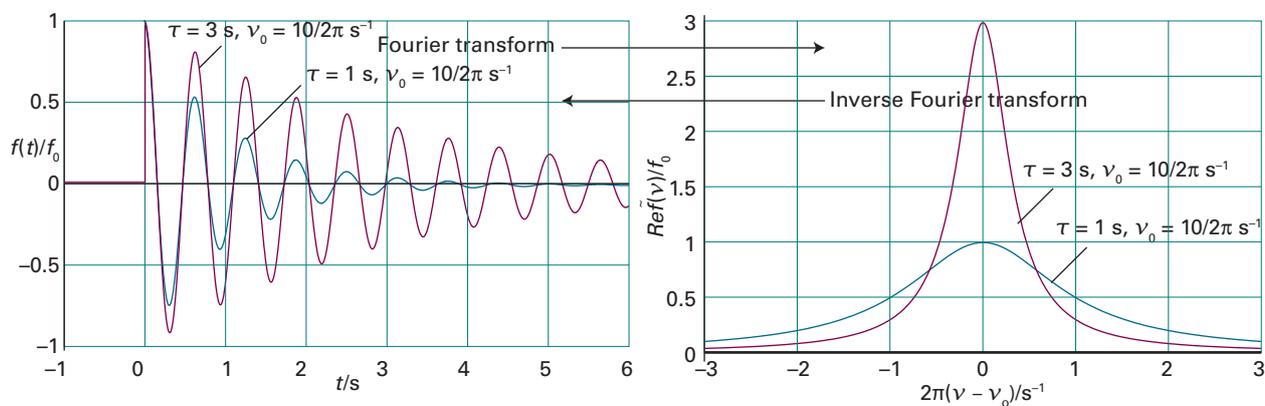The fraction of the right has both real an imaginary parts; they can be extracted by multiplying the numerator and denominator by the complex conjugate of the denominator

$$\tilde{f}(v) = f_0 \frac{-1}{i(\omega_0 - \omega) - 1/\tau} \times \frac{-i(\omega_0 - \omega) - 1/\tau}{-i(\omega_0 - \omega) - 1/\tau}$$

$$= f_0 \frac{i(\omega_0 - \omega) + 1/\tau}{(\omega_0 - \omega)^2 + 1/\tau^2} \qquad (28.7)$$

The real part of $\tilde{f}(v)$ is therefore

$$\mathrm{Re}\, \tilde{f}(v) = \frac{f_0/\tau}{(\omega_0 - \omega)^2 + 1/\tau^2} = \frac{f_0 \tau}{(\omega_0 - \omega)^2 \tau^2 + 1} \qquad (28.8)$$

which is essentially the same as in eqn 28.2 with the exception that the frequency $2\pi v = \omega$ has been replaced by $\omega_0 - \omega$.

Sketch 28.3

THE CHEMIST'S TOOLKIT 29 Electrostatics

A charge Q_1 (units: coulomb, C) gives rise to a Coulomb **potential** ϕ (units: volt, V) , as explained in *The chemist's toolkit 6*. The potential energy (units: joule, J, with $1\,\text{J} = 1\,\text{V C}$) of a second charge Q in that potential is

$$E_p = -Q\phi \tag{29.1}$$

In one dimension, the **electric field strength** (units: volt per metre, V m^{-1}), $\mathcal{E}$, is the negative of the gradient of the electric potential ϕ:

$$\mathcal{E} = -\frac{d\phi}{dx} \qquad \text{Electric field strength} \tag{29.2}$$

In three dimensions the electric field is a vector, and

$$\mathcal{E} = -\nabla\phi \tag{29.3}$$

The electric field between two plane parallel plates separated by a distance l, and between which there is a potential difference $\Delta\phi$, is uniform and given by

$$\mathcal{E} = -\frac{\Delta\phi}{l} \tag{29.4}$$

A charge Q experiences a force proportional to the electric field strength at its location:

$$F_{electric} = Q\mathcal{E} \tag{29.5}$$

A potential gives rise to a force only if it varies with distance.

THE CHEMIST'S TOOLKIT 30 Integration by the method of partial fractions

To solve an integral of the form

$$I = \int \frac{1}{(a-x)(b-x)} dx \tag{30.1}$$

where a and b are constants with $a \neq b$, use the **method of partial fractions** in which a fraction that is the product of terms (as in the denominator of this integrand) is written as a sum of fractions. To implement this procedure write the integrand as

$$\frac{1}{(a-x)(b-x)} = \frac{1}{b-a}\left(\frac{1}{a-x} - \frac{1}{b-x}\right)$$

Then integrate each term on the right. It follows that

$$I = \frac{1}{b-a}\left(\overbrace{\int \frac{dx}{a-x}}^{\text{Integral A.2}} - \overbrace{\int \frac{dx}{b-x}}^{\text{Integral A.2}}\right)$$

$$= \frac{1}{b-a}\left(\ln\frac{1}{a-x} - \ln\frac{1}{b-x}\right) + \text{constant} \tag{30.2}$$

Further information

Although the condition $a \neq b$ has been specified, the result is also valid for $a = b$ provided the equality is interpreted as the limit $b \to a$. Thus, write $b = a + \delta$, with $\delta \to 0$; then, by using $\ln(1 + z) = z + \frac{1}{2}z^2 + \ldots = z + O(z^2)$,

$$\lim_{\delta \to 0} \frac{1}{b-a}\left(\ln\frac{1}{a-x} - \ln\frac{1}{b-x}\right)$$

$$= \lim_{\delta \to 0} \frac{1}{a+\delta-a}\left(\ln\frac{1}{a-x} - \ln\frac{1}{a+\delta-x}\right)$$

$$= \lim_{\delta \to 0} \frac{1}{\delta}\ln\frac{a+\delta-x}{a-x} = \lim_{\delta \to 0} \frac{1}{\delta}\ln\left(1 + \frac{\delta}{a-x}\right)$$

$$= \lim_{\delta \to 0} \frac{1}{\delta}\left(\frac{\delta}{a-x} + O(\delta^2)\right) = \frac{1}{a-x}$$

That is, in this limit

$$\int \frac{1}{(a-x)^2} dx = \frac{1}{a-x} + \text{constant}$$