MATERIALS SCIENCE AND METALLURGY
Course C3: Mathematical Methods

AIMS OF THE COURSE

The principal aim of this course is revision since most of you will not have taken Mathematics for Natural Sciences in Part IB of the Natural Sciences Tripos.

Ideally you should find little new in this course from the Part IA Mathematics for Natural Sciences course and the mathematical components of the other Natural Sciences courses you have taken at Part IA and Part IB. You may come across some once familiar, but since forgotten, bits of mathematics. You may also find a familiar topic in materials science looked at in a new way in order to illustrate the usefulness of a particular piece of mathematics.

Although each lecture will concentrate on the mathematical topics listed, I will draw heavily throughout the course on examples taken from materials science where the ability to use mathematics as a tool can be usefully illustrated.

Note that while I shall be reasonably rigorous, I will not attempt to match the standards of rigour that would be expected in the Mathematical Tripos.

By the end of the course, i.e., after the lectures **and** after working through the Examples Class and the Question Sheet, I hope you will feel confident in dealing with mathematically-based topics wherever they arise in Part II and Part III Materials Science & Metallurgy.

If you do have particular difficulties which cannot be resolved through the Examples Class or in consultation with your supervisors, please contact me, e.g. by e-mail at kmk10@cam.ac.uk or by a note in my pigeon-hole.

When it comes to the examination I shall be surprised if the examiners set any questions based **solely** on these lectures and appalled if there are not lots of materials science and metallurgy questions involving some use of mathematics!

Kevin Knowles
October 2005

SYNOPSIS

Introduction (2 lectures)

Aims of the course, linkage to Examples Class, Question Sheets, supervisions and examination. Revision of some important points in algebra and calculus: integration, differentiation, hyperbolic functions, vector algebra and its use in crystallography, complex numbers, algebraic approximations.

Direction cosines, differentiation of exponentials, approximations based on series expansions and their application to models of ionic conductivity, kinetics of electrode reactions and crystal growth.

Taylor's series. Summation of geometric series and application to stepwise polymerisation. Solutions of simple second order linear differential equations: the damped simple harmonic oscillator and the shear lag equation. Fourier series and Fourier transforms and their application in materials science.

Matrix algebra (1 lecture)

Determinants, trace of a matrix, eigenvectors, eigenvalues, principal axes, coordinate transformations, similarity transformations, invariants of symmetric matrices, illustrated with real symmetric 2×2 and 3×3 matrices. Application to 3×3 stress and strain tensors and 2-dimensional stress states in thin laminae of long fibre composites.

First-order differential equations (1 lecture)

General solutions of simple first-order differential equations. Application to solidification, the electrical response of lossy dielectrics and viscoelastic behaviour.

Diffusion (1 lecture)

Fick's first law. Fick's second law. The diffusion equation and some specific solutions: the error function solution, the thin film solution, separation of variables and its application to diffusion out of a plane sheet. Steady state solidification.

Error analysis (1 lecture)

Errors and their treatment. Statistics and statistical distributions: Normal (Gaussian), Poisson, Weibull. Regression and curve fitting. Confidence and significance.

BOOK LIST

Your notes and any of the text books that you used for the Mathematics course in Part IA of the Natural Sciences Tripos will be helpful. The main text books recommended for the Part IA course include:

M.L. Boas, *Mathematical Methods in the Physical Sciences*, 2nd. edn., Wiley, 1983	Za52
E. Kreyszig, *Advanced Engineering Mathematics*, 7th. edn., Wiley 1993	Za77
K.F. Riley, M.P. Hobson and S.J. Bence, *Mathematical Methods for Physics and Engineering,* CUP, 1997	Za91
G. Stephenson, *Mathematical Methods for Science Students*, 2nd. edn., Longman, 1973	Za24

The following are examples of other helpful books:

G.B. Arfken and H.J. Weber *Mathematical Methods for Physicists*, 4th edn, Academic Press 1995	Za78
J. Crank, *The Mathematics of Diffusion*, Oxford, 1979 (reprinted 1999)	Lr66
G.L. Squires, *Practical Physics*, 4th. edn., Cambridge, 2001	LcZ26
J.R. Taylor, *An Introduction to Error Analysis*, University Science Books, 1982 (2nd edn 1997)	Za92

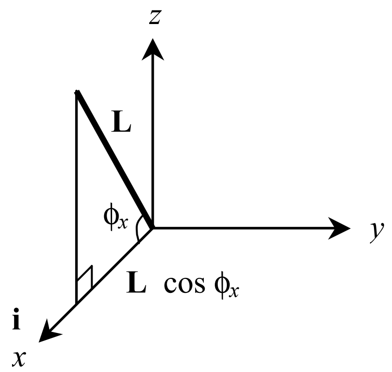
If you wish to check points from A-level Mathematics there are many books from which to choose. A good example is:

L. Bostock and S. Chandler, *Core Maths for A-level*, Stanley Thornes, 1994	Za76

Not all are available in the Departmental Library but they are widely available in College Libraries, including very similar older or newer editions in some cases.

Finally, you will find it useful to refer back to sections of your Part IA Materials & Mineral Sciences and Part IB Materials Science & Metallurgy lecture notes in which many of these mathematical topics were explicitly or implicitly used.

DIRECTION COSINES



The direction cosines, ℓ, m and n of the line **L** are the cosines of the angles between **L** and each of the axes in turn. Thus:

$$\ell = \cos \phi_x$$

$$m = \cos \phi_y$$

$$n = \cos \phi_z$$

For orthogonal axes

$$\boxed{\ell^2 + m^2 + n^2 = 1}$$

This can be shown by resolving **L** into components along the axes.

Component along x: $L_x = L \cos \phi_x \, \mathbf{i} = L\ell\,\mathbf{i}$, where **i** is a unit vector in the x direction.

Component along y: $L_y = L \cos \phi_y \, \mathbf{j} = Lm\mathbf{j}$, where **j** is a unit vector in the y direction.

Component along z: $L_z = L \cos \phi_z \, \mathbf{k} = Ln\mathbf{k}$, where **k** is a unit vector in the z direction.

Hence: $\mathbf{L} = L_x\,\mathbf{i} + L_y\,\mathbf{j} + L_z\,\mathbf{k} = L\ell\,\mathbf{i} + Lm\mathbf{j} + Ln\mathbf{k} = L(\ell\,\mathbf{i} + m\mathbf{j} + n\mathbf{k})$

‘Dotting’ a vector with itself gives a scalar equal to the square of the magnitude of the vector. Thus, for example, $\mathbf{i.i} = \mathbf{j.j} = \mathbf{k.k} = 1$.

Also, for three mutually perpendicular *orthogonal* axes, $\mathbf{i.j} = \mathbf{j.k} = \mathbf{k.i} = 0$.

Therefore: $L^2 = \mathbf{L.L} = L^2\left(\ell^2 + m^2 + n^2\right)$

from which it immediately follows that $\ell^2 + m^2 + n^2 = 1$

DELTA FUNCTIONS

$\delta(x-a) = 0$ unless $x = a$.

$$\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1$$

$$\int_{-\infty}^{\infty} \delta(x-a)\,f(x)\,\mathrm{d}x = f(a)$$

DIFFERENTIATION OF EXPONENTIALS

Preliminaries:

$$y = \exp\left(f(x)\right)$$

$$\ln y = f(x)$$

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}x}$$

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y\frac{\mathrm{d}f}{\mathrm{d}x}$$

$$\boxed{\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}x}\exp\left(f(x)\right)}$$

Application :

$$y = \exp\left(-\frac{Q}{kT}\right)$$

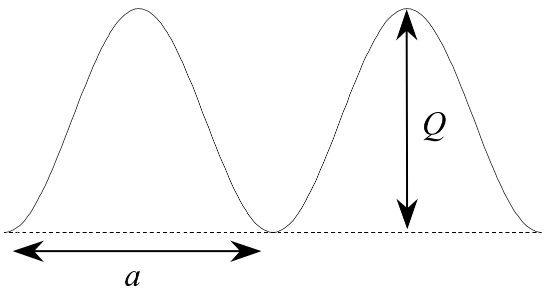
$$f = -\frac{Q}{kT}$$

$$\frac{\mathrm{d}f}{\mathrm{d}T} = +\frac{Q}{kT^2}$$

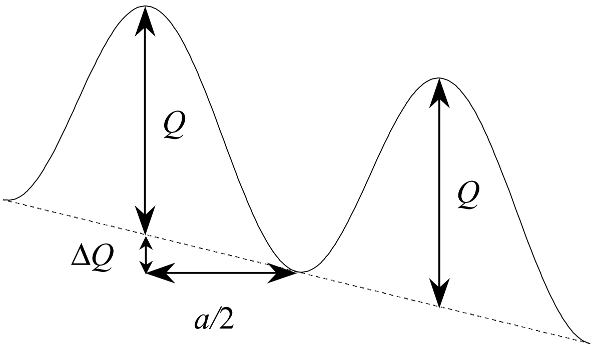
$$\boxed{\frac{\mathrm{d}y}{\mathrm{d}T} = \frac{Q}{kT^2}\exp\left(-\frac{Q}{kT}\right)}$$

APPROXIMATIONS BASED ON SERIES EXPANSIONS

Consider, for example, the effect of an electric field on the movement of an ion in a solid.



In zero field



In field *E*

Let the ions be of charge *ze*, in a field *E*, with sites distance *a* apart.

Change in barrier height: ΔQ = magnitude of field × distance × charge

$$\Delta Q = E\frac{a}{2}ze$$

Jump probabilities:

$$p_{left} = p_0 \exp\left(-\frac{(Q+\Delta Q)}{kT}\right)$$

$$p_{right} = p_0 \exp\left(-\frac{(Q-\Delta Q)}{kT}\right)$$

Net probability of jump to right

$$\Delta p = p_{right} - p_{left} = p_0 \exp\left(-\frac{(Q-\Delta Q)}{kT}\right) - p_0 \exp\left(-\frac{(Q+\Delta Q)}{kT}\right)$$

$$= p_0 \exp\left(-\frac{Q}{kT}\right)\left[\exp\left(\frac{\Delta Q}{kT}\right) - \exp\left(-\frac{\Delta Q}{kT}\right)\right]$$

$$= 2p_0 \exp\left(-\frac{Q}{kT}\right)\sinh\left(\frac{\Delta Q}{kT}\right) \qquad \text{using the identity } \sinh x = \frac{1}{2}\left(\mathrm{e}^{x} - \mathrm{e}^{-x}\right)$$

If *x* is small,

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots \approx x \qquad \text{(using data given in the Data Book)}$$

Is *x* small in this particular case? **Always check the validity of such approximations!!**

Here $$x = \frac{zeEa}{2kT}$$

e.g., for the gate oxide of a field effect transistor (FET),we can take

$E \sim 10^6$ V m^{-1} $a \sim 0.3$ nm
$ze = 1e = 1.6 \times 10^{-19}$ C $k = 1.4 \times 10^{-23}$ J T^{-1}

For $T \sim 300$ K, $x \sim 0.007$, so x is indeed small; however for $T \sim 4$ K, $x \sim 0.6$, so the approximation is much less good (but at so low a temperature, need we worry?)

Therefore we can say that to a good approximation the net probability of a jump to the right is:

$$\Delta p = 2p_0 \exp\left(-\frac{Q}{kT}\right)\frac{\Delta Q}{kT} = p_0 \exp\left(-\frac{Q}{kT}\right)\frac{zea}{kT}E$$

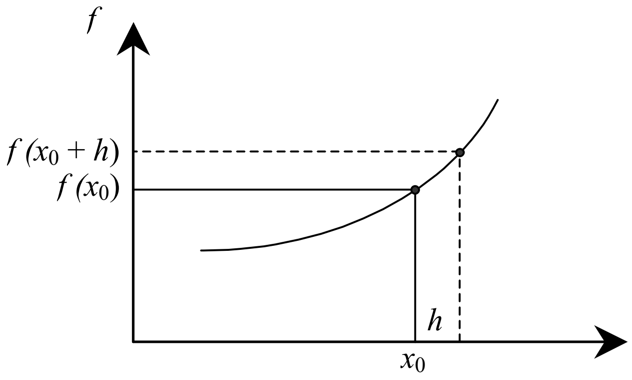
since $\Delta Q = E\frac{a}{2}ze$.

This means that the net jump rate varies linearly with the electric field.

Hence, the flux of ions and the electric current density both vary linearly with the applied field, and so using this simple model we have shown that ionically conducting solids exhibit Ohmic behaviour.

TAYLOR'S SERIES

Given the value of a function at some point, what is its value somewhere else?



$$f(x_0+h) = f(x_0) \qquad \text{first estimate}$$

$$+\,h\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} \qquad \text{to allow for slope}$$

$$+\frac{h^2}{2!}\left.\frac{\mathrm{d}^2f}{\mathrm{d}x^2}\right|_{x_0} \qquad \text{to allow for curvature}$$

$$\left[+\frac{h^3}{3!}\left.\frac{\mathrm{d}^3f}{\mathrm{d}x^3}\right|_{x_0} + \ldots.\right] \text{ assuming higher order terms can be ignored !!}$$

In many applications the "expansion" is about a minimum (e.g., an equilibrium position) or a maximum.

If so, the slope is zero, i.e., $\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} = 0$, and so the change in the function Δf is given by:

$$\Delta f = f(x_0+h) - f(x_0) \approx \frac{h^2}{2}\left.\frac{\mathrm{d}^2f}{\mathrm{d}x^2}\right|_{x_0}$$

and so the change is "*parabolic*" in h.

Remember: the expansion is carried out about a particular point (x_0) and the values for the differential coefficients *must be those corresponding to that point*.

BEWARE. Are the higher order terms really negligible? Take a good look!!

A good example of the application of Taylor's series is in the derivation of Fick's second law of Diffusion.

We shall keep the mathematics simple by deriving the 1D version of this law.

Suppose we have a concentration profile of atoms in 1D which varies with position and time, such as in the diffusion of carbon atoms into a piece of steel during a carburising operation.

From Fick's First Law,

$$J = -D\frac{\partial c}{\partial x}$$

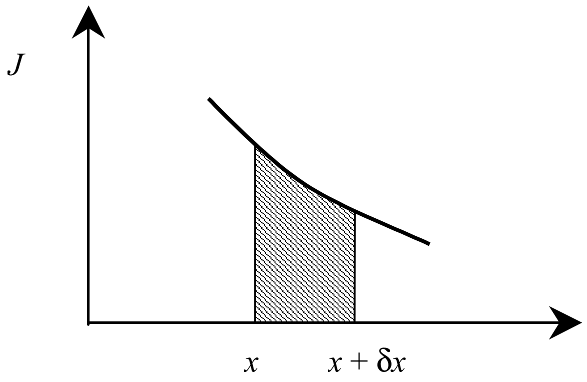
where J is the flux of atoms crossing unit area of interface in unit time, D is the diffusion coefficient for carbon atoms in steel and $\partial c/\partial x$ is the concentration gradient.

[Note that Crank in his book 'The Mathematics of Diffusion' states that the linear relationship between flux and concentration gradient is an hypothesis.

This contrasts with the approach often taken books in materials science in which atomic diffusion is modelled in terms of atoms jumping across adjacent atomic planes. From this, an estimate of the diffusion coefficient can be obtained in terms of jump frequency and the separation of atomic planes – see, for example, Porter and Easterling's book on 'Phase transformations in Metals and Alloys'. While it is implicit in this latter approach that second and higher order derivatives of concentration as a function of distance will be negligible, it is not often shown mathematically.]

The dimensions of J are $\mathrm{m}^{-2}\mathrm{s}^{-1}$, the dimensions of D are $\mathrm{m}^{2}\mathrm{s}^{-1}$ and the dimensions of $\partial c/\partial x$ are m^{-4}. Dimensional analysis is always useful to check you have written down equations correctly!

J will vary with distance, as in the schematic below:



Consider a cross section of area *A* perpendicular to the *x*-direction across which the atoms are jumping.

In time δt the net number of atoms diffusing into the hatched region between x and $x + \delta x$ from the left hand side = $J(x)\, A\, \delta t$.

In time δt the net number of atoms leaving the hatched region between x and $x + \delta x$ from the right hand side = $J(x + \delta x)\, A\, \delta t$.

Therefore the net concentration increase within the hatched area, δc, is given by

$$\delta c = \frac{\left(J(x) - J(x+\delta x)\right) A \delta t}{A \delta x}$$

Rearranging this,

$$\frac{\delta c}{\delta t} = \frac{J(x) - J(x+\delta x)}{\delta x}$$

Using the Taylor series expansion for $J(x + \delta x)$, we have

$$J(x+\delta x) = J(x) + \delta x \frac{\partial J}{\partial x} + \frac{(\delta x)^2}{2} \frac{\partial^2 J}{\partial x^2} + \ldots$$

and so

$$\frac{\delta c}{\delta t} = -\frac{\partial J}{\partial x} - \frac{\delta x}{2} \frac{\partial^2 J}{\partial x^2} - \ \ldots$$

so that as $\delta t \to 0$, we arrive at Fick's Second Law:

$$\frac{\partial c}{\partial t} = -\frac{\partial J}{\partial x}$$

Combining this with Fick's first law, we find in general that

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial c}{\partial x}\right)$$

If D is constant, we find that

$$\boxed{\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}}$$

Special cases where the Taylor series expansion is carried out about $x_0 = 0$ are called **Maclaurin series**.

Standard Maclaurin series expansions can be found in the texts recommended on the Book list. Some useful examples are:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \qquad \text{for all } x$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \qquad \text{for all } x$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \frac{x^6}{6} + \qquad \text{for } -1 < x \le 1$$

Other Maclaurin series expansions such as the one for sinh x quoted on page 6 can also be obtained straightforwardly. Remember that in the formulae for sin x and cos x that x is in **radians**.

Note that the Maclaurin series expansion for sinh x can also be derived from the one from sin x because

$$\sinh ix = i \sin x$$

and

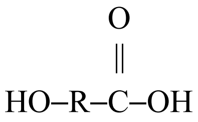
$$i \sinh x = \sin ix$$

STEPWISE POLYMERISATION

As an example of the use of geometric series in materials science to model physical phenomena mathematically, we shall investigate here the calculation of several characteristics of a polymer during stepwise polymerisation.

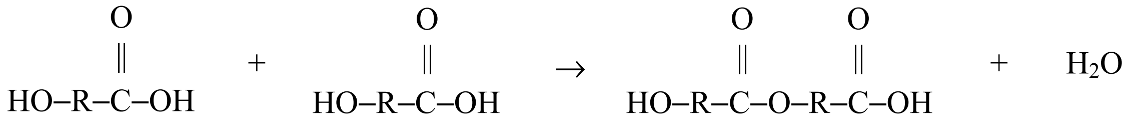
In stepwise polymerisation each polymer chain is just as reactive as the monomer.

For example, monomers of hydroxy acids of the general form



$$\mathrm{HO{-}R{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}OH}$$

can react with one another to produce a dimer with water of condensation:



$$\mathrm{HO{-}R{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}OH} \quad + \quad \mathrm{HO{-}R{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}OH} \quad \rightarrow \quad \mathrm{HO{-}R{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}O{-}R{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}OH} \quad + \quad H_2O$$

and it is apparent that the dimer is just as likely to react as the monomer. Such reactions can be very rapid. Stepwise polymerisation reactions also occur in the production of polyurethane foams.

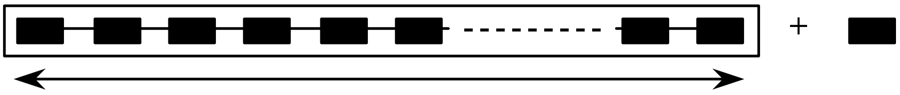
We shall look at four aspects of stepwise polymerisation:

1.	Mole fraction of x-mers	n_x
2.	Weight fraction of x-mers	w_x
3.	Number-average molecular weight	$\overline{M}_x$
4.	Weight-average molecular weight	$\overline{M}_w$

Consider a stepwise polymerisation process involving unit mers of mass M_0. Define p = probability of a given functional group having reacted in this stepwise polymerisation reaction, with $0 < p < 1$.

1. MOLE FRACTION of *x*-mers

$$n_x = \frac{\text{number of } x\text{ - mers}}{\text{Total number of molecules}}$$



$(x - 1)$ mers

To model stepwise polymerisation, imagine that in the above schematic to form an x-mer, an $(x - 1)$ mer reacts with a single mer. For the x-mer to remain an x-mer, both end mers must cease to react. In the x-mer, there are $(x - 1)$ links where reaction has taken place and two (the end mers) where reaction has not.

In practice, one of the two end mers of a polymer chain can be considered to be the ‘reacting’ end, so that in the above schematic it is the right hand side of the $(x–1)$ mer chain. If this reaction takes place, then to enable further polymerisation takes place, the ‘reacting’ end is again the new right hand side end.

From this logic it follows that the probability of forming an x-mer, P_x, is related to the probability of forming an $(x - 1)$ mer, P_{x-1}, through the formula

$$P_x = pP_{x-1}$$

The hydroxy acid mer is asymmetric, so without loss of generality, we can imagine the HO–R– part of this 1-mer to react with the –COOH part of the $(x - 1)$ mer to form the x-mer.

Now $P_1 = (1 - p)$ because it relates to the remaining 1-mer in the polymer soup after polymerisation has occurred which has not taken any part whatsoever in the polymerisation process. Hence, the probability of forming an x-mer, P_x, is

$$P_x = p^{x-1}(1 - p)$$

so that $P_1 = (1 - p)$, $P_2 = p(1 - p)$, etc. It follows that the probability of a single mer having reacted is the same as the probability of the given functional group reacting (which, in the way we have considered it here, is the HO–R– part).

Check: does $\sum_{x=1}^{\infty} P_x = 1$? (Sum of all probabilities must be unity.)

$$\sum_{x=1}^{\infty} P_x = \sum_{x=1}^{\infty} p^{x-1}(1 - p) = (1 - p)\sum_{x=1}^{\infty} p^{x-1}$$

Now,

$$\sum_{x=1}^{n} p^{x-1} = 1 + p + p^2 + p^3 + ... + p^{n-1} = \frac{1-p^n}{1-p}$$

and as $n \rightarrow \infty$, $p^n \rightarrow 0$, whence

$$\sum_{x=1}^{\infty} P_x = \sum_{x=1}^{\infty} p^{x-1}(1-p) = (1-p)\frac{1}{1-p} = 1, \text{as required.}$$

The formula quoted on the previous page for $\sum_{x=1}^{n} p^{x-1}$ can be checked using the principle of induction.

The result is true for $n = 1$ and $n = 2$ by inspection. Suppose the formula is also correct for a specific integer value of n, e.g. for $n = k$. For the formula to be of general applicability, it must also be true for $n = k + 1$.

Now, $\sum_{x=1}^{n+1} p^{x-1} = \left(\sum_{x=1}^{n} p^{x-1}\right) + p^n = \frac{1-p^n}{1-p} + p^n = \frac{1-p^n + p^n(1-p)}{1-p} = \frac{1-p^{n+1}}{1-p}$

and so if the formula is true for $n = k$, it is true for $n = k + 1$.

Hence, working forward from $n = 1$, it is apparent that the formula for $\sum_{x=1}^{n} p^{x-1}$ is correct.

From the above consideration of the probability of forming an x-mer in stepwise polymerisation, it is apparent that the mole fraction of x-mers, n_x, is the same as the probability of having an x-mer, so that

$$\boxed{n_x = P_x = p^{x-1}(1-p)}$$

2. WEIGHT FRACTION of *x*-mers

In this calculation we will make the approximation that the mass of an x-mer is simply x times the mass of a single mer. [This is an approximation because in polymerisation reactions there are products such as water which are produced, so that the mass of a dimer is slightly less than twice the mass of a single mer, and so on.]

With this approximation,

$$\begin{aligned} w_x &= \frac{\text{Total weight of } x\text{-mers}}{\text{Total weight of all mers}} \\ &= \frac{xM_0 \times \text{mole fraction of } x\text{-mers}}{\text{Total weight of all mers}} \\ &= \frac{xM_0 p^{x-1}(1-p)}{\sum_{x=1}^{\infty} xM_0 p^{x-1}(1-p)} \\ &= \frac{xp^{x-1}}{\sum_{x=1}^{\infty} xp^{x-1}} \end{aligned}$$

So to progress further, we need to evaluate $\sum_{x=1}^{\infty} xp^{x-1}$.

Now, $xp^{x-1} = \frac{\mathrm{d}}{\mathrm{d}p} p^x$, and we have already shown that $\sum_{x=1}^{\infty} p^{x-1} = \frac{1}{1-p}$.

Hence, $\sum_{x=1}^{\infty} p^x = p\sum_{x=1}^{\infty} p^{x-1} = \frac{p}{1-p}$, and so $\frac{\mathrm{d}}{\mathrm{d}p}\sum_{x=1}^{\infty} p^x = \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{p}{1-p}\right)$.

To evaluate $\frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{p}{1-p}\right)$:

Revision! $\frac{\mathrm{d}}{\mathrm{d}x}(uv) = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x}$

So with $u = p,\ v = \frac{1}{1-p}$,

we get: $\frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{p}{1-p}\right) = \left(\frac{1}{1-p}\right)\frac{\mathrm{d}}{\mathrm{d}p}(p) + p\frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{1}{1-p}\right) = \left(\frac{1}{1-p}\right) + p\left(+\frac{1}{(1-p)^2}\right)$

$$= \frac{1}{(1-p)^2}(1-p+p) = \frac{1}{(1-p)^2}$$

Therefore $\sum_{x=1}^{\infty} xp^{x-1} = \frac{1}{(1-p)^2}$ and so we find $\boxed{w_x = xp^{x-1}(1-p)^2}$

Check: does $\sum_{x=1}^{\infty} w_x = 1$? (Sum of all weight fractions must be unity.)

$$\text{Yes}: \sum_{x=1}^{\infty} w_x = \sum_{x=1}^{\infty} xp^{x-1}(1-p)^2 = (1-p)^2 \sum_{x=1}^{\infty} xp^{x-1} = (1-p)^2 \frac{1}{(1-p)^2} = 1$$

---

3. NUMBER-AVERAGE MOLECULAR WEIGHT

$$\overline{M_x} = \frac{\text{Mass of all molecules in a large sample}}{\text{Number of molecules, } N \text{, in the large sample}}$$

$$= \frac{\sum_{x=1}^{\infty} [(\text{Number of } x\text{-mers, i.e., } Nn_x) \times (\text{mass of one } x\text{-mer})]}{N}$$

$$= \frac{N\sum_{x=1}^{\infty} n_x x M_0}{N} = \frac{N\sum_{x=1}^{\infty} p^{x-1}(1-p)xM_0}{N}$$

making once again the approximation that the mass of an *x*-mer is simply *x* times the mass of a single mer. With this approximation,

$$\overline{M_x} = \frac{N\sum_{x=1}^{\infty} p^{x-1}(1-p)xM_0}{N} = (1-p)M_0 \sum_{x=1}^{\infty} xp^{x-1} = (1-p)M_0 \frac{1}{(1-p)^2} = \frac{M_0}{(1-p)}$$

i.e., the number-average molecular weight, $\overline{M_x}$, is given by

$$\boxed{\overline{M_x} = \frac{M_0}{(1-p)}}$$

This number-average molecular weight is obviously a useful quantity to measure for a polymer solution, but for physical properties such as viscosity, longer polymer chains have a greater influence than small chains, and it appropriate to consider instead the weight-average molecular weight, calculated in the next section.

4. WEIGHT-AVERAGE MOLECULAR WEIGHT

$$\overline{M_w} = \frac{\sum_{x=1}^{\infty}\left[(\text{Total mass of } x\text{ - mers in a large sample})\times(\text{mass of one } x\text{ - mer})\right]}{\text{Total mass of all molecules in the large sample}}$$

$$= \frac{\sum_{x=1}^{\infty}\left[(Nn_x xM_0)\times(xM_0)\right]}{\sum_{x=1}^{\infty} Nn_x xM_0} = \frac{\sum_{x=1}^{\infty} x^2 n_x}{\sum_{x=1}^{\infty} xn_x} M_0 = \frac{\sum_{x=1}^{\infty} x^2 p^{x-1}(1-p)}{\sum_{x=1}^{\infty} xp^{x-1}(1-p)} M_0$$

$$= \frac{\sum_{x=1}^{\infty} x^2 p^{x-1}(1-p)}{\sum_{x=1}^{\infty} xp^{x-1}(1-p)} M_0 = \frac{\sum_{x=1}^{\infty} x^2 p^{x-1}(1-p)}{\left((1-p)\big/(1-p)^2\right)} M_0 = (1-p)^2 M_0 \sum_{x=1}^{\infty} x^2 p^{x-1}$$

To evaluate $\sum_{x=1}^{\infty} x^2 p^{x-1}$ it is useful to note that $\sum_{x=1}^{\infty} x^2 p^{x-1} = \frac{\mathrm{d}}{\mathrm{d}p}\sum_{x=1}^{\infty} xp^x = \frac{\mathrm{d}}{\mathrm{d}p}\left(p\sum_{x=1}^{\infty} xp^{x-1}\right)$

Hence $\sum_{x=1}^{\infty} x^2 p^{x-1} = \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{p}{(1-p)^2}\right) = \frac{1}{(1-p)^2} + \frac{2p}{(1-p)^3} = \frac{1-p+2p}{(1-p)^3} = \frac{1+p}{(1-p)^3}$

Thus, we get:

$$\overline{M_w} = (1-p)^2 M_0 \frac{(1+p)}{(1-p)^3}$$

which simplifies to $\boxed{\overline{M_w} = \frac{(1+p)}{(1-p)} M_0}$

and so we find that

$$\boxed{\frac{\overline{M_w}}{\overline{M_x}} = 1+p}$$

The predicted ratio of 2 when p approaches unity is found experimentally for many stepwise polymerisations.

SIMPLE SECOND ORDER LINEAR DIFFERENTIAL EQUATIONS

The general second order linear differential equation takes the daunting form

$$a_0(x)\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + a_1(x)\frac{\mathrm{d}y}{\mathrm{d}x} + a_2(x)y = f(x)$$

but fortunately for many problems in materials science much simpler forms of this general equation are found to describe physical situations!

Particularly simple (and useful) forms occur when $a_0(x)$, $a_1(x)$ and $a_2(x)$ are constants and where $f(x)$ is either zero or a constant, as in the three examples below.

(1) Simple harmonic oscillator:

Here the differential equation is of the form

$$\boxed{a_0 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + a_2 y = 0}$$

where $a_0 > 0$ and $a_2 > 0$.

A solution of the form $y = \mathrm{e}^{imx}$ can be tried. Substituting this into the equation, we find that

$$-m^2 a_0 + a_2 = 0$$

and so $m^2 = (a_2/a_0)$. Hence the general solution is of the form

$$\boxed{y = A\mathrm{e}^{imx} + B\mathrm{e}^{-imx}} \quad \text{or} \quad \boxed{y = C\cos mx + D\sin mx}$$

where $m = \sqrt{a_2 / a_0}$.

Note that a common way of writing such differential equations where the differential is with respect to time is in the form

$$m\ddot{x} = -kx$$

(2) Damped simple harmonic oscillator:

Here the differential equation is of the form

$$\boxed{a_0 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_2 y = A\mathrm{e}^{i\omega x}}$$

where $a_0 > 0$, $a_1 > 0$, $a_2 > 0$ and A is also a positive constant.

A steady-state solution of the form $y = B\mathrm{e}^{i\omega x}$ can be tried. Note that when $A = 0$ other solutions exist for this equation (known as the 'complementary function') which die away as x increases.

For applications in materials science we are most likely to be interested in the steady-state solution of the differential equation when it takes the form

$$\ddot{x} + \gamma\dot{x} + \omega_0{}^2 x = A\mathrm{e}^{i\omega t}$$

Such a solution can be used to describe the variation of position coordinate x as a function of time, t, and hence the polarisation of a system of oscillators.

Resonance absorption in materials at optical frequencies is an example where this differential equation can be applied and used in materials science.

(3) Shear lag equation:

This equation is used to describe the response of a short fibre composite to an applied stress parallel to the length of the fibres.

Here the differential equation is of the form

$$\boxed{a_0 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - a_2 y = -A}$$

where $a_0 > 0$, $a_2 > 0$ and $A > 0$ is a constant. A solution of the form $y = B\mathrm{e}^{mx} + C$ can be tried where B and C are constants. There are two possible solutions for m, so that the general form of the solution is

$$\boxed{y = B_1\mathrm{e}^{mx} + B_2\mathrm{e}^{-mx} + C} \quad \text{or} \quad \boxed{y = B_3 \cosh mx + B_4 \sinh mx + C}$$

where $m = \sqrt{a_2 / a_0}$, the positive form of m.

FOURIER SERIES AND FOURIER TRANSFORMS

FOURIER SERIES

- Representation of periodic variations:

in space:	e.g.,	atomic potentials in crystals electron density in crystals
in time:	e.g.,	frequency analysis of vibrations

- Solutions of differential equations.

In this section we shall only consider the representation of periodic functions. The use of Fourier series in solutions of differential equations will be illustrated later in the course.

Representation of Periodic Functions

Suppose we have a periodic function $f(x)$, with period 2π (–π → +π). [Note that by simple substitution we can always make the period equal to 2π, see later.]

We seek to represent the function in terms of sines and cosines of period 2π/m (where the integer m = 1, 2, 3, ...), and a constant, a_0, thus:

$$\boxed{f(x) = \frac{1}{2}a_0 + \sum_{m=1}^{\infty}\left(a_m \cos mx + b_m \sin mx\right)}$$

To proceed further we need a few useful results:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \sin mx \cos nx \, \mathrm{d}x = 0$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \sin mx \sin nx \, \mathrm{d}x = \begin{cases} 0 & m \neq n \\ \frac{1}{2} & m = n \neq 0 \\ 0 & m = n = 0 \end{cases}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \cos mx \cos nx \, \mathrm{d}x = \begin{cases} 0 & m \neq n \\ \frac{1}{2} & m = n \neq 0 \\ 1 & m = n = 0 \end{cases}$$

These can be proved using the complex number representation of sine and cosine functions as demonstrated overleaf.

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\sin mx\cos nx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\mathrm{e}^{imx}-\mathrm{e}^{-imx}}{2i}\right)\left(\frac{\mathrm{e}^{inx}+\mathrm{e}^{-inx}}{2}\right)\mathrm{d}x$$

$$=\frac{1}{2\pi}\frac{1}{4i}\int_{-\pi}^{\pi}\left(\mathrm{e}^{i(m+n)x}+\mathrm{e}^{i(m-n)x}+\mathrm{e}^{i(-m+n)x}+\mathrm{e}^{-i(m+n)x}\right)\mathrm{d}x$$

All terms in this integral are of the form e^{ikx} where, for $m \neq n$, k is an non-zero integer.

Now, $\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{ikx}\,\mathrm{d}x=\frac{1}{2\pi}\left[\frac{\mathrm{e}^{ikx}}{ik}\right]_{-\pi}^{\pi}=\frac{1}{2\pi ik}\left(\mathrm{e}^{ik\pi}-\mathrm{e}^{-ik\pi}\right)=\frac{2i}{2\pi ik}\sin k\pi=\frac{\sin k\pi}{k\pi}=0$

since sin (integer π) = 0.

For $m = n \neq 0$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\sin mx\cos nx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\sin 2mx\,\mathrm{d}x=\frac{1}{4\pi}\left[-\frac{\cos 2mx}{2m}\right]_{-\pi}^{\pi}$$

$$=-\frac{1}{8\pi m}\left[\cos 2m\pi-\cos(-2m\pi)\right]=0$$

since $\cos 2m\pi = \cos(-2m\pi) = 1$.

For $m = n = 0$, $\sin mx = 0$, and so the integral over the product $\sin mx \cos nx$ must be zero. Hence we have shown that for **any** integers m and n,

$$\boxed{\frac{1}{2\pi}\int_{-\pi}^{\pi}\sin mx\cos nx\,\mathrm{d}x=0}$$

As a second example, we will look at the integral $\boxed{\frac{1}{2\pi}\int_{-\pi}^{\pi}\sin mx\sin nx\,\mathrm{d}x}$ for $m = n \neq 0$.

The integral is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\sin^2 mx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{(1-\cos 2mx)}{2}\mathrm{d}x\quad\text{using the result }\cos 2\theta=1-2\sin^2\theta$$

$$=\frac{1}{4\pi}\left[x-\frac{\sin 2mx}{2m}\right]_{-\pi}^{\pi}=\frac{1}{4\pi}\left[\pi-0-(-\pi-0)\right]=\frac{1}{2}$$

The other results can be obtained in a similar manner.

Returning to $f(x) = \frac{1}{2}a_0 + \sum_{m=1}^{\infty}(a_m \cos mx + b_m \sin mx)$; by multiplying both sides by cos *nx* or by sin *nx* and integrating from −π to +π and using the results we have just quoted, it follows that:

$$\boxed{\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x \quad (n = 0, 1, 2, \text{etc.}) \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,\mathrm{d}x \quad (n = 1, 2, \text{etc.}) \end{aligned}}$$

For example, suppose we multiply both sides by cos *nx* and integrate from −π to +π:

$$\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x = \int_{-\pi}^{\pi} \frac{1}{2}a_0 \cos nx\,\mathrm{d}x + \sum_{m=1}^{\infty}\int_{-\pi}^{\pi}(a_m \cos mx + b_m \sin mx)\cos nx\,\mathrm{d}x$$

The integral from −π to +π over cos *nx* is zero and the only non-zero term in the summation for $m = 1$ to ∞ is the a_m term where $m = n$.

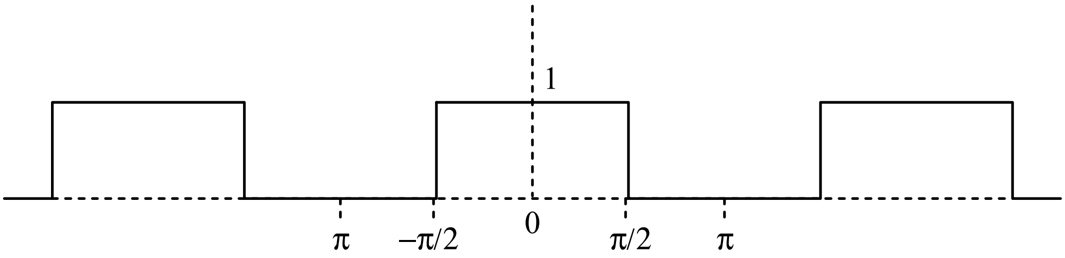
Hence, $\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x = \int_{-\pi}^{\pi} a_n \cos^2 nx\,\mathrm{d}x = \pi a_n$, and so the result quoted above follows.

Quite often, we can eliminate one or other set of a_n or b_n immediately because the function in which we are interested is either symmetrical or anti-symmetrical about the origin:

Even functions	**symmetrical about origin**	**all $b_n = 0$.**
Odd functions	**anti-symmetrical about origin**	**all $a_n = 0$.**

Simple Example of the Generation of Fourier Components

As a simple example, we will consider the periodic function $f(x)$ below which could represent (i) a series of slits of width π whose centres are spaced 2π apart or (ii) a series of simple potential wells of unit depth in a 1D lattice.



$f(x) = 1$ for $-\pi/2 \le x \le \pi/2$; $f(x) = 0$ for $-\pi \le x \le -\pi/2$ and $\pi/2 \le x \le \pi$

This is an even function (symmetrical about the origin), so that all $b_n = 0$. We need only determine a_0 and the a_n terms.

For $n = 0$, we have:

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,\mathrm{d}x = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} 1.\,\mathrm{d}x = \frac{1}{\pi}[x]_{-\pi/2}^{\pi/2} = 1$$

For $n \geq 0$, we have

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} 1.\cos nx\,\mathrm{d}x = \frac{1}{\pi}\left[\frac{\sin nx}{n}\right]_{-\pi/2}^{\pi/2} = \frac{1}{n\pi}\left(\sin\frac{n\pi}{2} - \sin\frac{-n\pi}{2}\right) = \frac{2}{n\pi}\sin\frac{n\pi}{2}$$

As n runs through the integers 1, 2, 3, etc., sin ($n\pi/2$) is zero for even n. For odd n sin ($n\pi/2$) takes the values 1 and –1 in alternation.

We seek to represent this alternation by $(-1)^p$ where $p = 0, 1, 2, 3$, etc.

This is achieved by setting $p = (n-1)/2$, as illustrated in the table below..

Values of n	1	3	5	7
Required value	+1	–1	+1	–1
Suitable p	0	1	2	3

Hence, we obtain

$$a_n = \frac{2}{n\pi}(-1)^{(n-1)/2} \quad \text{for } n \text{ odd}$$

$$= 0 \text{ for } n \text{ even } (> 0)$$

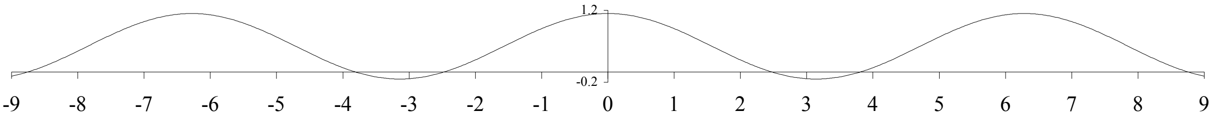
and so we can represent the periodic function $f(x)$ by the following Fourier series:

$$f(x) = \frac{1}{2}a_0 + \sum_{m=1}^{\infty}(a_m \cos mx + b_m \sin mx)$$

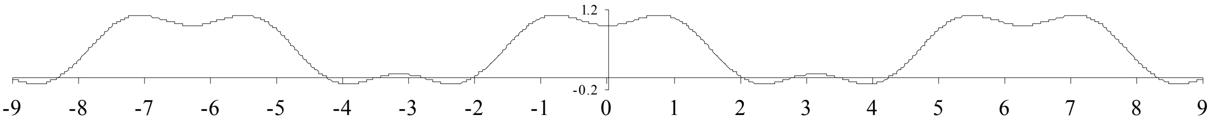
$$= \frac{1}{2} + \frac{2}{\pi}\cos x - \frac{2}{3\pi}\cos 3x + \frac{2}{5\pi}\cos 5x - \frac{2}{7\pi}\cos 7x + \ldots$$

The terms $\frac{2}{\pi}, -\frac{2}{3\pi}, \frac{2}{5\pi}$, etc. are known as the ‘Fourier coefficients’ in this representation of $f(x)$.

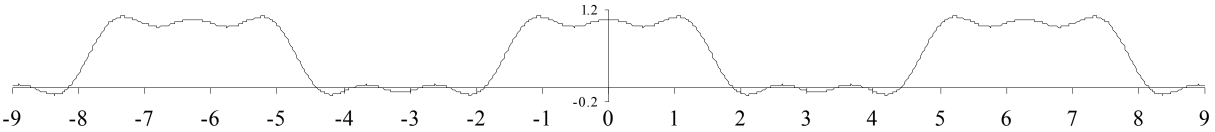
It is useful to look at how, as the number of terms in the Fourier series increases, the series converges to the example we have chosen:



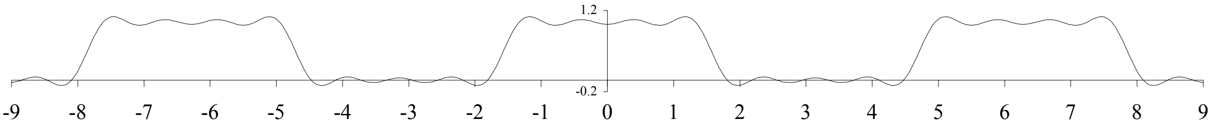
First two terms of the Fourier series



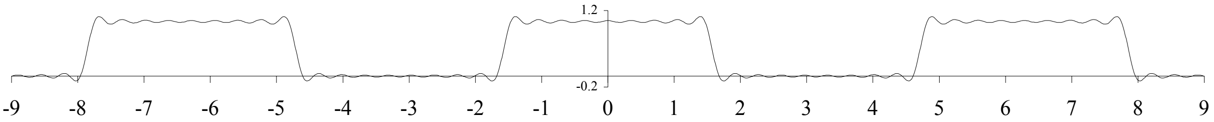
First three terms of the Fourier series



First four terms of the Fourier series



First five terms of the Fourier series



First ten terms of the Fourier series

Note the overshoot of the Fourier series at the values of $\pi/2$, $3\pi/2$, etc. This overshoot is known as Gibb's Phenomenon. In the limit of infinitely many terms in the Fourier series this overshoot sharpens to a spike with a magnitude of about 0.179 for the periodic top hat function shown here.

You will see a further example of the usefulness of Fourier coefficients in Part II in the Physical Properties course: the energy gaps at Brillouin zone boundaries are simply related to the Fourier coefficients representing the periodic potential in the crystal.

Use of complex exponentials

Instead of representing a periodic function $f(x)$ with period 2π (–π → +π). in terms of sines and cosines, we can choose to represent it in terms of complex exponentials.

In this representation,

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{inx}$$

and the Fourier coefficients c_n are then

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, \mathrm{e}^{-inx}\, \mathrm{d}x$$

General repeat periods

For a function of x', $f(x')$, with a periodic repeat of ℓ, it is convenient to use the substitution

$$\frac{2\pi x'}{\ell} = x$$

so that $f(x)$ is then periodic with period 2π (–π → +π).

FOURIER TRANSFORMS

Extending the ideas of Fourier series to general (not necessarily periodic) functions leads to the concept of Fourier transforms. These are particularly useful in handling diffraction phenomena in the Fraunhofer diffraction, which commonly applies to X-ray and electron diffraction. In Fraunhofer diffraction,

diffraction pattern = Fourier transform of the ‘object’

‘image’ = Fourier transform of the ‘diffraction pattern’

For a perfect lens, the image is an inverted (and magnified) version of the object. So, if we choose to ignore this trivial change in sign, we can also say that the object is also the Fourier transform of the diffraction pattern. Mathematically, if $f(x)$ represents the amplitude distribution in the object (or image) and $g(k)$ represents the amplitude distribution in the diffraction pattern, we have

$$\boxed{g(k)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)\,\mathrm{e}^{-ikx}dx} \quad \text{and} \quad \boxed{f(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(k)\,\mathrm{e}^{ikx}dk}$$

An example with which you should be familiar from Part IA Materials & Mineral Sciences arises is where $f(x)$ describes a single slit in real space. $g(k)$ is then the so-called ‘sinc’ function.

The analysis of diffraction experiments is often greatly simplified by use of the concept of convolution and the convolution theorem:

Convolution and the Convolution Theorem

The convolution, $f * g$ or $f \otimes g$, of two functions f and g is defined as:

$$f * g = \int_{0}^{t} f(\tau)\,g(t-\tau)\,\mathrm{d}\tau$$

The Fourier transform, of the convolution of two functions, ▯$(f * g)$, equals the product of their separate Fourier transforms:

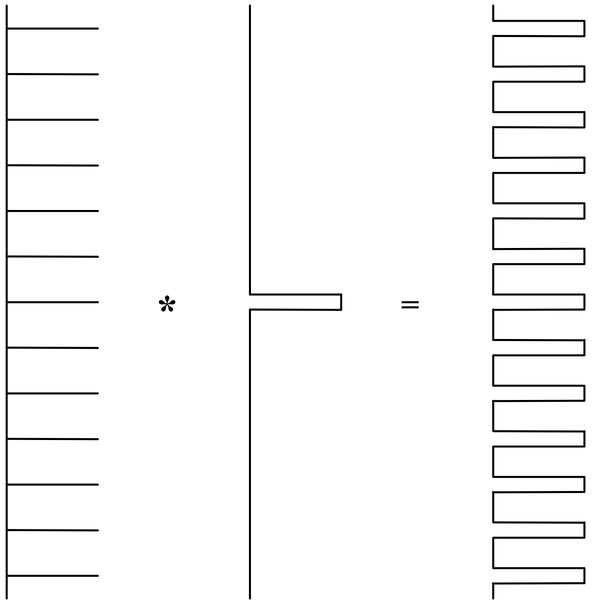
$$▯(f * g) = ▯(f)\,▯(g)$$

Similarly, the Fourier transform of the product of two functions f and g is equal to the convolution of the product of their separate Fourier transforms:

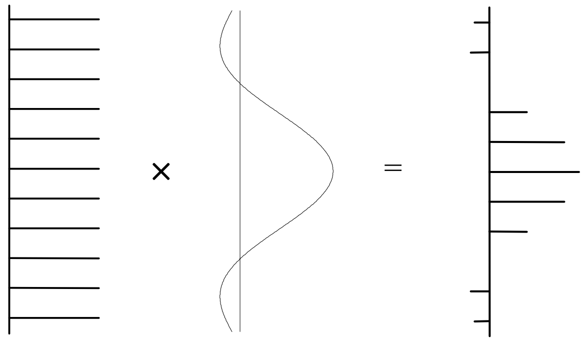
$$▯(fg) = ▯(f) * ▯(g)$$

More of this in Part III.

The amplitude distribution in the diffraction pattern of an infinite periodic array of delta functions is a periodic array of delta functions.



(a) Infinite grating of line sources
(b) Single-slit aperture function
(c) Infinite grating of slit sources

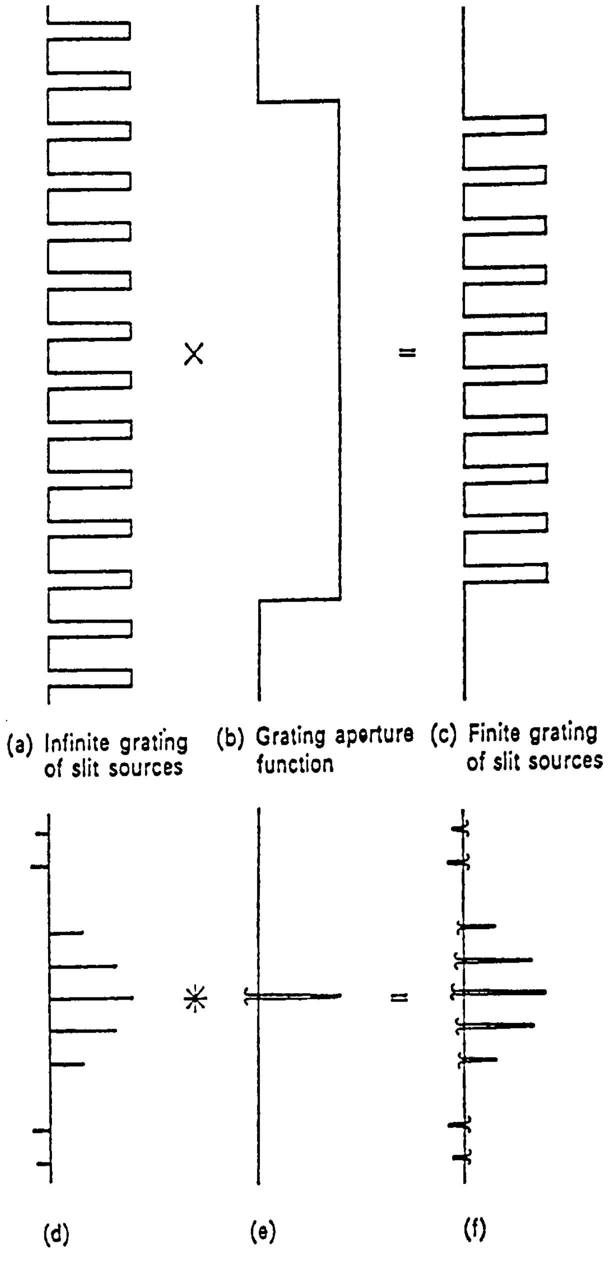


(d) Fourier Transform of (a)
(e) Fourier Transform of (b)
(f) Fourier Transform of (c)

As shown in (a) – (f) above, the amplitude distribution, (f), in the diffraction pattern of an infinite periodic array of narrow slits, (c), is the convolution of (i) the product of the Fourier transform of an infinite periodic array of delta functions and (ii) the Fourier transform of a single narrow slit, i.e., it is a periodic array of delta functions, the heights of which are modulated by a broad sinc function.

(d) is the Fourier transform of (a) and (e) is the Fourier transform of (b). Hence, (f), the amplitude distribution in the diffraction pattern of (c), is a convolution of (d) and (e) and is the product of (d) and (e).

The second example below shows the usefulness of the convolution theorem in explaining the effect of very small crystal sizes on the resultant diffraction pattern. Here, a small grating of thin slit sources (crystal) is obtained by multiplying a grating (a) by a 'top hat' function (b), producing the finite grating (c).



By the convolution theorem, (f), the amplitude distribution in the diffraction pattern of (c) is found by multiplying (d), the Fourier transform of (a), with (e), the Fourier transform of (b). The result is that diffraction peaks are broadened.

Remember that diffraction patterns, important to the electron microscopist and X-ray crystallographer, show intensities, not amplitudes.

MATRIX ALGEBRA

Consider first the properties of simple square matrices, *e.g.*, 3 × 3 ‘square’ matrices:

$$\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \mathbf{B}=\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

The product **C** = **AB** of these two matrices is then:

$$\mathbf{C}=\mathbf{AB}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}=\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$$

To calculate the elements of **C**, remember to go **<u>A</u>CROSS** and **<u>D</u>OWN** the i^{th} row and the k^{th} column.

An alternative mnemonic is ‘**RC**’ (e.g., **R**ow **C**olumn, **R**oman **C**atholic, **R**ed **C**ross, **R**einforced **C**oncrete) – the i^{th} **row** of matrix **A** is multiplied by the k^{th} **column** of matrix **B**.

Note that in general **AB ≠ BA**.

Formally:

$$c_{ik}=\sum_j a_{ij}b_{jk}=a_{ij}b_{jk} \text{ using the Einstein summation convention}$$

so that, for example,

$$c_{31}=a_{31}b_{11}+a_{32}b_{21}+a_{33}b_{31}$$

Note that the ‘outer’ indices remain fixed (here, 3 and 1 respectively), while the ‘inner’ ones run through 11, 22 and 33.

Similarly, $$c_{12}=a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32}$$

det **A** = ***Determinant*** of **A**

$$\det\mathbf{A}=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=a_{11}\left(a_{22}a_{33}-a_{32}a_{23}\right)-a_{12}\left(a_{21}a_{33}-a_{31}a_{23}\right)+a_{13}\left(a_{21}a_{32}-a_{31}a_{22}\right)$$

$\mathbf{M}_{ij}$ = ***Minor*** of a_{ij}

The minor is the determinant of the (smaller) matrix formed by eliminating the i^{th} row and j^{th} column of **A**.

Thus, $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ has the minor matrix $\mathbf{M}_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$ for example

$\mathbf{C}_{ij}$ = ***Cofactor*** of a_{ij}

$$\mathbf{C}_{ij} = (-1)^{i+j}\,\mathbf{M}_{ij}$$

The array of signs in this relationship can be represented simply by:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

and so for a 3 × 3 'square' matrix,

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = a_{1i}C_{1i} \text{ using the Einstein summation convention}$$

Transpose, $\mathbf{A}^{\mathrm{t}}$ (or $\mathbf{A}^{\mathrm{T}}$ or $\widetilde{\mathbf{A}}$ or $\mathbf{A}'$)

The transpose is obtained by interchanging rows and columns: $a_{ij}^{\ \mathbf{t}} = a_{ji}$

Complex Conjugate, **A***

$$a_{ij}^* = (a_{ji})^*$$

Unit Matrix, **I** (or **1** or **U** or **E**)

$$a_{ij} = \delta_{ij} \quad \text{(Kronecker } \delta\text{)}$$

e.g., for the unit 3 × 3 matrix:

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

Symmetric Matrix

$$\mathbf{A} = \mathbf{A}^{\mathbf{t}}$$

$$a_{ij} = a_{ji}$$

Antisymmetric Matrix

$$\mathbf{A} = -\,\mathbf{A}^{\mathbf{t}}$$

$a_{ij} = -\,a_{ji}$ (so that it follows that $a_{ii} = 0$)

Inverse Matrix, $\mathbf{A}^{-1}$

$$\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{A}^{-1}\,\mathbf{A} = \mathbf{I}$$

$\mathbf{A}^{-1} = \dfrac{1}{\det \mathbf{A}}\mathbf{C}^{\mathbf{t}}$, where $\mathbf{C}_{ij}$ is the cofactor of a_{ij} (see above).

Orthogonal Matrix, $\mathbf{A}$

Such matrices are defined by the two relationships

$$\mathbf{A}\,\mathbf{A}^{\mathbf{t}} = \mathbf{I}$$

$$\mathbf{A}^{\mathbf{t}} = \mathbf{A}^{-1}$$

Trace of a Matrix, Tr ($\mathbf{A}$)

$$\mathrm{Tr}\,(\mathbf{A}) = \sum_i a_{ii}$$

TWO USEFUL RESULTS:

1. $(\mathbf{AB})^{\mathbf{t}} = \mathbf{B}^{\mathbf{t}}\,\mathbf{A}^{\mathbf{t}}$

   Proof: $(\mathbf{AB})_{ik}^{\ \mathbf{t}} = (\mathbf{AB})_{ki} = \sum_j a_{kj} b_{ji} = \sum_j a_{jk}^{\ \mathbf{t}} b_{ij}^{\ \mathbf{t}} = \sum_j \ b_{ij}^{\ \mathbf{t}} a_{jk}^{\ \mathbf{t}} = (\mathbf{B}^{\mathbf{t}}\mathbf{A}^{\mathbf{t}})_{ik}$

2. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$

   Proof: first note that by definition $(\mathbf{AB})^{-1}\,(\mathbf{AB}) = \mathbf{I}$

   Now, $\mathbf{B}^{-1}\,\mathbf{A}^{-1}\,(\mathbf{AB}) = \mathbf{B}^{-1}\,(\mathbf{A}^{-1}\mathbf{A})\,\mathbf{B} = \mathbf{B}^{-1}\,\mathbf{B} = \mathbf{I}$

   Hence it follows that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$.

APPLICATION: COORDINATE TRANSFORMATIONS

EIGENVALUES, EIGENVECTORS, AND DIAGONALISATION

(Ref: [*e.g.*] Boas – Chapter 10)

Consider the linear transformations

$$X = ax + by$$
$$Y = cx + dy$$

In matrix form these can be written

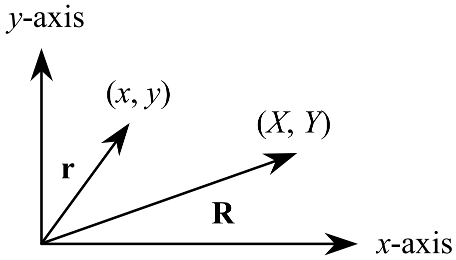
$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \text{ or } \mathbf{R} = \mathbf{M}\mathbf{r}$$

where

$$\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ \mathbf{R} = \begin{pmatrix} X \\ Y \end{pmatrix} \text{ and } \mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Same axes, change vector

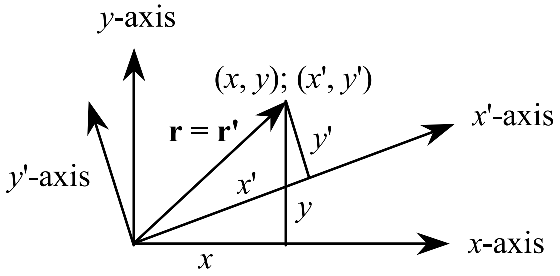
Old vector **r**, new vector **R**



$$\mathbf{R} = \mathbf{M}\mathbf{r}$$

Same vector, change axes

Old axes x, y ; new axes x', y'



$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

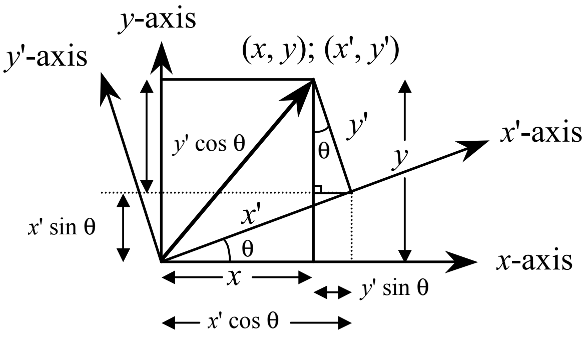
which is of the same form, i.e.,

$$\mathbf{r} = \mathbf{C}\mathbf{r}'$$

with

$$\mathbf{C} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ \mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix} \text{ and } \mathbf{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

Rotation of Axes



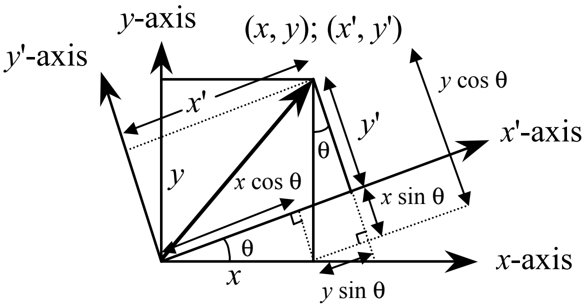
‘Forward’

$$x = x' \cos\theta - y' \sin\theta$$

$$y = x' \sin\theta + y' \cos\theta$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{C} \begin{pmatrix} x' \\ y' \end{pmatrix}$$



‘Reverse’

$$x' = x \cos\theta + y \sin\theta$$

$$y' = -x \sin\theta + y \cos\theta$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

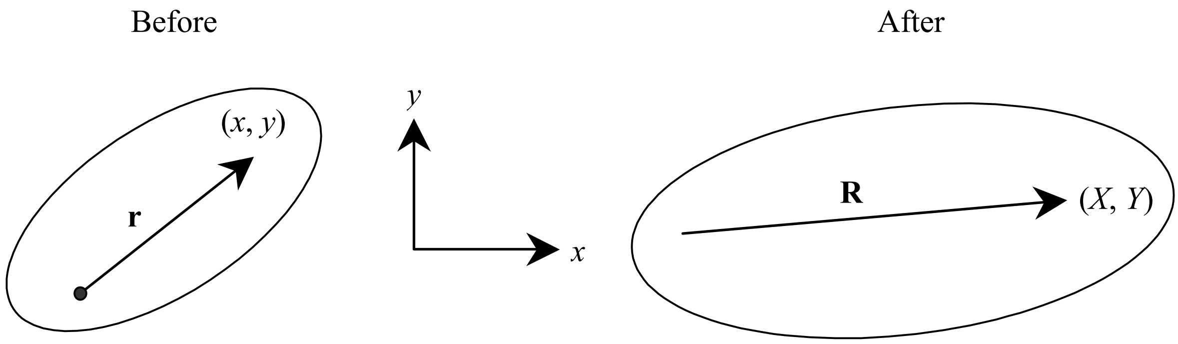
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{C}^{\mathbf{t}} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{C} \begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{C}\mathbf{C}^{\mathbf{t}} \begin{pmatrix} x \\ y \end{pmatrix}$$

The matrix **C** has the property that $\mathbf{C}\ \mathbf{C}^{\mathbf{t}} = \mathbf{I}$, so that $\mathbf{C}^{\mathbf{t}} = \mathbf{C}^{-1}$ and so this is an example of an ‘orthogonal’ transformation. Furthermore, it follows that

$$x^2 + y^2 = x'^2 + y'^2 \text{ and } \det \mathbf{C} = 1$$

EXAMPLE: A DEFORMED (PLANE) SHEET



Suppose

$$\mathbf{R} = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{Mr}$$

so that the matrix **M** describing the strain associated with the two-dimensional deformation is symmetric (see, for example, C4 for why this matrix has to be symmetric). Are there any vectors in the plane which have **not** changed in direction?

If so, they will be given by:

$$\mathbf{R} = \lambda\,\mathbf{r} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

so it follows that they satisfy the equations

$$\begin{aligned} 5x - 2y &= \lambda x \\ -2x + 2y &= \lambda y \end{aligned}$$

$$\begin{aligned} (5-\lambda)x - 2y &= 0 \\ -2x + (2-\lambda)y &= 0 \end{aligned}$$

$$\begin{pmatrix} (5-\lambda) & -2 \\ -2 & (2-\lambda) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

Either $x = y = 0$ (a solution we can ignore), or the '***characteristic equation***' must be satisfied:

$$\begin{vmatrix} (5-\lambda) & -2 \\ -2 & (2-\lambda) \end{vmatrix} = 0$$

Solving this, it follows that

$$\begin{aligned} (5-\lambda)(2-\lambda) - 4 &= 0 \\ \lambda^2 - 7\lambda + 6 &= 0 \\ (\lambda - 1)(\lambda - 6) &= 0 \end{aligned}$$

Thus the '***eigenvalues***' are $\lambda = 1$ and $\lambda = 6$.

Substituting back in:

$$\begin{pmatrix} (5-\lambda) & -2 \\ -2 & (2-\lambda) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

gives

$(\lambda = 1)$

$$\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

So **R** = **r** when $2x - y = 0$

$(\lambda = 6)$

$$\begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

So **R** = 6**r** when $x + 2y = 0$

The lines defined by $2x - y = 0$ and and $x + 2y = 0$ are known as the '***eigenvectors***'. Note that these lines are '*orthogonal*'. The product of their slopes (2 and $-^1/2$) equals – 1.

If we choose axes parallel to the eigenvectors we can get a much simpler description of the deformation.

Returning to:

$$\mathbf{R} = \mathbf{Mr} = \lambda\mathbf{r}$$

we can define eigenvectors of unit length associated with the two eigenvalues λ = 1 and λ = 6.

$(\lambda = 1)$

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}$$

$(\lambda = 6)$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$$

where x_1, y_1 satisfy the eigenvector equation $2x - y = 0$ for $\lambda = 1$ and x_2, y_2 satisfy the eigenvector equation $x + 2y = 0$ for $\lambda = 6$.

It follows from the definition of the eigenvalues and the eigenvectors that for this matrix **M**,

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$$

Note that the second matrix on the right-hand side (the one containing the eigenvalues) is 'diagonal'.

Substituting the values for x_1, y_1 and x_2, y_2 explicitly into this equation above gives:

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$$

which may be written in the form

$$\mathbf{MC} = \mathbf{CD}$$

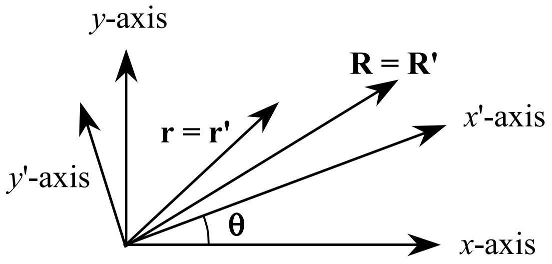
In which the columns of **C** are the components of the unit eigenvectors.

Multiplying each side by $\mathbf{C}^{-1}$ and simplifying, we get:

$$\mathbf{C}^{-1}\mathbf{MC} = \mathbf{C}^{-1}\mathbf{CD} = \mathbf{D}$$

D is '*similar*' to **M**. In this special case, **M** has been '*diagonalised*' by a '*similarity transformation*' using matrix **C** and its inverse.

MATRIX DIAGONALISATION IN TERMS OF ROTATION OF AXES



The **Rotation** of axes changes **r'** and **R'**: $\mathbf{r} = \mathbf{Cr'}$ and $\mathbf{R} = \mathbf{CR'}$, and so $\mathbf{R'} = \mathbf{C}^{-1}\mathbf{R}$

In terms of the two axis systems, the deformation is described in the following equivalent ways:

(i) With respect to the x, y axes: $\mathbf{R} = \mathbf{Mr}$

(ii) With respect to the x', y' axes: $\mathbf{R'} = \mathbf{C}^{-1}\mathbf{R} = \mathbf{C}^{-1}\mathbf{Mr} = \mathbf{C}^{-1}\mathbf{MCr'} = \mathbf{Dr'}$

The matrix $\mathbf{D} = \mathbf{C}^{-1}\mathbf{MC}$ is the matrix which describes in the (x', y') system the ***same deformation*** that **M** describes in the (x, y) system.

The matrix **C** which diagonalises **M** is the rotation matrix describing the rotation from the (x, y) axis system to the (x', y') axis system: the system in which the x', y' axes are along the eigenvectors of **M**.

Relative to axes along the eigenvectors

$$\mathbf{R'} = \mathbf{Dr'}$$

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

Hence, (i) $X' = x'$ and there is no change in length along this direction, and (ii) $Y' = 6y'$ and there is a six-fold change in length along this direction. In neither case is there a rotation of course.

More generally, we can invoke rotations which do not generate a diagonalised matrix describing the deformation in the new axis system. However, it might still be useful to describe a deformation in a new axis system by imposing a rotation between 'old' and 'new' axes.

You will see an example of this in C16 in which a 2D stress matrix describing deformation in the plane of a unidirectional fibre-reinforced lamina is redescribed in an axis system in which one axis is parallel to the fibres and the other perpendicular to the fibre direction in the plane of the lamina. The stress matrix in the 'old' axis system is related to the stress matrix in the 'new' axis system by a similarity transformation.

Finally, there are two useful results which follow from the mathematics of similarity transformations between two symmetric matrices **M** and **M'**:

(1) The determinants of **M** and **M'** are the same.

(2) The traces of **M** and **M'** are the same.

[So in the example that we have worked through here, det **M** = det **D** = 6, and Tr (**M**) = Tr (**D**) = 7. Another way of expressing (1) matrices is that the product of the eigenvalues of a symmetric matrix is a constant and is necessarily unaffected by a similarity transformation. Likewise (2) follows from the fact that the sum of the eigenvalues is also a constant and is necessarily unaffected by a similarity transformation.]

Note:

1. If **M** is *not* symmetric, the eigenvectors are not orthogonal and **C** is not orthogonal.
2. The conclusions of this 2D example apply in 3 dimensions (and higher dimensions).

Worked example:

Suppose we have a real symmetric tensor describing the stress applied in the plane of a sheet of a unidirectionally fibre-reinforced plastic:

Most generally the stress will have shear stress and tensile stress components and be of the form

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_x & \tau_{xy} \\ \tau_{xy} & \sigma_y \end{pmatrix}$$

(see C4: Tensor Properties).

It is useful in the theory describing the effect of such stresses on the strain characteristics of composite laminae to transform such a stress tensor to an axis system parallel and perpendicular to the axis of the fibres.

In terms of the diagrams on page 33, we can imagine the 'new' axis system to be the axes *x*', *y*' system and the 'old' axis system to be the *x, y* axis system.

Thus, for an anti-clockwise rotation of θ, we have from pages 33 and 36,

$$\mathbf{C} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

and the stress tensor referred to the 'new' axis system is then

$$\boldsymbol{\sigma}' = \begin{pmatrix} \sigma_1 & \tau_{12} \\ \tau_{12} & \sigma_2 \end{pmatrix} = \mathbf{C}^{-1}\boldsymbol{\sigma}\mathbf{C} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} \sigma_x & \tau_{xy} \\ \tau_{xy} & \sigma_y \end{pmatrix}\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

$$= \begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} c\sigma_x + s\tau_{xy} & -s\sigma_x + c\tau_{xy} \\ c\tau_{xy} + s\sigma_y & -s\tau_{xy} + c\sigma_y \end{pmatrix}$$

$$= \begin{pmatrix} c^2\sigma_x + s^2\sigma_y + 2sc\tau_{xy} & -sc\sigma_x + sc\sigma_y + \left(c^2 - s^2\right)\tau_{xy} \\ -sc\sigma_x + sc\sigma_y + \left(c^2 - s^2\right)\tau_{xy} & s^2\sigma_x + c^2\sigma_y - 2sc\tau_{xy} \end{pmatrix}$$

It is convenient to express the relationship between σ_1, σ_2, τ_{12}, σ_x, σ_y and τ_{xy} in the form

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{pmatrix} = \begin{pmatrix} c^2 & s^2 & 2sc \\ s^2 & c^2 & -2sc \\ -sc & sc & c^2 - s^2 \end{pmatrix}\begin{pmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{pmatrix}$$

as you will see in C16: Composite Materials.

FIRST ORDER ORDINARY DIFFERENTIAL EQUATIONS

$\frac{\mathrm{d}y}{\mathrm{d}x}$ (or equivalent) appears to the ***first*** *power*; and there are no higher differential coefficients.

Two main types: ***separable*** and ***linear***.

1. SEPARABLE

$$\frac{\mathrm{d}y}{\mathrm{d}x}=f(y)g(x)$$

$$\frac{\mathrm{d}y}{f(y)}=g(x)\,\mathrm{d}x$$

The case $f(y)$ = constant can be regarded as a special (simple) case of this type of equation.

2. LINEAR

$$\frac{\mathrm{d}y}{\mathrm{d}x}+P(x)y=Q(x)$$

To solve this, we search for a function $T(x)$ known as an 'integrating factor' to multiply both the left hand side of the equation and the right hand side of the equation, so that we have an equation of the form

$$\frac{\mathrm{d}}{\mathrm{d}x}(yT(x))=Q(x)T(x)$$

which must be valid for any $Q(x)$.

The case

$$\frac{\mathrm{d}}{\mathrm{d}x}(yT(x))=0$$

or, expanding the left hand side,

$$T(x)\frac{\mathrm{d}y}{\mathrm{d}x}+y\frac{\mathrm{d}T(x)}{\mathrm{d}x}=0$$

is a special case of the more general equation involving $Q(x)$ and can be compared with the special case of the original equation where $Q(x) = 0$:

$$\frac{\mathrm{d}y}{\mathrm{d}x}+P(x)y=0$$

Dividing through the first of these equations by $T(x)$ and equating coefficients of y, we find

$$P(x) = \frac{1}{T(x)}\frac{\mathrm{d}T(x)}{\mathrm{d}x}$$

is the equation from which we can determine $T(x)$.

Integrating this equation, we find

$$\int P(x)\,\mathrm{d}x = \int \frac{\mathrm{d}T(x)}{T(x)}$$

and so

$$\ln T(x) = \int P(x)\,\mathrm{d}x + c$$

where c is a constant of integration. Hence

$$T(x) = A\mathrm{e}^{\int P(x)\,\mathrm{d}x}$$

where $A = \mathrm{e}^c$ and is a constant. A can be taken to be unity for simplicity.

Hence the ‘integrating factor’ needed to solve our original equation is

$$\boxed{T(x) = \mathrm{e}^{\int P(x)\,\mathrm{d}x}}$$

and the original equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)$$

becomes

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\left(y\mathrm{e}^{\int P(x)\,\mathrm{d}x}\right) = Q(x)\mathrm{e}^{\int P(x)\,\mathrm{d}x}}$$

from which it follows after integration that the solution for y from the original equation is

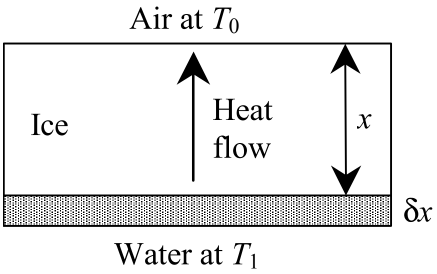
$$\boxed{y = \mathrm{e}^{-\int P(x)\,\mathrm{d}x}\int Q(x)\,\mathrm{e}^{\int P(x)\,\mathrm{d}x}\,\mathrm{d}x + C\mathrm{e}^{-\int P(x)\,\mathrm{d}x}}$$

for some constant of integration C. The *boundary* or *initial* conditions will determine the relevant value of C.

EXAMPLES OF THE SOLUTION OF FIRST ORDER DIFFERENTIAL EQUATIONS

1. Separable variables – a simple solidification model

(e.g., the growth of ice between air at fixed temperature T_0 and water at fixed temperature T_1)



Let Q be the heat evolved and t be time. The heat flow per unit area is then

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = k\left(\frac{T_1 - T_0}{x}\right)$$

(Fick's first law) where k is thermal conductivity. Growth of ice occurs at the ice/water interface.

The heat flow must balance the evolution of latent heat per unit area at the ice/water interface. In time δt the amount of heat evolved, δQ, is given by

$$\delta Q = \Delta H_f \, \delta x$$

where ΔH_f is the latent heat of fusion and δx is the increase in thickness of the solid in time δt. Hence, dividing this equation by δt and taking limits as δt tends to zero, we have the equation

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \Delta H_f \frac{\mathrm{d}x}{\mathrm{d}t}$$

arising from the condition that no heat is lost from the system.

Therefore,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = k\left(\frac{T_1 - T_0}{x}\right) = \Delta H_f \frac{\mathrm{d}x}{\mathrm{d}t}$$

and so we can generate a first order linear differential equation for x as a function of time:

$$k\left(\frac{T_1 - T_0}{x}\right) = \Delta H_f \frac{\mathrm{d}x}{\mathrm{d}t}$$

This equation is separable:

$$\frac{k(T_1 - T_0)}{\Delta H_f}\,\mathrm{d}t = x\,\mathrm{d}x$$

and integrates simply to

$$\frac{x^2}{2} = \frac{k(T_1 - T_0)}{\Delta H_f}t + \text{constant}$$

With the initial condition $x = 0$ at $t = 0$, the constant must be zero, giving the final result

$$x = \sqrt{\frac{2k(T_1 - T_0)}{\Delta H_f}t}$$

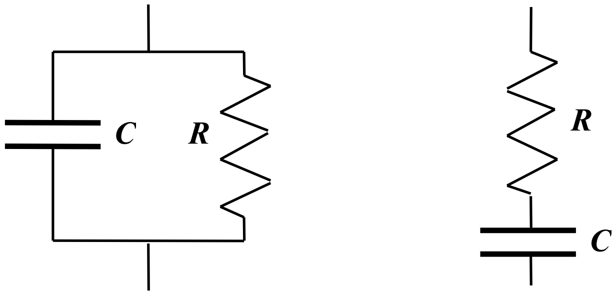
i.e. 'parabolic' growth, in which x is proportional to the square root of time.

Such growth is observed in oxidation and precipitation in the solid state too, as you have seen in Part IB. In these cases the relevant conservation equation is related to the conservation of atoms.

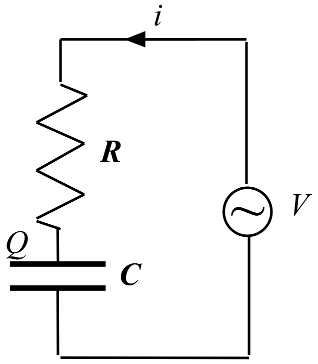
2. Integrating factor – 'Lossy' Dielectrics and the use of the *integrating factor*

Lossy dielectrics are imperfect (real world!) insulators in which heat is dissipated when they are being used. In terms of electric theory they can be considered as black boxes within which there are perfect capacitors and perfect resistors which together describe the behaviour of the dielectric as a function of applied frequency of the electric field.

The most simple representations are the two shown below, one in which there is a capacitor and resistor in parallel and one where these two elements are in series.



Here, we shall study the simple series arrangement in which the lossy dielectric is connected to an a.c. voltage supply:



Let Q be the charge on the capacitor at time t. The driving potential V is balanced by the potentials across the resistor and the capacitor. Hence,

$$Ri + \frac{Q}{C} = V$$

where $i = \frac{\mathrm{d}Q}{\mathrm{d}t}$ is the current flowing through the resistor at time t and V is an a.c. voltage.

Suppose

$$V = V_0 \cos \omega t = \mathrm{Re}\left(V_0 e^{j\omega t}\right)$$

where $j = \sqrt{-1}$ since we are using i to describe current (remember never to mix symbols so that they have two different meanings!).

We can choose to work with either the real form of V or the complex form. Since it is simple to work with the complex form of V, we shall do so.

Substituting for V and i, and dividing through by R, the differential equation becomes

$$\frac{dQ}{dt} + \frac{Q}{RC} = \frac{V_0}{R} e^{j\omega t}$$

and so in this case the integrating factor is simply $\mathrm{e}^{\int (1/RC)\mathrm{d}t} = \mathrm{e}^{t/RC}$.

Hence, using the general solution of a linear first order differential equation, we have

$$\boxed{Q = \mathrm{e}^{-t/RC} \int \mathrm{e}^{t/RC} \frac{V_0}{R} \mathrm{e}^{j\omega t} \mathrm{d}t + B\mathrm{e}^{-t/RC}}$$

where we have used B as our constant of integration because C is being used for capacitance.

Since we have used the complex form of V, the integral can be integrated easily:

$$Q = \frac{V_0}{R}\mathrm{e}^{-t/RC}\left(\frac{\mathrm{e}^{t/RC+j\omega t}}{\frac{1}{RC}+j\omega}\right) + B\mathrm{e}^{-t/RC} = \frac{V_0}{R}\frac{\mathrm{e}^{j\omega t}}{\frac{1}{RC}+j\omega} + B\mathrm{e}^{-t/RC} = CV_0\frac{\mathrm{e}^{j\omega t}}{1+j\omega RC} + B\mathrm{e}^{-t/RC}$$

and so the complex form of the current is simply

$$i = \frac{\mathrm{d}Q}{\mathrm{d}t} = j\omega CV_0\frac{\mathrm{e}^{j\omega t}}{1+j\omega RC} + B\mathrm{e}^{-t/RC}$$

This current is composed of two terms: an oscillatory term and a term which decays away exponentially to zero as $t \to \infty$.

Keeping the oscillatory term and ignoring the exponentially decaying term, we have

$$i = j\omega CV_0\frac{\mathrm{e}^{j\omega t}}{1+j\omega RC} = \frac{1}{\frac{1}{j\omega C}+R}V_0\mathrm{e}^{j\omega t} = \frac{V_0\mathrm{e}^{j\omega t}}{Z}$$

where Z is the known as the impedance of the circuit (which can be regarded as a complex resistance). It follows that Z is related to R and C through the formula

$$Z = \frac{1}{j\omega C} + R$$

and so the impedance of the resistor is R and the impedance of the capacitor is $\frac{1}{j\omega C}$.

A similar consideration for a simple electrical circuit in which R and C are in parallel gives the same result for the impedances of R and C in a parallel circuit. Thus, the formulae above for the impedance of R and C are general results.

The real part of the oscillatory term of the current is of interest for lossy dielectrics. This can be found straightforwardly by separating the oscillatory term of the complex current into real and imaginary parts:

$$i = j\omega CV_0 \frac{e^{j\omega t}}{1+j\omega RC} = j\omega CV_0\left(\frac{1-j\omega RC}{(1+j\omega RC)(1-j\omega RC)}\right)e^{j\omega t} = j\omega CV_0 e^{j\omega t}\left(\frac{1-j\omega RC}{1+\omega^2 R^2 C^2}\right)$$

$$= j\omega CV_0(\cos\omega t + j\sin\omega t)\left(\frac{1-j\omega RC}{1+\omega^2 R^2 C^2}\right)$$

The real part of this term is:

$$i = \omega CV_0\left(\frac{\omega RC\cos\omega t - \sin\omega t}{1+\omega^2 R^2 C^2}\right)$$

Does this seem a reasonable result?

We can check that the result is valid for a situation with which we are familiar. For a perfect dielectric (when $R = 0$), the formula predicts

$$i = -\omega CV_0 \sin\omega t$$

which is exactly what we would expect from the form of V. Alternatively, if $C \to \infty$, the current tends towards the value

$$i = \frac{V_0}{R}\cos\omega t$$

which is what we would expect for a resistor where the voltage drop across any capacitive component is zero.

It is always useful to check that your result is valid for a situation with which you are familiar. If it isn’t valid, you’ve probably made a mistake in your mathematics.

3. Viscoelasticity

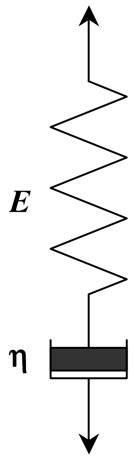
The response of polymers to an applied stress or strain depends on the rate or time period of loading. Simple mechanical models can be used to account for the dependence of the stress or strain on time. These models combine **springs** exhibiting linear elastic behaviour:

$$\sigma = E\varepsilon$$

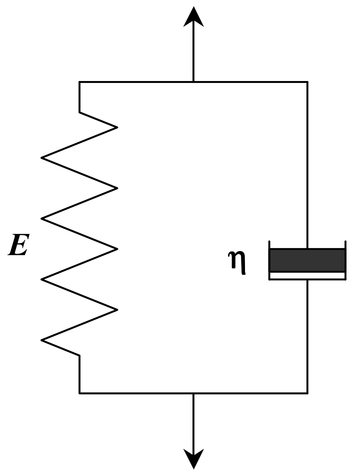
and **dashpots** exhibiting linear viscous behaviour

$$\sigma = \eta \frac{\mathrm{d}\varepsilon}{\mathrm{d}t}$$

Two particular models, the Maxwell model, (*a*), and the Voigt model, (*b*), are shown below.



(*a*): the Maxwell model



(*b*): the Voigt model

Let the stress in the spring be σ_1 and the strain be ε_1. Let the stress in the dashpot be σ_2 and the strain be ε_2.

In the Maxwell model, the stress on the spring and the stress on the dashpot are the same. Hence,

$$\sigma = E\varepsilon_1 = \eta \frac{\mathrm{d}\varepsilon_2}{\mathrm{d}t}$$

where σ is the overall stress. The overall strain, ε, is $\varepsilon_1 + \varepsilon_2$. Hence, the rate of change of strain, $\mathrm{d}\varepsilon/\mathrm{d}t$, is defined by the equation

$$\frac{\sigma}{\eta} + \frac{1}{E}\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{\mathrm{d}\varepsilon}{\mathrm{d}t}$$

where σ is the overall stress and ε is the overall strain. This Maxwell model is a useful model to describe the phenomenon of stress relaxation of a polymer.

In the Voigt model the spring and the dashpot have the same overall strain, ε. The overall stress is the sum of the individual stresses on the spring and the dashpot. Hence, in the Voigt model, we have

$$\eta \frac{\mathrm{d}\varepsilon}{\mathrm{d}t} + E\varepsilon = \sigma$$

where σ is the overall stress and ε is the overall strain. The Maxwell and Voigt models are considered further in C10: Structure and Properties of Polymers.

DIFFUSION

Fick's first law: Flux $J = -D\frac{\partial c}{\partial x}$ (1 dimension); $\mathbf{J} = -D \,\text{grad}\, c \equiv -D\nabla c$ (3 dimensions)

Fick's second law: $\frac{\partial J}{\partial x} = -\frac{\partial c}{\partial t}$ (1 dimension); $\text{div}\,\mathbf{J} \equiv \nabla.\mathbf{J} = -\frac{\partial c}{\partial t}$ (3 dimensions)

DIFFUSION EQUATION (constant D):

$$\boxed{\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}} \text{ (1 dimension);} \qquad \boxed{\frac{\partial c}{\partial t} = D\nabla^2 c} \text{ (3 dimensions)}$$

- We can represent ∇^2 in various coordinate systems (Cartesian, cylindrical, spherical polar)
- We can also generalise to anisotropic diffusion (when the diffusion coefficient becomes a tensor).

Solutions relevant for materials science diffusion problems are often in the form of:

- A series of 'error functions' (or related integrals)
- Trigonometric (Fourier) series

DIFFUSION IN ONE DIMENSION

$$\boxed{\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}}$$

Trial solution method

Choose a function that experience suggests might be a solution and check to see if indeed it is. For example, we can test the following trial solution:

$$c = \frac{A}{t^{1/2}}\exp\left(-\frac{x^2}{4Dt}\right)$$

Looking at the left hand side of the diffusion equation, we have:

$$\frac{\partial c}{\partial t} = -\frac{1}{2}\frac{A}{t^{3/2}}\exp\left(-\frac{x^2}{4Dt}\right) + \frac{A}{t^{1/2}}\frac{x^2}{4Dt^2}\exp\left(-\frac{x^2}{4Dt}\right) = \frac{A}{t^{1/2}}\exp\left(-\frac{x^2}{4Dt}\right)\left(-\frac{1}{2t} + \frac{x^2}{4Dt^2}\right)$$

Looking at the right hand side of the diffusion equation, we have:

$$\frac{\partial c}{\partial x} = \frac{A}{t^{1/2}}\left(-\frac{2x}{4Dt}\right)\exp\left(-\frac{x^2}{4Dt}\right)$$

$$\frac{\partial^2 c}{\partial x^2} = \frac{A}{t^{1/2}}\left(-\frac{2}{4Dt}\right)\exp\left(-\frac{x^2}{4Dt}\right) + \frac{A}{t^{1/2}}\left(-\frac{2x}{4Dt}\right)^2 \exp\left(-\frac{x^2}{4Dt}\right)$$

$$D\frac{\partial^2 c}{\partial x^2} = \frac{A}{t^{1/2}}\exp\left(-\frac{x^2}{4Dt}\right)\left(-\frac{1}{2t} + \frac{x^2}{4Dt^2}\right)$$

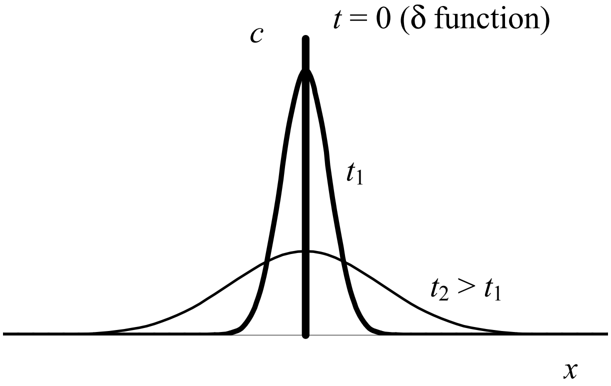
Hence,

$$\boxed{\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}}$$

as required, and so

$$\boxed{c = \frac{A}{t^{1/2}}\exp\left(-\frac{x^2}{4Dt}\right)}$$

is a possible solution of the diffusion equation in one dimension. Of course, it is not the **only** solution. The form of the solution is sketched below:



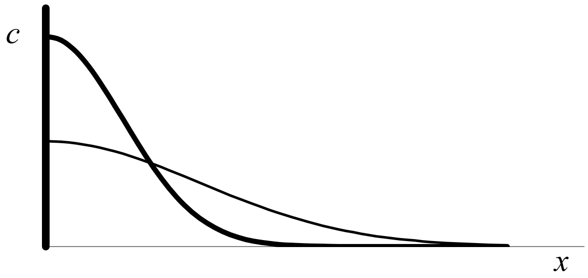
Note: for $x = 2\sqrt{Dt}$ the composition has fallen to $\frac{1}{\mathrm{e}}$ of its value at $x = 0$.

APPLICATIONS

1. DEPOSITION AND DRIVE-IN DIFFUSION

Deposit a thin layer on surface; then heat treat to produce appreciable diffusion into the solid (‘drive-in’ diffusion).

We can use the form of solution of the diffusion equation that we have just shown, noting that the diffusion can be in one direction only (say $+x$).



Since diffusion is only one dimensional, the total amount of solute in a column of unit cross-sectional area going into the sample must always equal the amount, M, initially deposited on unit area before the ‘drive-in’ process.

Hence,

$$M = \int_0^\infty \frac{B}{t^{1/2}} \exp\left(-\frac{x^2}{4Dt}\right) \mathrm{d}x$$

where B is to be determined.

To evaluate this integral, we can make the substitution $\zeta^2 = \dfrac{x^2}{4Dt}$, so that

$$x = \left(2\sqrt{Dt}\right)\zeta \; ; \qquad \mathrm{d}x = \left(2\sqrt{Dt}\right)\mathrm{d}\zeta$$

whence

$$M = \int_0^\infty \frac{B}{t^{1/2}} \exp\left(-\zeta^2\right)\left(2\sqrt{Dt}\right)\mathrm{d}\zeta = 2B\sqrt{D}\int_0^\infty \exp\left(-\zeta^2\right)\mathrm{d}\zeta$$

We can now substitute in a standard result that was derived in IA Mathematics for Natural Sciences:

$$\int_0^\infty \exp\left(-x^2\right)\mathrm{d}x = \frac{\sqrt{\pi}}{2}$$

so that we find

$$M = B\sqrt{\pi D}$$

Thus we can specify the parameter *B* in terms of the amount deposited:

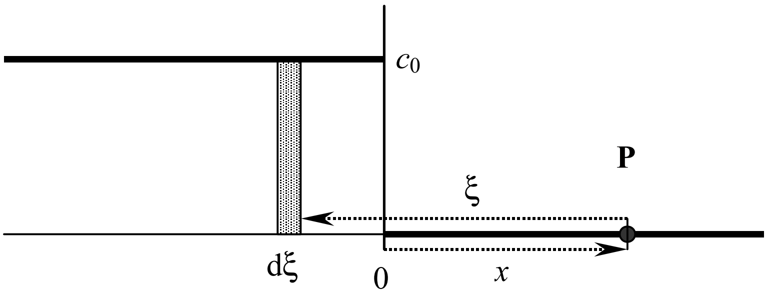
$$B = \frac{M}{\sqrt{\pi D}}$$

and so the composition at depth *x* and time *t* is given by:

$$\boxed{c = \frac{M}{\sqrt{\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right)}$$

---

2. TWO REGIONS WITH DIFFERENT INITIAL, UNIFORM CONCENTRATIONS



At $t = 0$, $c = c_0$ for $x \le 0$ and $c = 0$ for $x > 0$.

Consider the initial extended distribution to be a semi-infinite number of ‘line’ sources.

To calculate the contribution to the concentration at P at time *t*: due to source of thickness dξ at distance ξ from P, note that the initial amount in the source is $M = c_0 \mathrm{d}\xi$.

Thus, the contribution to the concentration at P, $\delta c(x,t)$ from the source of thickness dξ at distance ξ from P is:

$$\delta c(x,t) = \frac{c_0 \mathrm{d}\xi}{2\sqrt{\pi D t}} \exp\left(-\frac{\xi^2}{4Dt}\right)$$

The extra factor of 2 in the denominator arises because here we are dealing with an infinite solid whereas in the previous example we dealt with one that was semi-infinite.

The total concentration at *x* at time *t*, $c(x,t)$, is then obtained by integrating over ξ from *x* to ∞. [Note that because of the way we have set out the problem, ξ is measured from right to left on the figure above.]

Hence,

$$c(x,t)=\int_x^\infty \frac{c_0}{2\sqrt{\pi Dt}}\exp\left(-\frac{\xi^2}{4Dt}\right)\mathrm{d}\xi$$

Again, we make a substitution to simplify the exponent:

$$\zeta^2=\frac{\xi^2}{4Dt}\text{, so that }\xi=\left(2\sqrt{Dt}\right)\zeta\ ;\ \mathrm{d}\xi=\left(2\sqrt{Dt}\right)\mathrm{d}\zeta$$

so that

$$c(x,t)=\frac{c_0}{\sqrt{\pi}}\int_{\frac{x}{2\sqrt{Dt}}}^{\infty}\exp\left(-\zeta^2\right)\mathrm{d}\zeta$$

This is an example of a solution of the diffusion equation related to the 'Error Function' and one you have already met in Part IB Materials Science and Metallurgy.

THE ERROR FUNCTION (erf z) AND THE ERROR FUNCTION COMPLEMENT (erfc z)

Definition

Error Function

$$\boxed{\operatorname{erf} z=\frac{2}{\sqrt{\pi}}\int_0^z \exp\left(-x^2\right)\mathrm{d}x}$$

It follows at once that – erf z = erf (–z), erf (0) = 0 and erf (∞) = 1.

It also follows that

$$\frac{2}{\sqrt{\pi}}\int_z^\infty \exp\left(-x^2\right)\mathrm{d}x=\frac{2}{\sqrt{\pi}}\int_0^\infty \exp\left(-x^2\right)\mathrm{d}x-\frac{2}{\sqrt{\pi}}\int_0^z \exp\left(-x^2\right)\mathrm{d}x=1-\operatorname{erf} z$$

and this is defined as the error function complement, erfc z.

Error Function Complement

$$\boxed{\operatorname{erfc} z=\frac{2}{\sqrt{\pi}}\int_z^\infty \exp\left(-x^2\right)\mathrm{d}x=1-\operatorname{erf} z}$$

Hence the solution already derived for two regions with different initial and uniform concentrations can be written:

$$\boxed{c(x,t)=\frac{c_0}{2}\operatorname{erfc}\left(\frac{x}{2\sqrt{Dt}}\right)}$$

Simple extensions of this methodology enable solutions for (1) diffusion from a constant infinite source of concentration c_1 into an infinite medium with an initial concentration of c_0 (e.g., to model carburisation, covered in Part IB Materials Science and Metallurgy), and (2) interdiffusion between two semi-infinite blocks with initial concentrations of c_1 and c_0 respectively.

METHOD OF SEPARATION OF VARIABLES

Returning to the one dimensional diffusion equation with constant D:

$$\frac{\partial c}{\partial t}=D\frac{\partial^2 c}{\partial x^2}$$

We now look for solutions of the form

$$c(x,t)=X(x)T(t)$$

i.e., solutions which are the product of a function of x only and a function of t only.

Substituting into the diffusion equation, we obtain

$$X\frac{\mathrm{d}T}{\mathrm{d}t}=DT\frac{\mathrm{d}^2X}{\mathrm{d}x^2}$$

replacing the partial differentials by ordinary ones because X is a function of x only and T is a function of t only.

Rearranging this,

$$\frac{1}{T}\frac{\mathrm{d}T}{\mathrm{d}t}=\frac{D}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2}$$

The left hand side of this expression depends on t only and the right hand side depends on x only. Therefore, the two sides must each equal the same constant, independent of x and t, say $-\lambda^2 D$.

$$\frac{1}{T}\frac{\mathrm{d}T}{\mathrm{d}t}=\frac{D}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2}=-\lambda^2 D$$

From this we have solutions such as

$$(1) \qquad T = C\mathrm{e}^{-\lambda^2 Dt}$$

$$(2) \qquad X = A\cos\lambda x + B\sin\lambda x$$

where A, B, and C are constants which depend on the boundary conditions. C will only give a constant of proportionality to A and B, so we arrive at a solution for $c(x, t)$ of

$$c(x,t) = \left(A\cos\lambda x + B\sin\lambda x\right)\mathrm{e}^{-\lambda^2 Dt}$$

Because the diffusion equation is linear, there will be lots of solutions of this form.

The most general solution of this type is:

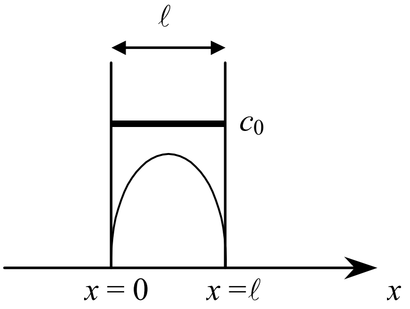
$$\boxed{c(x,t) = \sum_{m=1}^{\infty}\left(A_m\cos\lambda_m x + B_m\sin\lambda_m x\right)\mathrm{e}^{-\lambda_m{}^2 Dt}}$$

where the A_m, B_m and λ_m are determined by the initial and boundary conditions.

Those of you who take the 'Joining of Materials' module in Part III will see a form of this solution of the diffusion equation applied to the mathematical modelling of microsegregation in dendrites and its homogenisation by heat treatment.

APPLICATION

DIFFUSION OUT OF A PLANE SHEET



At $t = 0$, $c = c_0$ for $0 < x < \ell$; $c = 0$ elsewhere.

For $t > 0$, $c = 0$ at $x = 0$ and $x = \ell$.

Using the method of separation of variables, we look for a solution of the form

$$c(x,t) = \sum_{m=1}^{\infty}\left(A_m\cos\lambda_m x + B_m\sin\lambda_m x\right)\mathrm{e}^{-\lambda_m{}^2 Dt}$$

Applying the boundary conditions, we require:

(i) $c = 0$ at $x = 0$ at all times; therefore there can be no contributions from the cosine terms, and so all the $A_m = 0$;

(ii) $c = 0$ at $x = \ell$ at all times; all the sine terms are zero at $x = \ell$; $\lambda_m = \dfrac{m\pi}{\ell}$ where m is an integer.

Applying the initial condition, that at $t = 0$, $c = c_0$ for $0 < x < \ell$, it follows that:

$$c_0 = \sum_{m=1}^{\infty} B_m \sin\frac{m\pi x}{\ell}$$

and so the coefficients B_m are the relevant Fourier coefficients from the Fourier series expansion of c_0.

In the discussion of Fourier series, the following useful results were quoted relevant to the determination of the B_m:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \sin mx \sin nx \, \mathrm{d}x = \begin{cases} 0 & m \neq n \\ \frac{1}{2} & m = n \neq 0 \\ 0 & m = n = 0 \end{cases}$$

If we focus on the result for $m = n \neq 0$, we can change the variable within the integral to x', where

$$\frac{\pi x'}{\ell} = x \quad \text{and} \quad \frac{\pi \, \mathrm{d}x'}{\ell} = \mathrm{d}x$$

so that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \sin^2 mx \, \mathrm{d}x = \frac{1}{\pi}\int_{0}^{\pi} \sin^2 mx \, \mathrm{d}x = \frac{1}{\ell}\int_{0}^{\ell} \sin^2 \frac{m\pi x'}{\ell} \, \mathrm{d}x' \equiv \frac{1}{2}$$

Hence, multiplying both sides of the expression

$$c_0 = \sum_{m=1}^{\infty} B_m \sin\frac{m\pi x}{\ell}$$

by $\sin\dfrac{n\pi x}{\ell}$ and integrating both sides from 0 to ℓ, we find that

$$\int_0^{\ell} c_0 \sin\frac{n\pi x}{\ell}\,\mathrm{d}x = \sum_{m=1}^{\infty}\int_0^{\ell} B_m \sin\frac{m\pi x}{\ell}\sin\frac{n\pi x}{\ell}\,\mathrm{d}x = \frac{B_n \ell}{2}$$

Since c_0 is a constant, the left hand integral can be easily evaluated. Changing the variable back from n to m, we have:

$$\int_0^{\ell} c_0 \sin\frac{m\pi x}{\ell}\,\mathrm{d}x = c_0\left[-\frac{\ell}{m\pi}\cos\frac{m\pi x}{\ell}\right]_0^{\ell} = -\frac{c_0 \ell}{m\pi}\left[\cos m\pi - 1\right]$$

Now cos $m\pi = 1$ if m is even and -1 if m is odd.

Thus, if m is even, $B_m = 0$.

If m is odd, e.g, $m = 2p+1$ for $p = 0, 1, 2, 3, 4$, etc., we have:

$$\frac{2c_0 \ell}{m\pi} = \frac{B_m \ell}{2} \quad \text{for } m \text{ odd}$$

i.e.,

$$B_p = \frac{4c_0}{(2p+1)\pi} \quad \text{for } p = 0, 1, 2, 3, 4, \text{ etc.}$$

Thus, the required solution describing diffusion out of a plane sheet is

$$\boxed{c(x,t) = \frac{4c_0}{\pi}\sum_{p=0}^{\infty}\left[\frac{1}{(2p+1)}\sin\frac{(2p+1)\pi x}{\ell}\exp\left(-\frac{(2p+1)^2\pi^2 Dt}{\ell^2}\right)\right]}$$

As Crank demonstrates in his book on the mathematics of diffusion, this solution is equivalent to one which makes use of error functions instead.

It turns out that for numerical computation the error function form of the solution is most useful for the early stages of diffusion, while the above solution converges satisfactorily for moderate and large times.

STEADY STATE SOLIDIFICATION

You have already met this in Part IB Materials Science and Metallurgy. In this model of solidification, an observer moving with the interface sees a constant solute distribution ahead of the interface arising from the condition that the concentrations of solute in the solid and liquid need to be in equilibrium locally at the solid-liquid interface.

Suppose the liquid-solid interface advances at a constant velocity v along the x-axis, and that this interface is at $x = 0$ when $t = 0$. If we move to a reference frame travelling with the interface, we can define a new axis system, x', so that the coordinate of a point x referred to the x' system at time t is

$$x' = x - vt$$

We can now express the 1D diffusion equation

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$$

in the new x' coordinate system.

To do this we need to make use of the ‘chain rule’ which enables us to change from the independent variables x and t to x'.

$$\frac{\partial}{\partial t} = \frac{\mathrm{d}x'}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}x'} = -v\frac{\mathrm{d}}{\mathrm{d}x'} \;; \qquad \frac{\partial}{\partial x} = \frac{\mathrm{d}x'}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}x'} = 1.\frac{\mathrm{d}}{\mathrm{d}x'} = \frac{\mathrm{d}}{\mathrm{d}x'}$$

Hence,

$$\frac{\partial c}{\partial t} = \frac{\mathrm{d}x'}{\mathrm{d}t}\frac{\mathrm{d}c}{\mathrm{d}x'} = -v\frac{\mathrm{d}c}{\mathrm{d}x'} \;; \qquad \frac{\partial^2 c}{\partial x^2} = \frac{\mathrm{d}}{\mathrm{d}x'}\left(\frac{\mathrm{d}c}{\mathrm{d}x'}\right) = \frac{\mathrm{d}^2 c}{\mathrm{d}x'^2}$$

and so the 1D diffusion equation becomes

$$-v\frac{\mathrm{d}c}{\mathrm{d}x'} = D\frac{\mathrm{d}^2 c}{\mathrm{d}x'^2}\text{, i.e.,} \qquad \frac{\mathrm{d}^2 c}{\mathrm{d}x'^2} + \frac{v}{D}\frac{\mathrm{d}c}{\mathrm{d}x'} = 0$$

If we now drop the suffix on the x, the equation becomes

$$\boxed{\frac{\mathrm{d}^2 c}{\mathrm{d}x^2} + \frac{v}{D}\frac{\mathrm{d}c}{\mathrm{d}x} = 0}$$

which is the equation you met in Course A of Part IB Materials Science & Metallurgy. Solving this equation with the relevant boundary conditions is straightforward:

Let $f = \frac{\mathrm{d}c}{\mathrm{d}x}$. We have

$$\frac{\mathrm{d}f}{\mathrm{d}x} + \frac{v}{D}f = 0$$

which is a simple first order differential equation where we can separate the variables. Hence

$$\frac{\mathrm{d}f}{f} = -\frac{v}{D}\mathrm{d}x$$

and so

$$\ln f = -\frac{v}{D}x + \lambda$$

where λ is a constant of integration. Therefore,

$$f = \frac{\mathrm{d}c}{\mathrm{d}x} = A\exp\left(-\frac{vx}{D}\right)$$

where λ = ln A. Integrating once more,

$$c = -\frac{DA}{v}\exp\left(-\frac{vx}{D}\right) + A_2 \equiv A_1\exp\left(-\frac{vx}{D}\right) + A_2$$

where A_1 and A_2 are determined by the boundary conditions.

If the concentration of solute in the liquid is C_0, the boundary conditions are

$$\text{(i)}\;\; c(\infty) = C_0; \;\; \text{(ii)}\;\; c(0) = C_0/k$$

where $k = C_S/C_L$ is the partition coefficient.

[In words: The solute concentration in the liquid infinitely far away from the interface is invariant and simply the overall solute concentration of the liquid. The concentration of the solid which forms from the liquid is also assumed to be uniform and the same as the overall solute concentration of the liquid before solidification began. The condition for the concentration of solute in the liquid at $x = 0$, $c(0)$, follows.]

Applying the boundary conditions:

$$\text{(i)}\; C_0 = A_2$$

$$\text{(ii)}\; C_0/k = A_1 + A_2$$

and so

$$\boxed{c = C_0 + \frac{C_0(1-k)}{k}\exp\left(-\frac{vx}{D}\right)}$$

which is the result quoted in Part IB Materials Science & Metallurgy.

PROPAGATION OF ERRORS / UNCERTAINTIES

As part of standard good practice experimentally, we need to be able to assess the uncertainty in quantities, q, which are functions of experimentally measured quantities x, …, z with uncertainties in measurement δx, …, δz. Assuming that the uncertainties in measurement are ***independent*** and ***random***, the following general expression can be obtained:

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \ldots + \left(\frac{\partial q}{\partial z}\delta z\right)^2}$$

Note that, even if the independent and random condition is not satisfied:

$$\delta q \not> \left|\frac{\partial q}{\partial x}\delta x\right| + \ldots + \left|\frac{\partial q}{\partial z}\delta z\right|$$

[Note that the sign $\not>$ means ‘cannot be greater than’].

EXAMPLES

1.

$$q = x + y$$

$$\delta q = \sqrt{\delta x^2 + \delta y^2}$$

N.B. beware of the hazards that can arise when trying to determine accurately $q = x - y$ when x and y are nearly equal, because the formula for δq is the same!

When x and y are nearly equal quantities, δq can be as big as or bigger than q.

2.

$$q = x^2$$

$$\delta q = \sqrt{(2x\delta x)^2} = 2x\delta x \quad ; \quad \frac{\delta q}{q} = 2\frac{\delta x}{x}$$

3.

$$q = xy$$

$$\delta q = \sqrt{(y\delta x)^2 + (x\delta y)^2} \quad ; \quad \frac{\delta q}{q} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2}$$

NEVER FORGET! There may be ***systematic*** errors of unknown magnitude and sign.

STATISTICS

Some definitions

For *N* observations:

$$\textit{Mean},\ \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\textit{‘Sample’ standard deviation},\ \sigma_x = \sqrt{\frac{1}{(N-1)}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2}$$

Note the (*N* – 1) term.

The ‘– 1’ arises because one degree of freedom has been used in calculating the mean, $\bar{x}$.

$$\textit{Standard error of the mean},\ \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}$$

Two useful points:

1 To reduce $\sigma_{\bar{x}}$ it is often wiser to seek to reduce σ_x rather than to increase *N*; a rather large increase in *N* would be needed because of the square root, and the apparent improvement in the standard error of the mean might well be meaningless (see the next point).

2. Use common sense: e.g. suppose you take 6 readings of the time required for (say) 10 oscillations of a pendulum using a stop-clock that reads to the nearest 0.1 s and obtain:

   10.0, 10.0, 10.1, 10.0, 10.0, 10.0 (all in s)

   My calculator tells me that, to four significant figures (supposedly!), $\bar{t} = 10.02 \pm 0.04$ s. How confident can I be about the significance of the 0.04?

INTERPRETATION OF ± δq AND ± $\sigma_{\bar{q}}$

Individual result, ± δq

Suppose we measure x, ... , z and then calculate q and δq to arrive at the result $q \pm \delta q$. What does this tell us? **If** the errors are normally distributed there is a 68% chance approximately that the 'correct' value of q lies in the range $q - \delta q$ to $q + \delta q$. It is very important to recognise that δq is not so large that this range includes all possible 'correct' values of q, but that there is an approximately two-thirds chance that the 'correct' value lies within the range.

Average of several results, ± $\sigma_{\bar{q}}$

If we make several measurements and determine $\bar{q}$ and $\sigma_{\bar{q}}$, the standard error of the mean, there is a 68% chance approximately that the 'correct' value of $\bar{q}$ lies in the range $\bar{q} - \sigma_{\bar{q}} \leftrightarrow \bar{q} + \sigma_{\bar{q}}$.

CONFIDENCE/SIGNIFICANCE

Here we are concerned with measures of the probability that the difference between two quantities could have arisen simply because of random chance.

For example, to assess the difference between the mean, $\bar{x}$, of N observations, with standard deviation σ_x, worked out from those observations, and a given value, X, evaluate:

$$t = \left| \frac{\bar{x} - X}{\sigma_x / \sqrt{N}} \right| = \left| \frac{\bar{x} - X}{\sigma_{\bar{x}}} \right|$$

Then consult Tables to find the '*Confidence limits*'. For example, for 10 observations there is a:

5%	[1/20]	chance that	$t = 5.1$
1%	[1/100]	chance that	$t = 10.6$
0.1%	[1/1000]	chance that	$t = 22.9$

simply through random chance.

If $t = 12$ (say), the difference between $\bar{x}$ and X is termed '*significant at the 1% level*'.

REJECTION OF DATA – CHAUVENET’S CRITERION

When, if ever, can we justify rejecting/ignoring a result?

Suppose we have a set of data and we feel that one value looks a bit odd. How do we assess the oddness?

We can assume a normal distribution and estimate the number of measurements that might be expected to be at least as bad as the one in question. If that number turns out to be less than $^1/_2$ out of the total number of measurements, the suspect measurement should be considered for rejection **BUT** first always try to find out if there is a reasonable explanation for its oddness.

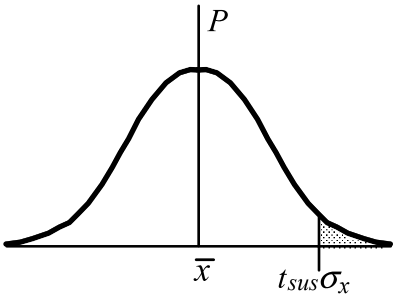
Procedure:

Gather the data $x_1, \ldots, x_N$, and calculate $\bar{x}$ and σ_x. For the suspect measurement, calculate:

$$t_{\text{suspect}} = \left| \frac{x_{\text{suspect}} - \bar{x}}{\sigma_x} \right|$$

i.e, how many standard deviations the suspect measurement is away from the mean. In statistical tables, look up the probability, P_{suspect}, of obtaining a value at least that far from $\bar{x}$, i.e.,

$$P_{\text{suspect}} = P(x \geq \bar{x} + t_{\text{suspect}}\sigma_x)$$



Suppose we have made N total measurements. The probable number, n_{suspect}, of measurements with $x \geq \bar{x} + \mathrm{t}_{\text{suspect}}\sigma_x$ is predicted to be NP_{suspect}.

If $n_{\text{suspect}} < {}^1/_2$, the suspect measurement is to be regarded with suspicion.

Of course the value $^1/_2$ is arbitrary, but nevertheless it is a useful choice (see the discussion about Chauvenet’s criterion in Taylor’s book).

If you do decide to reject the ‘odd’ measurement, recalculate $\bar{x}$ and σ_x, but do **not** use these new values in a second application of Chauvenet's criterion. [If you think you have a second dubious data point, test it against the original distribution.]

Probability of exceeding $t_{\text{suspect}}\sigma_x$

t	P (%)	*Roughly*
1	16	1/6
2	2.3	1/40
3	0.13	1/1000

PROBABILITY DISTRIBUTIONS

Normal or **Gaussian** [mean $\overline{x}$, standard deviation σ_x]

$$\boxed{f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{\left(x - \overline{x}\right)^2}{2{\sigma_x}^2} \right)}$$

Here x is a continuous variable.

For integrals of this function, see standard tables, **BUT** note that some are based on e^{-x^2}, some on $e^{-x^2/2}$. This is the distribution used most routinely in materials science where repeated measurements are made of a particular quantity (e.g., density, length, yield stress of a metal, hardness, etc.).

The Normal distribution is most appropriate to use when an experiment with a finite probability of success, p, is repeated a very large number of times, n, so that the expected outcome, np, is also large. Thus, for example, in the National Lottery, each of the 49 balls has a probability of $^1/_7$ of being drawn each time the National Lottery draw is made (six balls plus the 'bonus' ball).

The distribution in the number of times that each of the 49 balls has been drawn conforms to the Normal distribution.

Poisson [mean a, standard deviation $\approx \sqrt{a}$]

$$\boxed{P(m) = \frac{a^m e^{-a}}{m!}}$$

Here m is the number of occurrences.

The Poisson distribution is appropriate when the number of experiments, n, tends to infinity and the probability of success of each experiment, p, tends to zero in such a way that the product $np = a$ remains finite.

For example, in the National Lottery, the chance of winning the jackpot is

$$\frac{1}{{}^{49}C_6},$$

or roughly 1 in 14 million. Thus, if 28 million lottery tickets are sold, $a = 2$ (i.e., on average two winners are expected).

Hence using the formula above for $P(m)$, $P(0) = 0.135$, $P(1) = P(2) = 0.271$, whereas $P(10) = 4 \times 10^{-5}$. So the chances of a lottery rollover are quite high, while the chances of 10 people sharing the jackpot are very small.

For large a, the Poisson distribution approximates to a Normal distribution with mean a and standard deviation $\sqrt{np(1-p)} \approx a$.

In materials science, applications of the Poisson distribution include counting systems, e.g., X-ray detectors. We can think of each quantum of time (e.g., a millisecond, a second) as a separate experiment, success being one or more X-rays received during that time period. The Poisson distribution is also the relevant distribution to use when describing the number of times electrons are scattered inelastically when travelling through a transmission electron microscopy specimen.

The Poisson distribution and the Normal distribution are different limiting examples of the binomial distribution, as shown in the Appendix.

Weibull

This distribution is applied (for example) to the fracture of ceramics. In this context, from a consideration of the distribution of ‘weak links’, the probability, $P_s(V)$ that a volume V survives an applied stress, σ, is given by:

$$\boxed{P_s(V) = \exp\left(-V\left[\frac{\sigma - \sigma_u}{\sigma_0}\right]^m \right)}$$

σ_u: the ‘threshold stress’ below which there is zero probability of fracture (usually taken as $\sigma_u = 0$);

σ_0: a normalising parameter (to ensure that the integral of the probability distribution over all σ comes to unity);

m: the ***Weibull Modulus***. If m is low (e.g, < 10), there is a very wide scatter of fracture stresses. If m is high (e.g., > 50), there is a narrow scatter of fracture stresses.

An important result derived from this distribution in the case $\sigma_u = 0$ is obtained by taking natural logarithms twice:

$$\boxed{\ln\ln\left(\frac{1}{P_s}\right) = \ln V + m\ln\sigma - m\ln\sigma_0}$$

This is the form commonly used in analysing the results of measurements of the fracture stress of ceramics.

The Weibull Modulus is then obtained from the slope of a graph of ln ln $(1/P_s)$ against ln σ. [*caution:* if you use base-10 logarithms, remember the difference from natural logarithms.]

REGRESSION LINES AND CURVE FITTING

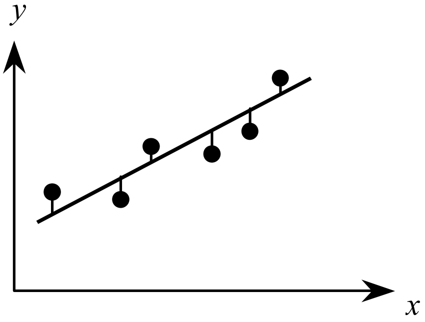
Regression – least squares fit for a line

Suppose we have N pairs of values x_i, y_i

We wish to fit these data points (shown as large dots on the graph below) to a straight line of the form

$$y = A + Bx$$

where it is assumed that the x values are known exactly.



[Remember that useful 'eye-ball' estimates can be obtained by pairing points.]

In the least squares method, we minimise the function

$$\sum_{i=1}^{N} (\text{deviations of } y_i \text{ from line})^2 = \sum_{i=1}^{N} {d_i}^2$$

$$\text{where } d_i = y_i - A - Bx_i$$

It can be shown through the mathematics of maximum likelihood estimation that this minimisation is achieved with:

$$A = \frac{\left(\sum_{i=1}^{N} {x_i}^2\right)\left(\sum_{i=1}^{N} y_i\right) - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} x_i y_i\right)}{N\left(\sum_{i=1}^{N} {x_i}^2\right) - \left(\sum_{i=1}^{N} x_i\right)^2} \qquad B = \frac{N\left(\sum_{i=1}^{N} x_i y_i\right) - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right)}{N\left(\sum_{i=1}^{N} {x_i}^2\right) - \left(\sum_{i=1}^{N} x_i\right)^2}$$

The uncertainties in the parameters *A* and *B* are given by:

$$\sigma_A = \sigma_y \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}} \qquad \sigma_B = \sigma_y \sqrt{\frac{N}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}} \qquad \text{where } \sigma_y = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N} d_i^2}$$

[The factor ($N - 2$) arises in σ_y because two quantities have already been calculated from the data.]

The ‘goodness of fit’ of a straight line to a set of data can be measured by the linear correlation coefficient, *r*:

$$r = \left| \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \sum_{i=1}^{N} (y_i - \bar{y})^2}} \right|$$

where *r* can range from 0 to 1. If *r* is close to 1, the points do indeed lie close to a straight line, while if *r* is close to 0, the points have little or no tendency to lie on a straight line.

Functions which work out *A*, *B* and *r* from a set of data are readily available in most pocket calculators, including the University’s ‘examination’ calculators.

The analysis can straightforwardly be extended to ‘weighted’ data (see the books by Squires and Taylor).

WARNING: Always look at the graphical fit; don’t simply rely on the output from a computer/ calculator!

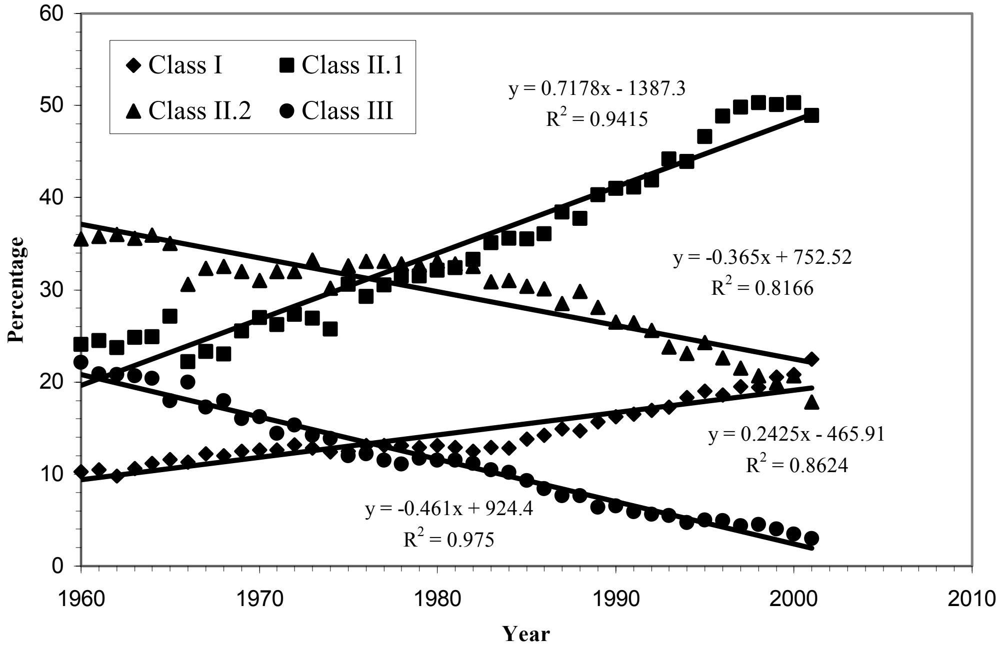
In analysing the data, ask yourself the following:

(i) Does the calculated line look plausible? Are there any ‘odd’ points seriously distorting the fit?

(ii) Should $A = 0$ for sound scientific reasons? If so, is the fit consistent with that?

An example of a least squares fit is shown below, taken from data on Tripos examinations in Cambridge from the Cambridge University Reporter.

Cambridge University Honours Examination Results 1960-2001



Note that just because a reasonable least squares fit can be obtained with data, as in the example above, it does not mean that the fit implies a cause and effect.

Also beware of the dangers of extrapolating from data (*e.g.*, to predict the future given records over *n* years of a variable). This happens routinely on radio and T.V. programmes. Not all extrapolations turn out to be accurate!

Least squares fits for other functions

Linearise if possible; e.g. for

$$y = A\mathrm{e}^{Bx}$$

change to

$$z = \ln y = \ln A + Bx$$

and proceed as above. However, note that the distribution of uncertainties in z is **not** the same as in y. Remedial action may be required.

For a function such as

$$y = Ax + Bx^3$$

we can divide both sides by x and plot a graph of (y/x) against x^2 and proceed as above. Again, note that the the distribution of uncertainties in (y/x) is **not** the same as in y.

APPENDIX

Binomial distribution

The binomial distribution is the distribution which gives the probability, $P(m)$, of m 'successes' and $(n-m)$ 'failures' in n repetitions of an experiment where the probability of success is p and the probability of failure is $(1-p)$:

$$P(m) = {}^{n}C_{m}\, p^{m}(1-p)^{n-m} = \frac{n!}{m!(n-m)!}\, p^{m}(1-p)^{n-m}$$

and so the 'expected' value is $m = pn$.

Two limits of this distribution are

(i) the Normal distribution when both n and pn are large.

(ii) the Poisson distribution where $n \to \infty$ and $p \to 0$ in such a way that the product $pn = a$, a finite number.

Both of these distributions can be derived straightforwardly from the binomial distribution.

Normal distribution

For this limit, we need to use the accurate form of Stirling's approximation:

$$n! = \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^{n}\left(1 + \frac{1}{12n} + \frac{1}{288n^{2}} + \ldots\right)$$

and keep the first term in the asymptotic expansion in the brackets (see for example, Riley, Hobson and Bence, p. 981). [Note that this is more accurate than the form commonly used in the Data Books and calculations in Materials Science and Metallurgy.]

Hence, substituting for the factorials in the binomial distribution, we have

$$P(m) = \frac{\sqrt{2\pi n}\; n^{n}\mathrm{e}^{-n}}{\sqrt{2\pi m}\; m^{m}\mathrm{e}^{-m}\sqrt{2\pi(n-m)}\;(n-m)^{(n-m)}\mathrm{e}^{-(n-m)}}\, p^{m}(1-p)^{n-m}$$

$$= \frac{n^{n+0.5}}{\sqrt{2\pi}\; m^{m+0.5}\;(n-m)^{(n-m)+0.5}}\, p^{m}(1-p)^{n-m}$$

$$= \frac{1}{\sqrt{2\pi n}}\left(\frac{m}{n}\right)^{-m-0.5}\left(\frac{n-m}{n}\right)^{-n+m-0.5} p^{m}(1-p)^{n-m}$$

after some rearrangement.

Making substitutions of the form

$$a^b = \mathrm{e}^c$$

for each of the power terms, it is apparent that the required values of c are $c = b \ln a$. The expression for $P(m)$ becomes

$$P(m) = \frac{1}{\sqrt{2\pi n}} \exp\left((-m - 0.5)\ln\left(\frac{m}{n}\right) + (-n + m - 0.5)\ln\left(\frac{n-m}{n}\right) + m \ln p + (n-m)\ln(1-p) \right)$$

The Normal distribution is then obtained from this daunting expression by letting $m = np + \xi$ where $\xi << np$, and, after some initial tidying up, expanding the logarithms in the expression as power series, keeping only the dominant terms.

Thus,

$$\ln\left(\frac{m}{n}\right) = \ln\left(p + \frac{\xi}{n}\right) = \ln\left(p\left(1 + \frac{\xi}{np}\right)\right) = \ln p + \ln\left(1 + \frac{\xi}{np}\right)$$

and

$$\ln\left(\frac{n-m}{n}\right) = \ln\left(1 - p - \frac{\xi}{n}\right) = \ln\left((1-p)\left(1 - \frac{\xi}{n(1-p)}\right)\right) = \ln(1-p) + \ln\left(1 - \frac{\xi}{n(1-p)}\right)$$

and so substituting for these in the above expression for $P(m)$, we find

$$P(m) = \frac{1}{\sqrt{2\pi n}} \exp\left((-m - 0.5)\ln\left(1 + \frac{\xi}{np}\right) + (-n + m - 0.5)\ln\left(1 - \frac{\xi}{n(1-p)}\right) - 0.5\ln p - 0.5\ln(1-p) \right)$$

$$= \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left((-np - \xi - 0.5)\ln\left(1 + \frac{\xi}{np}\right) + (-n(1-p) + \xi - 0.5)\ln\left(1 - \frac{\xi}{n(1-p)}\right) \right)$$

We can now use the Maclaurin's series expansion for the two logarithm terms:

$$\ln\left(1 + \frac{\xi}{np}\right) = \frac{\xi}{np} - \frac{\xi^2}{2n^2p^2} + O(\xi^3)$$

and

$$\ln\left(1 - \frac{\xi}{n(1-p)}\right) = -\frac{\xi}{n(1-p)} - \frac{\xi^2}{2n^2(1-p)^2} - O(\xi^3)$$

Hence, within the exponential term, we then have

$$(-np-\xi-0.5)\left(\frac{\xi}{np}-\frac{\xi^2}{2n^2p^2}\right)+(-n(1-p)+\xi-0.5)\left(-\frac{\xi}{n(1-p)}-\frac{\xi^2}{2n^2(1-p)^2}\right)$$

ignoring terms $O(\xi^3)$ in the series expansions for the two logarithm terms.

The term in ξ sums to:

$$\left(-1-\frac{1}{2np}\right)\xi+\left(1+\frac{1}{2n(1-p)}\right)\xi=\frac{\xi}{2n}\left(-\frac{1}{p}+\frac{1}{1-p}\right)=\frac{\xi}{2n}\left(\frac{2p-1}{p(1-p)}\right)$$

and the term in ξ^2 sums to

$$\left(-\frac{1}{np}+\frac{(np+0.5)}{2n^2p^2}\right)\xi^2+\left(-\frac{1}{n(1-p)}+\frac{(n(1-p)+0.5)}{2n^2(1-p)^2}\right)\xi^2$$

from which the dominant term here is clearly

$$\boxed{-\frac{1}{2np}\xi^2-\frac{1}{2n(1-p)}\xi^2=-\frac{1}{2np(1-p)}\xi^2}$$

and this term dominates over the term in ξ since $\xi^2 > (2p - 1)\,\xi$. By inspection, the term in ξ^3 in the exponential will be of order $1/n^2$ and therefore smaller than the term in ξ^2.

Hence, in the Normal distribution limit of the binomial expansion, we find

$$\boxed{P(m)=\frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-\frac{\xi^2}{2np(1-p)}\right)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\xi^2}{2\sigma^2}\right)}$$

where σ is a measure of the width of the distribution.

Poisson Distribution

Returning to the expression for $P(m)$,

$$P(m) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}$$

if $m << n$ and $np = a$ while $p \to 0$, we can make the approximations,

$$\frac{n!}{(n-m)!} \to n^m \qquad \text{and} \qquad (1-p)^{n-m} \to (1-p)^n = (1-p)^{a/p} \to \mathrm{e}^{-a}$$

whence

$$\boxed{P(m) = \frac{a^m}{m!} \mathrm{e}^{-a}}$$