Mathematics

Series:

Binomial series

$$(x+y)^N = \sum_{n\approx 0}^{\infty} C_n^N\ x^n\ y^{N-n} \qquad \text{where} \quad C_n^N = \frac{N!}{n!(N-n)!} \qquad (1)$$

e.g. $(x + y)^4 = x^4 + 4\,x^3\,y + 6\,x^2\,y^2 + 4\,x\,y^3 + y^4$

in particular

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \ldots\ldots\ldots \qquad (2)$$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 \quad \ldots\ldots = 1 - x \quad \text{for } x \ll 1 \qquad (3)$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 \quad \ldots\ldots\ldots = 1 + x \quad \text{for } x \ll 1 \qquad (4)$$

Logarithmic series [by integration of (3) and (4)]

$$\ell n(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 \ \ldots\ldots = x \qquad \text{for } x \ll 1 \qquad (5)$$

$$\ell n(1-x) = -x - \tfrac{1}{2}x^2 - \tfrac{1}{3}x^3 \ \ldots\ldots = -x \qquad \text{for } x \ll 1 \qquad (6)$$

Geometric series [derived from (4)]

$$\frac{a}{1-x} = a + ax + ax^2 + ax^3 + \ldots\ldots \text{to } \infty \qquad (7)$$

and so $$\frac{a(1-x^{N+1})}{1-x} = a + ax + ax^2 + ax^3 + \ldots\ldots + ax^N \qquad (8)$$

Exponential series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\ldots \qquad = 1 + x \qquad \text{for } x \ll 1 \qquad (9)$$

hence from $e^{ix} = \cos(x) + i\sin(x)$ (10)

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots\ldots \qquad = 1 - \frac{x^2}{2!} \qquad \text{for } x \ll 1 \text{ radian} \qquad (11)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \ \ldots\ldots \qquad = x \qquad \text{for } x \ll 1 \text{ radian} \qquad (12)$$

Vectors

Unit vectors

A *unit vector* has unit length. The length of a vector **a** is designated $|\mathbf{a}|$ or simply a. A unit vector in the direction of **a** is designated $\hat{\mathbf{a}}$. So

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|}$$

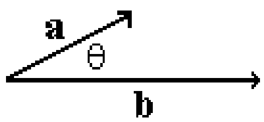
The unit vectors in the x, y and z directions of a conventional right handed set of axes are designated **i**, **j** and **k**.

It follows that the vector, **r**, from the origin to the point (x,y,z) is

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

Products of vectors

There are two that we need to be familiar with:



(i) The *scalar* product, **a.b**, of two vectors **a** and **b** is a *scalar* given by:

$$\mathbf{a.b} = ab\cos(\theta)$$

So, for the unit vectors **i, j, k**, it follows that:

$$\mathbf{i.i} = 1, \quad \mathbf{i.j} = 0, \quad \mathbf{i.k} = 0 \text{ and so on.}$$

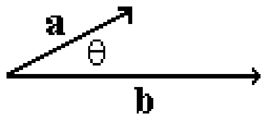
Since $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}$ and $\mathbf{b} = b_x\mathbf{i} + b_y\mathbf{j} + b_z\mathbf{k}$ it follows that the scalar product, **a.b,** can also be expressed as:

$$\mathbf{a.b} = a_xb_x + a_yb_y + a_zb_z$$

An example of the scalar product is the work, W, done by a force, **F**, when it displaces its point of action by a vector **r:**

$$\mathbf{W} = \mathbf{F.r} = F_xx + F_yy + F_zz$$

(ii) The v*ector* product $\mathbf{a}\times\mathbf{b}$

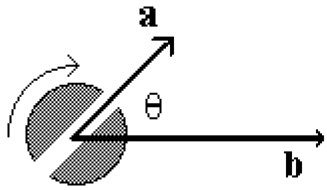


The *magnitude* of $\mathbf{a}\times\mathbf{b}$ is the area of the parallelogram formed by the vectors **a** and **b.** So:

$$|\mathbf{a}\times\mathbf{b}| = ab\sin(\theta)$$

The *direction* of $\mathbf{a}\times\mathbf{b}$ is perpendicular to the flat area of the parallelogram formed by the vectors **a** and **b.** A convention is needed here to choose between the two oppositely directed alternatives.

The direction of $\mathbf{a}\times\mathbf{b}$ is defined as that in which a ***right hand screw***, pointing perpendicular to **a** and **b**, would progress if the slot in the head were rotated from alignment with **a** to alignment with **b** (by the shorter route). So, in the diagram shown, $\mathbf{a}\times\mathbf{b}$ points ***into*** the paper.



It follows that $\mathbf{b}\times\mathbf{a}$ is oppositely directed to $\mathbf{a}\times\mathbf{b}$, i.e. $\mathbf{b}\times\mathbf{a} = -\mathbf{a}\times\mathbf{b}$.

From the definition of $|\mathbf{a}\times\mathbf{b}|$ we see that $\mathbf{i}\times\mathbf{i} = \mathbf{j}\times\mathbf{j} = \mathbf{k}\times\mathbf{k} = 0$.

In sympathy with the right hand screw rule, the conventional ***right-handed*** set of axes is one for which $\mathbf{i}\times\mathbf{j} = \mathbf{k}$ (and, of course $\mathbf{j}\times\mathbf{k} = \mathbf{i}$ and $\mathbf{k}\times\mathbf{i} = \mathbf{j}$).

An alternative statement of the vector cross product of **a** and **b**, as you can verify, is

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = (a_y b_z - a_z b_y)\mathbf{i} + (a_z b_x - a_x b_z)\mathbf{j} + (a_x b_y - a_y b_x)\mathbf{k}$$

Examples:

The force on a charge (q) moving with velocity (**v**) in a magnetic field (**B**) is:

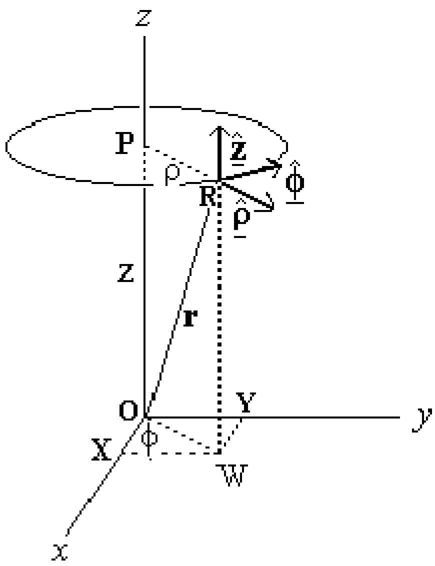
$$\mathbf{F} = q\,\mathbf{v}\times\mathbf{B}$$

The torque exerted by a force (**F**) at position **r** is $\mathbf{T} = \mathbf{r}\times\mathbf{F}$. (This is the torque about an axis in the direction of $\mathbf{r}\times\mathbf{F}$ through the origin.)

The angular momentum of a particle at **r** with linear momentum **p** is $\mathbf{L} = \mathbf{r}\times\mathbf{p}$.

Cylindrical polar co-ordinates

The location of a point in 3-D space can be specified by its cartesian coordinates (x,y,z) or by any other specification that determines the point uniquely. The *cylindrical polar* system, shown in the diagram below, is one possibility.



The vector **OR = r** locates the point that has coordinates (x,y,z) or (ρ,ϕ,z). The two sets of coordinates can, of course, be derived from eachother.

To express x, y and z in terms of ρ, ϕ and z we proceed as follows:

OW = ρ

x = OX = OW cosϕ = ρ cosϕ

y = OY = OW sinϕ = ρ sinϕ

z = OZ = z

The inverse relationships are

$$\rho = +\sqrt{x^2 + y^2}$$

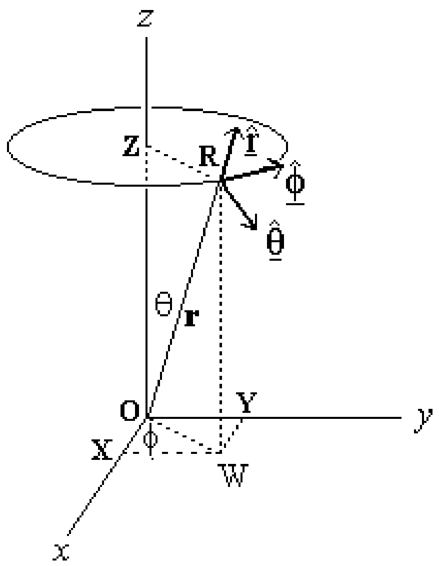
$$\phi = \tan^{-1}\left(\frac{y}{x}\right)$$

$$z = z$$

The diagram also shows the conventional (right handed) set of unit vectors $\underline{\hat{\boldsymbol{\rho}}}$, $\underline{\hat{\boldsymbol{\varphi}}}$ and $\underline{\hat{\mathbf{z}}}$ at the point (ρ,ϕ,z) for cylindrical polar coordinates.

Spherical polar co-ordinates

The location of a point in 3-D space can be specified by its cartesian coordinates (x,y,z) or by any other specification that determines the point uniquely. The *spherical polar* system, shown in the diagram below, is one possibility.



The vector **OR = r** locates the point that has coordinates (x,y,z) or (r,θ,φ). The two sets of coordinates can, of course, be derived from eachother.

To express x, y and z in terms of r, θ and φ we proceed as follows:

OW = r sinθ

x = OX = OW cosφ = r sinθ cosφ

y = OY = OW sinφ = r sinθ sinφ

z = OZ = r cosθ

The inverse relationships are

$$r = +\sqrt{x^2 + y^2 + z^2}$$

$$\theta = \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right)$$

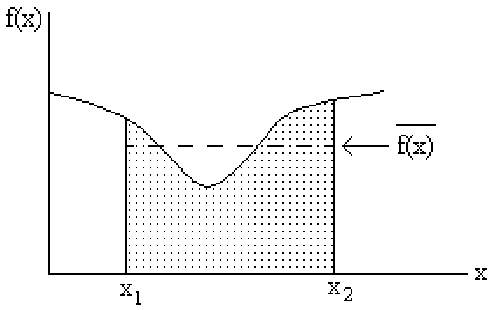
$$\phi = \tan^{-1}\left(\frac{y}{x}\right)$$

The diagram also shows the conventional (right handed) set of unit vectors $\underline{\hat{\mathbf{r}}}$, $\underline{\hat{\boldsymbol{\theta}}}$ and $\underline{\hat{\boldsymbol{\varphi}}}$ at the point (r,θ,φ) for spherical polar coordinates.

Averages

The average of a continuous function is the constant value that has the same area under it as the function itself. So, in the diagram, the average value of f(x), i.e. $\overline{f(x)}$, between x_1 and x_2 is given by

$$\overline{f(x)} \times (x_2 - x_1) = \text{shaded area} = \int_{x_1}^{x_2} f(x)dx$$



This can be stated more generally as

$$\overline{f(x)} = \frac{\int_{x_1}^{x_2} f(x)\,dx}{\int_{x_1}^{x_2} dx}$$

The ***deviation*** of f(x) from its mean value is defined as $f(x) - \overline{f(x)}$. The average of the *deviation*, over a particular range of x, is identically zero; i.e.

$$\overline{f(x) - \overline{f(x)}} = 0$$

Standard deviation.

The ***standard deviation*** of f(x) is a useful measure of the extent to which the function deviates from its mean. Since the average of the deviation, $f(x) - \overline{f(x)}$, is *always* zero we turn to the average of the square of the deviation to get a non-vanishing result.

The *mean square deviation* is $\overline{\left(f(x) - \overline{f(x)}\right)^2}$

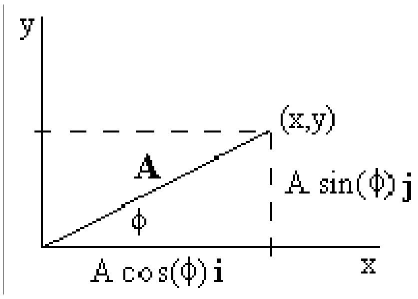
The *standard deviation*, SD, is the *root mean square deviation*:

$$SD = \sqrt{\overline{\left(f(x) - \overline{f(x)}\right)^2}}$$

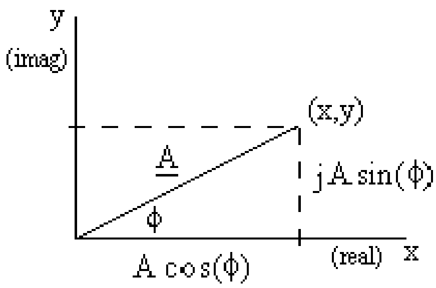
Complex numbers

Complex numbers are two-dimensional objects, so we first draw attention to the common ground between complex numbers and two dimensional vectors.

2-D vectors (unit vectors **i**, **j**)



Complex numbers



The 'vector' from the origin to the point (x,y) is:

$\mathbf{A} = x\,\mathbf{i} + y\,\mathbf{j}$

$= A\cos(\phi)\,\mathbf{i} + A\sin(\phi)\,\mathbf{j} \qquad A = |\mathbf{A}|$

$\underline{A} = x + j\,y$

$= A\cos(\phi) + j\,A\sin(\phi) \qquad A = |\underline{A}|$

$= A[\cos(\phi) + j\sin(\phi)]$

$= A\,e^{j\phi}$

In either case:

$$|\mathbf{A}| = |\underline{A}| = A = \sqrt{x^2 + y^2} \qquad \tan(\phi) = \frac{y}{x}$$

The algebra of complex numbers

This is quite different from 2-D vectors (products of vectors are discussed elsewhere).

Multiplication:

$e^{j\alpha}\,e^{j\phi} = e^{j(\alpha+\phi)}$. So multiplying the complex number $\underline{A} = Ae^{j\phi}$ by $e^{j\alpha}$ produces the new complex number $Ae^{j(\phi+\alpha)}$, and so the representative vector is rotated *anticlockwise* by the angle α.

In particular multiplying a vector by j (= $e^{j\pi/2}$) rotates it anticlockwise by $\pi/2$, and multiplying twice by j rotates it *anticlockwise* by π (180^o), so

$$j^2\,\underline{A} = -\underline{A}$$

Thinking of 'j' as a 90^o rotation operator is wholly legitimate, in which case $j^2 = -1$ is quite intuitive. However, in complex arithmetic 'j' **is** $\sqrt{-1}$. [$j^2 = e^{j\pi/2} \times e^{j\pi/2} = e^{j\pi} = -1$]

The vector $\underline{B} = B\,e^{j\omega t}$ is a vector, of length B ,rotating anticlockwise at angular frequency ω.

Complex conjugate:

If $x + jy = Ae^{j\phi}$ then $x - jy = Ae^{-j\phi}$ and so:

$$(x + jy)(x - jy) = Ae^{j\phi}\, Ae^{-j\phi} = A^2 = x^2 + y^2$$

Division:

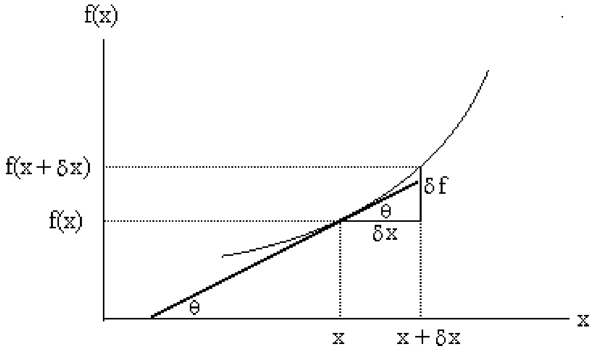
If $x + jy = Ae^{j\phi}$ then $\frac{1}{x + jy} = \frac{1}{Ae^{j\phi}} = \frac{1}{A}e^{-j\phi} = \frac{x - jy}{\sqrt{x^2 + y^2}}$

Differentiation

Differentiation is about how the value of a function changes in response to small changes in the parameters on which it depends.

One independent parameter:

We will call the independent variable x, and the dependent variable f(x).



As indicated on the graph, a change, δx, in the value of x, produces a change, δf, in the value of f. It is apparent from the diagram that as δx is made very small, the ratio $\frac{\delta f}{\delta x}$ tends to the limiting value tan(θ), where θ is the angle that the tangent at x makes with the x-axis.

The notation $\frac{df}{dx}$ is used as a shorthand form for: $\text{Limit}_{\delta x \to 0}\left(\frac{\delta f}{\delta x}\right)$.

$\frac{df}{dx}$ [= tan(θ)] is an indivisible quantity and cannot be split into a 'df' and a 'dx'.

The same argument shows that:

$$\frac{dx}{df} = \text{Limit}_{\delta f \to 0}\left(\frac{\delta x}{\delta f}\right) = \cot(\theta) = \frac{1}{\tan(\theta)} = \frac{1}{\left(\frac{df}{dx}\right)}$$

Example:

For $f(x) = e^{-x^2}$, we can approach the evaluation of $\frac{dx}{df}$ in two ways.

indirectly

$$\frac{df}{dx} = -2x\,e^{-x^2}$$

$$\frac{dx}{df} = \frac{1}{\left(\frac{df}{dx}\right)} = -\frac{1}{2xe^{-x^2}}$$

directly

$$x(f) = \sqrt{-\ln(f)}$$

$$\frac{dx}{df} = \tfrac{1}{2}\left(-\ln(f)\right)^{-\frac{1}{2}}\left(\frac{-1}{f}\right)$$

Approximations to the value of f(x)

If we were to estimate the value of the function at x + δx, our ***first*** approximation would be

$$f(x + \delta x) = f(x) + \delta x \tan(\theta) = f(x) + \left(\frac{df}{dx}\right)_x \delta x$$

where $\left(\frac{df}{dx}\right)_x$ is the value of $\frac{df}{dx}$ at the point x.

i.e. $$\delta f = f(x + \delta x) - f(x) = \left(\frac{df}{dx}\right)_x \delta x$$

This is a ***linear*** approximation since it is based on the assumption that f(x) is a straight line, with gradient $\frac{df}{dx}$, in the vicinity of x.

However, if the gradient of f(x) is changing, then $\frac{d}{dx}\left(\frac{df}{dx}\right) = \frac{d^2f}{dx^2}$ is non-zero, and f(x) is bending away from the tangent at x. The ***second***, or quadratic, approximation to the value of f(x + δx) turns out to be

$$f(x + \delta x) = f(x) + \left(\frac{df}{dx}\right)_x \delta x + \frac{1}{2!}\left(\frac{d^2f}{dx^2}\right)_x (\delta x)^2$$

The full expression for f(x + δx), containing all the terms resulting from the higher derivatives of f(x) at the point x, is called **Taylor's series**. It takes the form:

$$f(x + \delta x) = f(x) + \left(\frac{df}{dx}\right)_x \delta x + \frac{1}{2!}\left(\frac{d^2f}{dx^2}\right)_x (\delta x)^2 + \frac{1}{3!}\left(\frac{d^3f}{dx^3}\right)_x (\delta x)^3 + \text{etc.}$$

$\frac{df}{dx}$ and $\frac{\delta f}{\delta x}$

Strictly speaking $\frac{df}{dx}$ is to be regarded as an indivisible quantity with a specific value (tan(θ)). It cannot be split into a 'df' and a 'dx'. However, we have seen that, for sufficiently small δx,

$$\frac{df}{dx} \approx \frac{\delta f}{\delta x}$$

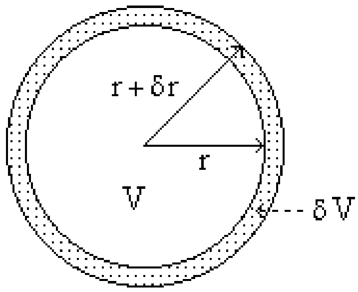
Since δf and δx are *separate* small increments, they can be split up and the above equation (as we have already noted) is equivalent to

$$\delta f \approx \frac{df}{dx}\delta x$$

Consider the example of the volume of a sphere; i.e. $V(r) = \frac{4}{3}\pi r^3$.

This give $\frac{dV}{dr} = 4\pi r^2$. However, for sufficiently small δr, we may approximate

$$\frac{\delta V}{\delta r} \approx \frac{dV}{dr} = 4\pi r^2$$



and so a small change, δr in radius, produces a small change, δV in volume, given by:

$$\delta V \approx 4\pi r^2\,\delta r$$

This is no surprise, since the increase in volume for a very thin layer must be the surface area $(4\pi r^2) \times$ thickness (δr).

[You may wish to satisfy yourself that

$$\delta V = \tfrac{4}{3}\pi(r + \delta r)^3 - \tfrac{4}{3}\pi r^3$$

leads to the same result in the limit of small δr. In this case 'small' means $\delta r \ll r$.]

Generally, physicists do not agonise much about the distinction between $\frac{df}{dx}$ (an indivisible quantity) and $\frac{\delta f}{\delta x}$ (the ratio of two small quantities), and will use df and dx to represent small increments in f and x.

Thus $$\frac{dV}{dr} = 4\pi r^2$$

is rewritten directly as $$dV = 4\pi r^2\,dr$$

'dV' and 'dr' now being used to represent the incremental forms.

Some rules:

Chain rule:

If g = g(f) and f = f(x), then: $\frac{dg}{dx} = \frac{dg}{df} \times \frac{df}{dx}$

Differentiation of a product:

If h(x) = f(x)×g(x), then: $\frac{dh}{dx} = f \times \frac{dg}{dx} + g \times \frac{df}{dx}$

A quotient like $h(x) = \frac{u(x)}{v(x)}$ can be treated as a product: $h(x) = u(x) \times \left(\frac{1}{v(x)}\right)$

Partial Differentiation

When a function depends on more than one parameter, there are correspondingly more derivatives.

If f(x,y,z) is a function of three ***independent*** variables x, y and z, then an extension of our earlier argument shows that small changes δx, δy and δz in x, y and z, produce a small change in the value of f given by

$$\delta f = \left(\frac{\partial f}{\partial x}\right)_{y,z} \delta x + \left(\frac{\partial f}{\partial y}\right)_{x,z} \delta y + \left(\frac{\partial f}{\partial z}\right)_{x,y} \delta z$$

It is necessary to specify ***independent*** variables because $\left(\frac{\partial f}{\partial x}\right)_{y,z}$ implies changing x while keeping y and z fixed. If y depended on x through y = y(x) then it would be impossible to change x without changing y also.

A famous example is the first law of thermodynamics, which relates to the entropy, S(E,V,N), expressed as a function of the independent variables E, the energy of the system, V, the volume of the system and N, the number of particles in the system.

Small changes δE, δV and δN produce a change in S given by:

$$\delta S = \left(\frac{\partial S}{\partial E}\right)_{V,N} \delta E + \left(\frac{\partial S}{\partial V}\right)_{E,N} \delta V + \left(\frac{\partial S}{\partial N}\right)_{E,V} \delta N$$

[To put this in context you have to know that $\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N}$ and $\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N}$, so, for a system containing a fixed number of particles (δN = 0), the above equation becomes the familiar

$$\delta E = T\,\delta S - P\,\delta V$$

]

Second derivatives:

$$\frac{\partial^2 f}{\partial x \partial y} \text{ strictly means } \left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)_x\right]_y$$

It can be shown that the order of partial differentiation is unimportant, so

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Differentiating sums

Consider the example $Z = Z(\beta, E_1, E_2, E_3, \ldots\ldots, E_N) = \sum_{i=1}^{N} e^{-\beta E_i}$ (i)

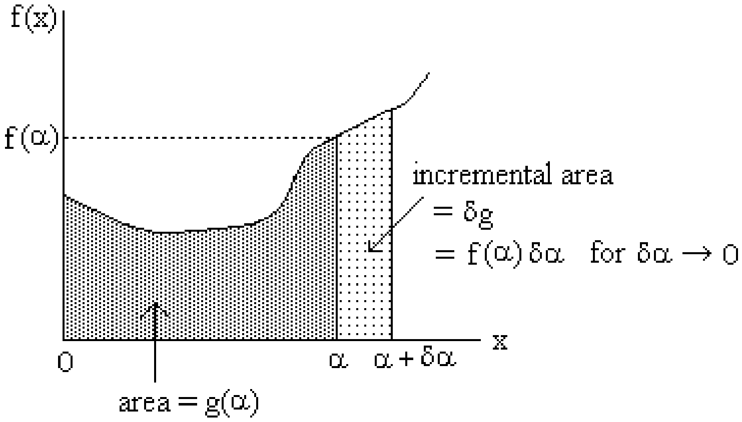
Then $\frac{\partial Z}{\partial \beta} = -\sum_{i=1}^{N} E_i e^{-\beta E_i}$ [since β appears in every term in (i)]

and $\frac{\partial Z}{\partial E_i} = -\beta e^{-\beta E_i}$ [since E_i appears in only one term in (i)]

Differentiating integrals

If $g(\alpha) = \int_0^{\alpha} f(x)\, dx$

then $\frac{dg(\alpha)}{d\alpha} = f(\alpha)$



Proof:

$$\frac{dg(\alpha)}{d\alpha} = L_{\delta\alpha \to 0}\left(\frac{g(\alpha+\delta\alpha) - g(\alpha)}{\delta\alpha}\right) = \frac{\int_0^{\alpha+\delta\alpha} f(x)\,dx - \int_0^{\alpha} f(x)\,dx}{\delta\alpha} = \frac{\int_{\alpha}^{\alpha+\delta\alpha} f(x)\,dx}{\delta\alpha} = \frac{f(\alpha)\,\delta\alpha}{\delta\alpha} = f(\alpha)$$

A small increase in α produces a change in g given by $\delta g = \frac{dg(\alpha)}{d\alpha}\delta\alpha = f(\alpha)\delta\alpha$, as the diagram shows.

Example:

If $g(T) = \int_0^{\frac{\theta}{T}} \frac{x^3}{e^x - 1}\, dx$, where θ is a constant, evaluate $\frac{dg}{dT}$.

First put $\frac{\theta}{T} = \alpha$. g is now a function of α, namely $g(\alpha) = \int_0^{\alpha} \frac{x^3}{e^x - 1}\, dx$

Now $$\frac{dg}{dT} = \frac{dg}{d\alpha}\frac{d\alpha}{dT} = \frac{dg}{d\alpha}\left(-\frac{\theta}{T^2}\right) = \left(\frac{\alpha^3}{e^{\alpha} - 1}\right)\left(-\frac{\theta}{T^2}\right) = -\frac{\theta^4}{T^5\left(e^{\frac{\theta}{T}} - 1\right)}$$

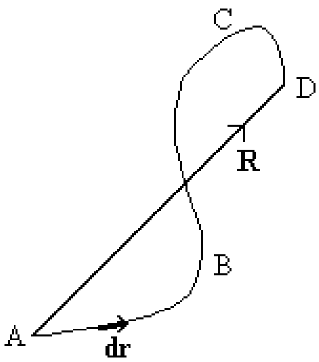
Path Integrals

An *integral* is the name we give to the addition of (infinitesimally) small quantities.

A *path* is a defined route from one point to another. The diagram below shows a *path*, ABCD, from A to D.

A *path integral* is the summation of small quantities defined for each (infinitesimal) element of length of the path. **dr** is an infinitesimal **vector** element of the path.

We will designate the length of **dr** by dr.



NOTE:

$$\int_{\text{path ABCD}} dr$$ means summing the *lengths* of the elements.

So $$\int_{\text{path ABCD}} dr = \text{the total length of the path ABCD.}$$

whereas

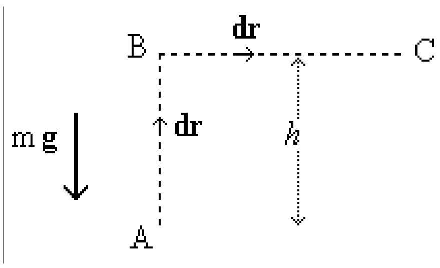
$$\int_{\text{path ABCD}} \mathbf{dr}$$ means summing the small vectors **dr.**

So $$\int_{\text{path ABCD}} \mathbf{dr} = \mathbf{R} \qquad [\text{ the vector } \underline{AD}\ ]$$

Of course, the integrand usually takes a more complicated form. For example, the total work, W, required to move a particle along the path ABCD involves adding up the elements **–F.dr**, where **F** is the force which acts on the particle at each path element **dr**.

So $$W = -\int_{\text{path ABCD}} \mathbf{F.dr}$$

The calculation of W is simplified if the force, **F**, is the same at all points on the path; for example, when moving a mass m in the gravitational field near the earth's surface. In that case **F** is the constant vector m**g**. For the path ABC, we find:



at all elements on AB, $-\mathbf{F.dr} = -m\mathbf{g.dr} = +mg\,dr$;

at all elements on BC, $-\mathbf{F.dr} = -m\mathbf{g.dr} = 0$

So:

$$W = -\int_{\text{path ABC}} \mathbf{F.dr} = mg \int_{\text{path AB}} dr = mgh.$$

Closed paths:

The path integral round a closed path (i.e. one that starts and finishes at the same point) is designated $\oint$.

So $$\oint dr = \text{the circumference of the path,}$$

and $$\oint \mathbf{dr} = 0.$$

Vector calculus

Scalar fields:

A scalar field $\phi(x,y,z)$ is a function of position that has a single numerical value (i.e. a magnitude) at each point. For example, the scalar field 'height' has a single value at each point (x,y) on a contour map.

At any point (x,y,z) a function $\phi(x,y,z)$ may have different slopes, $\left(\frac{\partial\phi}{\partial x}\right)_{y,z}$, $\left(\frac{\partial\phi}{\partial y}\right)_{x,z}$ and $\left(\frac{\partial\phi}{\partial z}\right)_{x,y}$, in the x, y and z directions. From these a **vector** may be constructed at any point according to the prescription

$$\mathbf{i}\frac{\partial\phi}{\partial x}+\mathbf{j}\frac{\partial\phi}{\partial y}+\mathbf{k}\frac{\partial\phi}{\partial z}=\nabla\phi$$

∇ is shorthand for the differential operator ($\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}$). The symbol ∇ is pronounced DEL or GRAD.

The direction of the **vector** $\nabla\phi$ at any point is unique for any well-behaved function.

We shall now prove that the **vector** $\nabla\phi$ points in the direction in which ϕ increases most rapidly and its ***magnitude*** is the rate of change of ϕ in that direction. We prove the second statement first

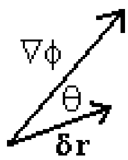
Proof:

In a small step $\boldsymbol{\delta}\mathbf{r} = \mathbf{i}\delta x + \mathbf{j}\delta y + \mathbf{k}\delta z$ from the point (x,y,z) the change in ϕ is

$$\phi(x+\delta x,y+\delta y,z+\delta z) - \phi(x,y,z) = \delta\phi = \frac{\partial\phi}{\partial x}\delta x+\frac{\partial\phi}{\partial y}\delta y+\frac{\partial\phi}{\partial z}\delta z$$

$$= \nabla\phi\mathbf{.}\boldsymbol{\delta}\mathbf{r} \qquad \text{[from the definition of the scalar product]}$$

$$= |\nabla\phi|\,|\boldsymbol{\delta}\mathbf{r}|\cos(\theta)$$



The dependence on $\cos(\theta)$ shows that the greatest change in ϕ for a fixed step length is achieved when $\theta = 0$, i.e. when the step, $\boldsymbol{\delta}\mathbf{r}$, is taken in the direction of $\nabla\phi$.

The maximum change in ϕ in a step of length $|\boldsymbol{\delta}\mathbf{r}|$ is then,

$$\delta\phi_{max} = |\nabla\phi||\boldsymbol{\delta}\mathbf{r}|$$

So the magnitude of the vector $\nabla\phi$ is $|\nabla\phi| = \frac{\delta\phi_{max}}{|\boldsymbol{\delta}\mathbf{r}|}$, i.e. the maximum rate of change of ϕ. [QED]

We shall now prove that the ***direction*** of $\nabla\phi$ is perpendicular to the contours of ϕ. Note that in 2-D the contours are lines (as on a map), but in 3-D the contours are surfaces.

Proof:

If the step $\delta\mathbf{r}$ is taken along a contour (i.e. in a direction in which the value of ϕ does not change), call it $\delta\mathbf{r}_c$, then

$$\nabla\phi.\delta\mathbf{r}_c = \delta\phi = 0$$

It follows that the angle between $\nabla\phi$ and $\delta\mathbf{r}_c$ is $\pi/2$ (90^o).

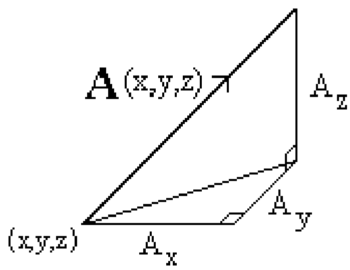
So the direction of steepest ascent (or descent) at a point (x,y) on a map is perpendicular to the contour line through that point (any fool knows that!).

In 3-D the directions normal to the unique direction $\nabla\phi$ at the point (x,y,z) define a local *surface* of constant ϕ.

Vector fields

A vector field $\mathbf{A}$(x,y,z) has a ***magnitude*** and a ***direction*** at any point.

$$\mathbf{A}(x,y,z) = \mathbf{i}\, A_x(x,y,z) + \mathbf{j}\, A_y(x,y,z) + \mathbf{k}\, A_z(x,y,z)$$



To define a vector field the three (scalar) components of the vector must be specified at each point; i.e. each of A_x, A_y and A_z is a ***scalar*** field.

The magnitude of $\mathbf{A}$ is

$$A = |\mathbf{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2}$$

A ***vecto***r field is defined by three ***scalar*** fields A_x, A_y and A_z. So, whereas a scalar field has three independent derivatives, a vector field has nine, namely

$$\begin{array}{ccc} \frac{\partial A_x}{\partial x} & \frac{\partial A_x}{\partial y} & \frac{\partial A_x}{\partial z} \\ \frac{\partial A_y}{\partial x} & \frac{\partial A_y}{\partial y} & \frac{\partial A_y}{\partial z} \\ \frac{\partial A_z}{\partial x} & \frac{\partial A_z}{\partial y} & \frac{\partial A_z}{\partial z} \end{array}$$

This is not as bad as it seems! When we look at the physical properties of vector fields we find that only two combinations of the above derivatives are required. One, the ***divergence***, embodies three of the nine derivatives and the other, the ***curl***, embodies the remaining six.

Divergence (DIV, ∇.)

The divergence of a vector field **A** is:

$$\nabla.\mathbf{A} = (\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}).(\mathbf{i}\,A_x + \mathbf{j}\,A_y + \mathbf{k}\,A_z)$$

i.e.

$$\nabla.\mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

Note that ∇.**A** is a ***scalar*** field and involves the derivatives on the diagonal of the above array.

Fields for which ∇·**A** = 0 everywhere are called ***solenoidal***.

Curl (∇×)

The curl of a vector field **A** is:

$$\nabla\times\mathbf{A} = (\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z})\times(\mathbf{i}\,A_x + \mathbf{j}\,A_y + \mathbf{k}\,A_z)$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}$$

$$= \mathbf{i}(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}) + \mathbf{j}(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}) + \mathbf{k}(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y})$$

Note that ∇×**A** is a ***vector*** field and involves the off-diagonal derivatives in the array on the previous page.

Fields for which ∇×**A** = 0 everywhere are called ***irrotational***.