Mathematical Methods I

Three dimensional vectors and vector calculus

Define vector $\underline{r} = (q_1, q_2, q_3)$ w.r.t some basis. Let this basis be related to cartesians (x, y, z) s.t we have 3 simultaneous equations for x, y, z in terms of q_1, q_2, q_3. i.e x, y, z are functions of q_1, q_2, q_3.

Hence by the chain rule: $dx = \sum_{i=1}^{3} \frac{\partial x}{\partial q_i} dq_i$, $dy = \ldots$

Now in cartesians: $d\underline{r} = dx\,\underline{e}_x + dy\,\underline{e}_y + dz\,\underline{e}_z$ were $\underline{e}_x, \underline{e}_y, \underline{e}_z$ are basis vectors: unitary and mutually orthogonal: $\rightarrow$ ORTHONORMAL. So substituting above result:

$$d\underline{r} = \left(\sum_{i=1}^{3} \frac{\partial x}{\partial q_i} dq_i\right)\underline{e}_x + \left(\sum_{i=1}^{3} \frac{\partial y}{\partial q_i} dq_i\right)\underline{e}_y + \left(\sum_{i=1}^{3} \frac{\partial z}{\partial q_i} dq_i\right)\underline{e}_z$$

$$= \sum_{i=1}^{3} \left(\frac{\partial x}{\partial q_i}\underline{e}_x + \frac{\partial y}{\partial q_i}\underline{e}_y + \frac{\partial z}{\partial q_i}\underline{e}_z\right) dq_i = \sum_{i=1}^{3} \underline{h}_i(\underline{r})\, dq_i$$

Define $\underline{h}_i(\underline{r}) = h_i(\underline{r})\, \underline{e}_i(\underline{r})$ i.e, $|\underline{e}_i| = 1$

Hence ANY curvilinear coordinate system can be used to describe a vector in differential form if it can be related to a cartesian basis. If this coordinate system or BASIS is orthogonal, i.e, $\underline{e}_i \cdot \underline{e}_j = \delta_{ij}$ then:

(for basis (q_1, q_2, q_3)): $\boxed{d\underline{r} = \sum_{i=1}^{3} h_i\, dq_i\, \underline{e}_i}$

where $\boxed{h_i^2 = \left(\frac{\partial x}{\partial q_i}\right)^2 + \left(\frac{\partial y}{\partial q_i}\right)^2 + \left(\frac{\partial z}{\partial q_i}\right)^2}$

(i.e $\underline{h}_i \cdot \underline{h}_i = h_i^2 (\underline{e}_i \cdot \underline{e}_i) = h_i^2$).

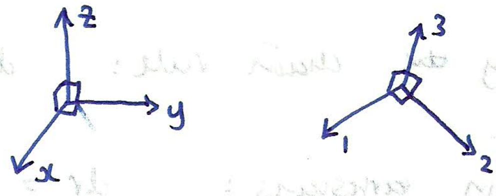
Using this new basis can define a volume element $d\tau$ as:

$$d\tau = (d\underline{r}_1 \times d\underline{r}_2)\cdot d\underline{r}_3$$ Vector triple product - volume of a llepiped.

$$= (h_1\underline{e}_1 dq_1 \times h_2\underline{e}_2 dq_2)\cdot h_3\underline{e}_3 dq_3$$

$$= h_1h_2h_3\, dq_1 dq_2 dq_3 (\underline{e}_1\times\underline{e}_2)\cdot\underline{e}_3$$

Now $\underline{e}_1 \times \underline{e}_2 = \underline{e}_3$ in RH basis of cartesians



So $\boxed{d\tau = h_1h_2h_3\, dq_1 dq_2 dq_3}$

Now for a scalar field ϕ of x,y,z [cartesians]

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz = \nabla\phi\cdot d\underline{r}$$

i.e. vector operator $\underline{\nabla}$ is defined in cartesians by:

$$\underline{\nabla} = \underline{i}\,\partial/\partial x + \underline{j}\,\partial/\partial y + \underline{k}\,\partial/\partial z$$

Interpretation: For surfaces of constant ϕ $d\phi = 0 \Rightarrow \nabla\phi$ or $d\underline{r} = 0$. (or both). Since $d\underline{r}$ could represent any displacement in ϕ; $d\phi = 0 \Rightarrow$ two characteristics of $\nabla\phi$.

(1) $\nabla\phi \perp d\underline{r}$ i.e. $\nabla\phi$ points normal to any SURFACE of constant ϕ.

'$\Rightarrow$' (2) If surface is a local stationary region then since $d\underline{r}$ could be anything; $d\phi = 0 \Rightarrow \nabla\phi = 0$. So $\nabla\phi = 0$ is the equation of a stationary point. Now if $d\phi \neq 0$ clearly it is maximised when $\nabla\phi \,//\, d\underline{r}$. i.e. $\nabla\phi$ points in the direction where ϕ is changing most rapidly. Hence the name G[rad]

Curl and divergence of vector fields generalised by $\underline{F} = \underline{F}(\underline{r})$ are given by:

$$\text{curl } \underline{F} = \underbrace{\underline{\nabla} \times \underline{F}}_{\text{vector}} \qquad \text{div } \underline{F} = \underbrace{\underline{\nabla} \cdot \underline{F}}_{\text{scalar}}$$

Note curl grad $= 0$ i.e, $\nabla \times \nabla = 0$

Now $\underline{\nabla} \cdot \underline{\nabla}\phi \equiv \nabla^2 \phi$ which is the (scalar) Laplacian. We can write $\underline{\nabla}$, curl, div, ∇^2 in general curvilinear coordinates using $d\phi = \nabla\phi \cdot d\underline{r}$ as the defining property of $\underline{\nabla}$. For $\phi(q_1, q_2, q_3)$:

$$\text{i.e,} \quad d\phi = \sum_{i=1}^{3} \frac{\partial \phi}{\partial q_i} dq_i = \sum_{i=1}^{3} \left(\frac{1}{h_i} \frac{\partial \phi}{\partial q_i} \right) h_i \, dq_i$$

Now since $d\underline{r} = \sum_{i=1}^{3} h_i \, dq_i \, \underline{e}_i$

$$\Rightarrow \text{ as } d\phi = \nabla\phi \cdot d\underline{r} \; ; \quad \nabla \phi = \sum_{i=1}^{3} \left(\frac{1}{h_i} \frac{\partial \phi}{\partial q_i} \right) \underline{e}_i$$

$$\text{i.e} \quad \underline{\nabla} = \sum_{i=1}^{3} \frac{\underline{e}_i}{h_i} \frac{\partial}{\partial q_i}.$$

Using this result plus $\nabla \times \nabla = 0$ and other vector formulae: we can generate the following table of vector operators in curvilinear coordinates.

... scalar	grad	$\underline{\nabla} = \sum_{i=1}^{3} \frac{\underline{e}_i}{h_i} \frac{\partial}{\partial q_i} \ldots$ (operates on scalars)
... vector	div	$\underline{\nabla} \cdot \ldots = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1}(h_2 h_3 \ldots_1) + \frac{\partial}{\partial q_2}(h_3 h_1 \ldots_2) + \frac{\partial}{\partial q_3}(h_1 h_2 \ldots_3) \right]$
... vector	curl	$\underline{\nabla} \times \ldots = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \underline{e}_1 & h_2 \underline{e}_2 & h_3 \underline{e}_3 \\ \partial/\partial q_1 & \partial/\partial q_2 & \partial/\partial q_3 \\ h_1 \ldots_1 & h_2 \ldots_2 & h_3 \ldots_3 \end{vmatrix}$
... scalar	Laplacian	$\nabla^2 \ldots = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1}\left(\frac{h_2 h_3}{h_1} \frac{\partial}{\partial q_1} \ldots \right) + \frac{\partial}{\partial q_2}\left(\frac{h_1 h_3}{h_2} \frac{\partial}{\partial q_2} \ldots \right) + \frac{\partial}{\partial q_3}\left(\frac{h_1 h_2}{h_3} \frac{\partial}{\partial q_3} \ldots \right) \right]$

2

Vector operators are made especially useful by their presence in Stokes' and Divergence theorems.

Stokes' Theorem (Generalised).

For open surface S bounded by closed curve C:

$$\oint_C d\underline{l} \begin{Bmatrix} \cdot \underline{A} \\ \times \underline{A} \\ \phi \end{Bmatrix} = \int_S (d\underline{s} \times \nabla) \begin{Bmatrix} \cdot \underline{A} \\ \times \underline{A} \\ \phi \end{Bmatrix}$$

$d\underline{l}$ is a line element of C, $d\underline{s}$ is a surface element of S, $\underline{A}$ is a vector field and ϕ is a scalar field assumed present in the vicinity of S.

i.e. $$\oint_C d\underline{l} \cdot \underline{A} = \int_S (d\underline{s} \times \underline{\nabla}) \cdot \underline{A} = \int_S (\underline{\nabla} \times \underline{A}) \cdot d\underline{s}$$

Divergence Theorem (Generalised)

For a closed surface S having volume V:

$$\oint_S d\underline{s} \begin{Bmatrix} \cdot \underline{A} \\ \times \underline{A} \\ \phi \end{Bmatrix} = \int_V d\tau \, \underline{\nabla} \begin{pmatrix} \cdot \underline{A} \\ \times \underline{A} \\ \phi \end{pmatrix}$$

$d\tau$ is a volume element of V.

i.e. $$\oint_S d\underline{s} \cdot \underline{A} = \int_V d\tau \, \underline{\nabla} \cdot \underline{A}$$

Integral Transforms : Fourier ($\mathcal{F}$) and Laplace ($\mathcal{L}$)

Fourier transforms

A function, $f(x)$ may be expressed as a Fourier series of harmonic functions of period T.

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/T} = \sum_{-\infty}^{\infty} c_n e^{i\omega_n x}$$

where $\omega_n = \dfrac{2\pi n}{T}$ so as $T \to \infty$; if $f(x)$ fulfils the DIRICHLET conditions:

(1) $f(x)$ periodic with period T ← Always ok if $T \to \infty$!

(2) Single valued and continuous - except at a FINITE number of finite discontinuities.

(3) Finite number of stationary points (max, min) within one period

(4) $\int_{-T/2}^{T/2} |f(x)|\,dx$ is convergent, i.e, $\neq \infty$.

then changes in ω_n become vanishingly small → ω_n becomes continuous. ($\Delta\omega = 2\pi/T \to 0$).

Now as $c_n = \dfrac{1}{T}\int_{-T/2}^{T/2} f(x)\, e^{-2\pi i n x/T}\,dx = \dfrac{\Delta\omega}{2\pi}\int_{-T/2}^{T/2} f(x) e^{-i\omega x}\,dx$

$$\Rightarrow \quad f(x) = \sum_{-\infty}^{\infty} \frac{\Delta\omega}{2\pi}\left[\int_{-T/2}^{T/2} f(u)\, e^{-i\omega_n u}\,du\right] e^{i\omega_n x}$$

dummy variable u - no x dependence of c_n (only $w/$ is valid. Don't want to confuse with x!).

Now as $T \to \infty$, $\Delta\omega \to 0$ so first sum → integral. ($\Delta\omega \to d\omega$)*

* so $f(x)$ is not zero.

i.e, $$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, e^{i\omega x}\int_{-\infty}^{\infty} f(u)\, e^{-i\omega u}\,du \qquad (*)$$

Now if we define the Fourier Transform of $f(x)$, $\tilde{f}(\omega)$ as:

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx$$

then (*) $\Rightarrow$ $$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega x} d\omega$$

This is Fourier's inversion theorem.

Proof: $$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega\, e^{i\omega x} \int_{-\infty}^{\infty} du\, f(u) e^{-i\omega u}$$ TRUE.

Now defining $\tilde{f}(\omega)$ above: let $f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega x} d\omega$

$$\Rightarrow f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx \right] e^{i\omega x} d\omega$$

$\rightarrow$ use dummy variables to not confuse integrals over $e^{\pm i\omega x}$

$$\Rightarrow f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega\, e^{i\omega x} \int_{-\infty}^{\infty} f(u) e^{-i\omega u} du$$ i.e (*).

So the Fourier inversion theorem works.

Important functions that behave in interesting ways when Fourier Transformed.

(i) Gaussian. $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-x^2/2\sigma^2\right)$$

$$f[f(x)] = \tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\sigma^2\omega^2}{2}\right)$$ i.e another Gaussian.

Note 'width' of Gaussian $= \sigma$ and $1/\sigma$ for FT.

$\Rightarrow$ $\Delta\omega\Delta x = 1$ where 'Δ' $\Rightarrow$ 'spread' in ω and x.

(2) Dirac δ function

The Dirac δ function is defined:

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

$f(x)\,\delta(x-a) = 0\,,\ x \neq a$
so at $x = a$, $f(a) =$ constant
so is not integrated.

i.e, $\delta(x) = 0\,,\ x \neq 0$ and $\int_{-\infty}^{\infty} \delta(x)\,dx = 1$

(An infinitely sharp pulse at some 'origin' with unit integral over all space).

Now it is clear from the definition that $\delta(x)$ is REAL and SYMMETRIC. i.e, $\delta(x) = \delta(-x)$ and $\delta^*(x) = \delta(x)$.

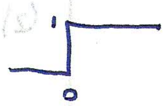
Now $f[\delta(x)] = \tilde{\delta}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega x}\,\delta(x)\,dx$

Now $e^{-i\omega x}\,\delta(x) = 0$ when $x \neq 0$; $e^0 = 1$ when $x = 0$

So $$\tilde{\delta}(\omega) = \frac{1}{\sqrt{2\pi}}$$

$\delta(x)$ is related to two other functions.

(1) Heaviside step function $H(x)$. $H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$



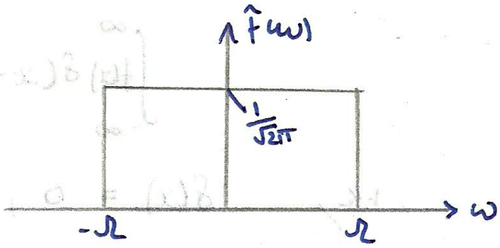
$$H'(x) = \delta(x).$$

Proof: $f(0) = \int_{-\infty}^{\infty} \delta(x)\,f(x)\,dx = \int_{-\infty}^{\infty} H'(x)\,f(x)\,dx$ is true.

i.e, $H'(x)$ has the same property of 'picking out $f(0)$' as $\delta(x)$ hence the equality.

(2) Sinc function, $f(x) = \frac{\sin x}{x}$

Consider a Fourier transform $\tilde{f}(\omega)$ which is a top hat of width 2Ω and height $\frac{1}{\sqrt{2\pi}}$



Now $f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega x} \, d\omega$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \frac{1}{\sqrt{2\pi}} e^{i\omega x} \, d\omega$$

$$= \frac{1}{2\pi} \left[\frac{1}{ix} e^{i\omega x} \right]_{-\Omega}^{\Omega}$$

$$= \frac{1}{2\pi} \left[\frac{1}{ix} e^{ix\Omega} - \frac{1}{ix} e^{-i\Omega x} \right] = \frac{\sin \Omega x}{\pi}$$

Now if $\Omega \to \infty$, $\tilde{f}(\omega) \to \tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}}$ $|\omega| \leq \infty$.

which is the 'definition'* of the $f[\delta(x)]$. Hence

$$\delta(x) = \lim_{\Omega \to \infty} \left(\frac{\sin \Omega x}{\pi} \right)$$

* or proven result more ...ire!

Properties and applications of Fourier Transforms.

(1) Differentiation $\quad f[f^{(n)}(x)] = (i\omega)^n \tilde{f}(\omega)$

(2) Integration $\quad f\left[\int f(x)\,dx\right] = \frac{1}{i\omega}\tilde{f}(\omega) + 2\pi c\,\delta(\omega)$

(c = integration constant.)

(3) Scaling $\quad f[f(ax)] = \frac{1}{a}\tilde{f}\left(\frac{\omega}{a}\right)$

(4) Translation $\quad f[f(x+a)] = e^{ia\omega}\tilde{f}(\omega)$

(5) Exponential multiplication $\quad f[e^{\alpha x} f(x)] = \tilde{f}(\omega + i\alpha)$

Convolution

For functions $f(x)$ and $g(x)$ ('f' and 'g') : convolution $f * g$ is defined by:

$$f * g = \int_{-\infty}^{\infty} dy\, f(y)\, g(x-y)$$

(y again a dummy variable to distinguish it from x).

Note $f * g = g * f$.

Convolution theorem: $$\mathcal{F}[fg] = \frac{1}{\sqrt{2\pi}} \tilde{f} * \tilde{g}$$

AND $$\mathcal{F}[f * g] = \sqrt{2\pi}\, \tilde{f}\tilde{g}$$

This is handy in the theory of optics/waves.

Parseval's Theorem $$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\tilde{f}(\omega)|^2 d\omega$$

Useful for evaluating definite integrals over all space where FT is integratable and original function is not / hard to integrate.

Laplace Transform For functions defined $x > 0$ and have convergence problems when $x \to \infty$ another integral transform is used instead of Fourier transforms. (Non convergence as $x \to \infty$ infringes the Dirichlet conditions). This is the Laplace Transform.

$$\bar{f}(s) = \int_{0}^{\infty} f(x)\, e^{-sx} dx$$

Inverse is NOT easy to find. (No 'Laplace Inversion theorem'). → Require contour integration in complex plane.
→ Use comparison with standard table of LT's for solutions.

Laplace Transforms are useful in solving ODE's by reducing them to linear equations. These can be compared to a table of LT's for the solution to be read off. This makes use of the property:

$$\mathcal{L}\left[f^{(n)}(x)\right] = s^n \bar{f}(s) - s^{n-1} f(0) - s^{n-2} f'(0) - \ldots - f^{(n-1)}(0)$$

other properties and 'standard' Laplace transforms.

Function $f(x)$ $x>0$	LT $\bar{f}(s)$
$\delta(x)$	1
$H(x)$	$1/s$
x^n	$n!/s^{n+1}$
$\sqrt{x}$	$\frac{1}{2}(\pi/s^3)^{1/2}$
$x^{-1/2}$	$(\pi/s)^{1/2}$
e^{-ax}	$1/(s+a)$
$\sin \omega x$	$\omega/(s^2+\omega^2)$
$\cos \omega x$	$s/(s^2+\omega^2)$
$\sinh \omega x$	$\omega/(s^2-\omega^2)$
$\cosh \omega x$	$s/(s^2-\omega^2)$
$e^{-ax} f(x)$	$\bar{f}(s+a)$
$f(x-\tau)\, H(x-\tau)$	$e^{-s\tau}\bar{f}(s)$
$x f(x)$	$-d\bar{f}(s)/ds$
$\int_0^x f(\tau)\, d\tau$	$\bar{f}(s)/s$

Convolution theorem for Laplace Transforms: For f, g

$$\mathcal{L}\left[f * g = g * f\right] = \bar{f}\bar{g}.$$

Note inverse is not true because the lack of a Laplace Transform Inversion theorem.

Matrices and Vector spaces

In general terms a vector is a single columned matrix with N rows corresponding to the N DIMENSIONS of the VECTOR SPACE that we use to describe it.

A vector space of N dimensions is defined by a set of N LINEARLY INDEPENDANT vectors which form a BASIS for the vector space.

linearly independent vector sets $\{\underline{u}_1, \underline{u}_2 \dots \underline{u}_N\}$ that form an N dimensional basis satisfy the following:

$$\text{if } \sum_{i=1}^{N} a_i \underline{u}_i = 0 \Rightarrow a_i = 0 \quad \text{for scalars } a_i.$$

Hence if we introduce another vector $\underline{v}$ into the vector space it becomes linearly dependant and ∴ we can write $\underline{v}$ in terms of all the other vectors

$$\text{i.e., } \underline{v} = \sum_{i=1}^{N} a_i \underline{u}_i$$

linear independence of $\{\underline{u}_i\} \Rightarrow \{a_i\}$ must be unique so $\underline{v}$ is uniquely defined by $\sum_{i=1}^{N} a_i \underline{u}_i$.

so we can express any vector $\underline{v}$ in terms of its components and the basis of the vector space that $\underline{v}$ exists in.

i.e for 3D Euclidean (cartesian) vector space

$$\underline{v} = v_x \underline{e}_x + v_y \underline{e}_y + v_z \underline{e}_z$$

basis $= \{\underline{e}_x, \underline{e}_y, \underline{e}_z\}$.

Note a vector is independent of the basis used to describe it. The reality of a vector (in 3D* a line of finite magnitude and some concept of direction — an arrow if you like) is independent of description though its existence must be defined by some basis which can be arbitrary. Sounds a bit philosophical....

E

Scalar product is an operation between two vectors which yields a scalar.

Defined as $\underline{u}\cdot\underline{v}$ for vectors $\underline{u}$, $\underline{v}$.

Properties: (1) $\underline{u}\cdot\underline{v} \in \mathbb{C}$. (2) $(\underline{u}\cdot\underline{v}) = (\underline{v}\cdot\underline{u})^*$

(3) $\underline{u}\cdot(a\underline{v}_1 + b\underline{v}_2) = a\underline{u}\cdot\underline{v}_1 + b\underline{u}\cdot\underline{v}_2$

(4) $\underline{v}\cdot\underline{v} \geq 0$ i.e. can write $\underline{v}\cdot\underline{v} = |\underline{v}|^2$.

(5) $|\underline{v}| = 0 \Rightarrow \underline{v} = 0$.

Note $(a\underline{u}_1 + b\underline{u}_2)\cdot\underline{v} = a^*\underline{u}_1\cdot\underline{v} + b^*\underline{u}_2\cdot\underline{v}$

and $a\underline{u}\cdot b\underline{v} = a^* b\, \underline{u}\cdot\underline{v}$

So scalar product of two vectors $\underline{a} = \sum_{i=1}^{N} a_i \underline{e}_i$

$\underline{b} = \sum_{j=1}^{N} b_j \underline{e}_j$ w.r.t N dimensional vector space:

$\{\underline{e}_j\} \leftarrow$ basis vector set.

is $\underline{a}\cdot\underline{b} = \left(\sum_{i=1}^{N} a_i \underline{e}_i\right)\cdot\left(\sum_{j=1}^{N} b_j \underline{e}_j\right)$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} a_i^* b_j\, \underline{e}_i\cdot\underline{e}_j$$

Now if basis is orthogonal: $\underline{e}_i\cdot\underline{e}_j = \delta_{ij}$

Hence $\underline{a}\cdot\underline{b} = \sum_{i=1}^{N} a_i^* b_i$

Now if $\underline{e}_i \cdot \underline{e}_j \neq \delta_{ij}$ then define a matrix $\underline{\underline{G}}$ such that $G_{ij} = \underline{e}_i \cdot \underline{e}_j$.

So $\underline{a} \cdot \underline{b} = \sum_{i=1}^{N} \sum_{j=1}^{N} a_i^* G_{ij} b_j$

Now although we have introduced the 'dummy suffix' j to avoid confusion with i the above proves ⇒ we can interpret $\underline{a} \cdot \underline{b}$ as a matrix multiplication with i, j corresponding to rows, columns.

$$\underline{a} \cdot \underline{b} = \underline{a}^{+} \underline{\underline{G}}\, \underline{b} \qquad \underline{a}^{+} = (a_1^* \; a_2^* \; a_3^* \ldots a_n^*)$$

So for orthogonal basis: $\underline{\underline{G}} = \underline{\underline{I}} \Rightarrow \underline{a} \cdot \underline{b} = \underline{a}^{+} \underline{b}$

To ensure correct properties of scalar products: $\underline{\underline{G}}$ must be Hermitian. i.e. $\underline{\underline{G}} = \underline{\underline{G}}^{+}$

Now if vector space is real then all vectors described by it are real. Hence $\underline{\underline{G}}$ is also real.

∴ as $G^{+} = G$ and $G^{*} = G$ ⇒ $G = G^{T}$ which is the defining property of a real symmetric matrix.

Matrices and their properties

Matrices are Associative but non-commutative (usually).

They multiply according to the rule: $\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{C}}$

$$C_{ij} = (AB)_{ij} = \sum_{k=1}^{l} A_{ik} B_{kj}$$ for $\underline{\underline{A}} = n \times l$ $\underline{\underline{B}} = l \times m$ and $\underline{\underline{C}}$ is $n \times m$ order.

One can perform operations on matrix elements: for matrix $\underline{\underline{A}}$ $(n \times m)$:

$$\underline{\underline{A}} \to \underline{\underline{A}}^* \Rightarrow A_{ij} \to A_{ij}^*$$

$$\underline{\underline{A}} \to \underline{\underline{A}}^T \Rightarrow A_{ij} \to A_{ji}$$

$$\underline{\underline{A}} \to \underline{\underline{A}}^\dagger \Rightarrow A_{ij} \to A_{ji}^*$$

and for square matrices: $\underline{\underline{A}} \to \hat{\underline{\underline{A}}}$ $\Rightarrow A_{ij} = C_{ij}^T$ (Transpose of cofactors*)

* see below.

For square $(n \times n)$ matrices we have further properties and special types of matrices. Let $\underline{\underline{A}}$ = $n \times n$ square matrix.

$$\underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{I}} \quad ; \quad I_{ij} = \delta_{ij}$$

$\underline{\underline{A}}^{-1}$ is called the <u>inverse</u>. Now $A_{ij}^{-1} = \dfrac{C_{ij}^T}{|\underline{\underline{A}}|}$

C_{ij} is the 'determinant' of a matrix formed from the reduced/'minor' matrix formed when row i and column j are removed from the matrix.

'determinant' = minor determinant $\times (-1)^{i+j}$

Determinant of a 2×2 matrix is defined as:

For $\underline{\underline{A}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$: $|\underline{\underline{A}}| = ad - bc$.

For higher orders of matrices we use a Laplace expansion to find the determinants in terms of 2×2 ones.

$$|\underline{\underline{A}}| = \sum_{j=1}^{n} a_j C_{ij}$$ (This is one way of doing it).

Each C_{ij} can be found by considering a Laplace expansion of the matrix $\underline{\underline{C}}'$ (ignore $(-1)^{i+j}$ bit.) $|\underline{\underline{A}}|$ is $\therefore$ evaluated as a sum of 2×2 determinants.

Inverses of matrices $\underline{\underline{A}}\,\underline{\underline{B}}\,\underline{\underline{C}} \ldots \underline{\underline{Z}}$ have the property

$$(\underline{\underline{A}}\,\underline{\underline{B}}\,\underline{\underline{C}} \ldots \underline{\underline{Z}})^{-1} = \underline{\underline{Z}}^{-1} \ldots \underline{\underline{C}}^{-1}\underline{\underline{B}}^{-1}\underline{\underline{A}}^{-1}$$

Determinants of matrices $\underline{\underline{A}}\,\underline{\underline{B}}\,\underline{\underline{C}} \ldots \underline{\underline{Z}}$ have the property

$$|\underline{\underline{A}}\,\underline{\underline{B}}\,\underline{\underline{C}} \ldots \underline{\underline{Z}}| = |\underline{\underline{A}}||\underline{\underline{B}}||\underline{\underline{C}}| \ldots |\underline{\underline{Z}}|$$

Special matrices:

(1) Symmetric $\quad \underline{\underline{A}} = \underline{\underline{A}}^T$

(2) Anti-symmetric $\quad \underline{\underline{A}} = -\underline{\underline{A}}^T$

Can express any matrix $\underline{\underline{M}}$ (square) as a sum of symmetric and anti-symmetric matrices:

$$\underline{\underline{M}} = \underbrace{\tfrac{1}{2}(\underline{\underline{M}} + \underline{\underline{M}}^T)}_{\text{Symmetric}} + \underbrace{\tfrac{1}{2}(\underline{\underline{M}} - \underline{\underline{M}}^T)}_{\text{Antisymmetric}}$$

(3) Real $\quad \underline{\underline{A}} = \underline{\underline{A}}^*$

(4) Pure imaginary $\quad \underline{\underline{A}} = -\underline{\underline{A}}^*$

(5) Orthogonal $\quad \underline{\underline{A}}^{-1} = \underline{\underline{A}}^T$ $\quad$ ($\underline{\underline{A}}^T$ is also orthogonal and $|\underline{\underline{A}}| = \pm 1$).

(6) Hermitian $\quad \underline{\underline{A}} = \underline{\underline{A}}^+$

anti - " $\quad \underline{\underline{A}} = -\underline{\underline{A}}^+$

As: $$\underline{\underline{M}} = \underbrace{\tfrac{1}{2}(\underline{\underline{M}} + \underline{\underline{M}}^+)}_{\text{Hermitian}} + \underbrace{\tfrac{1}{2}(\underline{\underline{M}} - \underline{\underline{M}}^+)}_{\text{anti-Hermitian}}$$

we can write any matrix $\underline{\underline{M}}$ (square) as a sum of Hermitian and anti-hermitian matrices.

(7) Unitary $\underline{\underline{A}}^{+} = \underline{\underline{A}}^{-1}$

(8) Normal $\underline{\underline{A}}\,\underline{\underline{A}}^{+} = \underline{\underline{A}}^{+}\underline{\underline{A}}$

Eigenvectors and Eigenvalues of a square matrix

For a square matrix $\underline{\underline{M}}$: $\underline{\underline{M}}\,\underline{e} = \lambda \underline{e}$
defines scalar eigenvalues λ and eigenvectors $\underline{e}$.

Now since $\lambda \underline{e} = \lambda \underline{\underline{I}}\,\underline{e} \Rightarrow \underline{\underline{M}}\,\underline{e} - \lambda \underline{\underline{I}}\,\underline{e} = 0$

Hence $(\underline{\underline{M}} - \underline{\underline{I}}\lambda)\underline{e} = 0$ and since $\underline{e}$ is non zero

$$\Rightarrow \quad |\underline{\underline{M}} - \underline{\underline{I}}\lambda| = 0$$

This characteristic equation yields a polynomial of order of $\underline{\underline{M}}$ which allows $\{\lambda\}$ to be found. (There will be a maximum of N λ's if N is the order of $\underline{\underline{M}}$).

Now let $\underline{\underline{A}}$ = matrix of eigenvectors of a matrix $\underline{\underline{M}}$. $(\underline{\underline{A}})_{ij} = e_{(j)i}$ i.e. jth eigenvector, component i.

i.e. $$\underline{\underline{A}} = \begin{pmatrix} e_{(1)1} & e_{(2)1} & \cdots & e_{(n)1} \\ e_{(1)2} & e_{(2)2} & \cdots & \vdots \\ e_{(1)3} & e_{(3)3} & \cdots & \vdots \\ \vdots & \vdots & & e_{(n)n} \end{pmatrix}$$

Now as $\underline{\underline{M}}\,\underline{e} = \lambda \underline{e} \Rightarrow \sum_{k=1}^{n} M_{jk} e_{(i)k} = \lambda_i e_{(i)j}$

$$\Rightarrow \sum_{k=1}^{n} M_{jk} A_{ki} = A_{ji}\lambda_i = \sum_{k=1}^{n} A_{jk}\delta_{ki}\lambda_i$$

$$\Rightarrow \underline{\underline{M}}\,\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{D}} \quad \text{where } (\underline{\underline{D}})_{ij} = \lambda_{(i)}\delta_{ij}$$

let $A_{ij} = e_{(j)i}$

Better: $(\underline{\underline{M}}\,\underline{\underline{A}})_{ij} = \sum_k M_{ik} A_{kj} = \sum_k M_{ik} e_{(j)k}$ Now $\underline{\underline{M}}\,\underline{e}_{(j)} = \lambda_{(j)}\underline{e}_{(j)} \Rightarrow \sum_k M_{ik} e_{k(j)} = \lambda_{(j)} e_{i(j)} = \lambda_{(j)} A_{ij}$

$\therefore (\underline{\underline{M}}\,\underline{\underline{A}})_{ij} = \lambda_{(j)} A_{ij} = \sum A_{ik}\delta_{kj}\lambda_{(j)}$

So any square matrix $\underline{\underline{M}}$ can be written in the form

$$\underline{\underline{M}} = \underline{\underline{A}}\,\underline{\underline{D}}\,\underline{\underline{A}}^{-1} \qquad (*)$$

or equivalently, any square matrix $\underline{\underline{M}}$ can be diagonalised by $\underline{\underline{A}}^{-1}\underline{\underline{M}}\,\underline{\underline{A}}$. $(= \underline{\underline{D}})$.

Now $(*)$ is useful as $\underline{\underline{M}}^n = \underline{\underline{A}}\,\underline{\underline{D}}^n\underline{\underline{A}}^{-1}$.

$\underline{\underline{D}}^n$ is easily evaluated because it is diagonal.

$$(\underline{\underline{D}}^n)_{ij} = \lambda_{(i)}^n \delta_{ij}$$

Hermitian matrices have interesting/specific eigenvector, eigenvalue properties. ($\underline{\underline{H}}$ = Hermitian matrix, $\underline{\underline{H}} = \underline{\underline{H}}^+$).

(1) eigenvalues are real
(2) eigenvectors are orthogonal if eigenvalues are different.
('orthogonal' $\Rightarrow$ $\underline{e}_{(j)} \cdot \underline{e}_{(i)} = 0$).
(3) $\underline{e}_{(j)}$ and $\underline{e}_{(i)}$ are linearly independent.
(4) If $\{\underline{e}_{(i)}\}$ are orthonormal eigenvectors of $\underline{\underline{H}}$ then $\underline{\underline{A}} = (\underline{e}_{(1)}\ \underline{e}_{(2)}\ \underline{e}_{(3)} \dots)$ is unitary.

Eigenvectors and Principle axis

For equation $\underline{r}^T\underline{\underline{A}}\,\underline{r} = c$.

c = constant ; $\underline{r} = \begin{pmatrix} x \\ y \end{pmatrix}$; $\underline{\underline{A}} = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$ i.e real symmetric

$\Rightarrow$ $\alpha x^2 + 2\beta xy + \gamma y^2 = c$ (equation of conic section).

change basis : $\underline{r} = \underline{\underline{R}}\,\underline{r}'$ where $\underline{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ $\longrightarrow$ more and $\underline{\underline{R}}$ is a real orthogonal rotation matrix

$\Rightarrow$ by substitution: $(\underline{\underline{R}}\,\underline{r}')^T \underline{\underline{A}}\,\underline{\underline{R}}\,\underline{r}' = \underline{r}'^T \underline{\underline{R}}^T \underline{\underline{A}}\,\underline{\underline{R}}\,\underline{r}'$

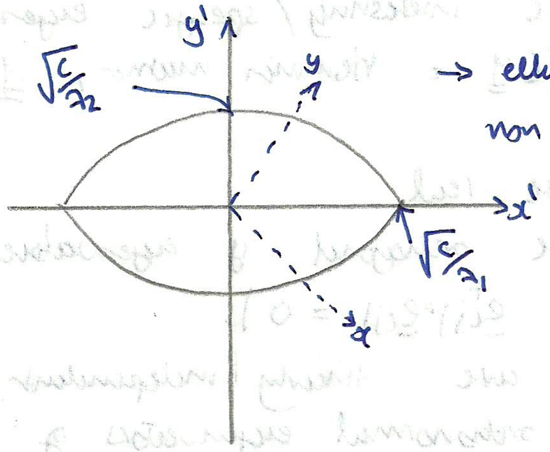
Now since $\underline{\underline{R}}$ is orthogonal: $\underline{\underline{R}}^T = \underline{\underline{R}}^{-1}$

so if $\underline{\underline{R}} = (\{\underline{e}_{(i)}\})$ where $\{\underline{e}_{(i)}\}$ are the eigenvectors of $\underline{\underline{A}}$ then $\underline{\underline{R}}^T \underline{\underline{A}}\,\underline{\underline{R}} = \underline{\underline{D}}$

i.e, $D_{ij} = \lambda_{(i)} \delta_{ij}$ where $\{\lambda_{(i)}\}$ = eigenvalues of $\underline{\underline{A}}$.

So in new basis conic becomes $\underline{r}'^T \underline{\underline{D}}\,\underline{r}' = c$

i.e, $\lambda_1 x'^2 + \lambda_2 y'^2 = c.$ which is much easier to visualise.



$\rightarrow$ ellipse if λ_1, λ_2 are +ve and non zero.

Note $\underline{r}$ could be in higher dimensions where in that case we have 3 eigenvalues and a $\underline{\underline{A}}$ surface.

Because of the squared dependence on the new basis these surfaces are called 'Quadric Surfaces'.

Now eigenvectors of $\underline{\underline{A}}$ are aligned with its 'principle axis' – i.e points where the distance to the quadric surface from the origin is 'stationary'. (i.e min, max).

This leads to the expression

$$\lambda(\underline{e}) = \frac{\underline{e}^T \underline{\underline{A}}\,\underline{e}}{\underline{e}^T \underline{e}}$$

which computes the eigenvalues

of symmetric matrix $\underline{\underline{A}}$ gives the eigenvectors.
If $\{ \underline{e}_j \}$ are known this may be a quicker method of finding $\{ \lambda_j \}$ than solving the characteristic equation.

For Hermitian matrices, $\underline{\underline{H}}$

$$\lambda(\underline{e}) = \frac{\underline{e}^{+} \underline{\underline{H}}\, \underline{e}}{\underline{e}^{+} \underline{e}}$$

Infinite series and convergence

Given an infinite sequence of complex numbers $u_1, u_2 \ldots$ define partial sum S_N by

$$S_N = \sum_{n=1}^{N} u_n$$

If $\lim_{N \to \infty} S_N$ exists and is finite then

$$S = \sum_{n=1}^{\infty} u_n \quad \text{converges.}$$

ABSOLUTE convergence is if $\sum_{n=1}^{\infty} |u_n|$ converges also otherwise convergence is conditional on the sign of u_n.

i.e. $S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \ldots\ldots$ converges to $\ln 2$.
but $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}$ is not convergent.

Tests for convergence

(1) Comparison test: If $\sum_{n=1}^{\infty} v_n$ converges then series

$S = \sum_{n=1}^{\infty} u_n$ converges if $|u_n| < C|v_n|$ for some n independent C.

Further tests exploit proven convergence of a geometric series:

$$S = 1 + x + x^2 + x^3 + \ldots \qquad |x| < 1 \Rightarrow S \neq \infty.$$

Now for $S_N = \underset{1st}{1} + \underset{2nd}{x} + \ldots \underset{Nth\ term}{x^{N-1}} = \dfrac{1-x^N}{1-x}$

Proof/ [Induction] $S_0 = \dfrac{1-x^0}{1-0} = 0$ $\quad S_1 = \dfrac{1-x}{1-x} = 1$ ✓

Not essential but helpful!

Assume kth term is true, i.e $S_k = \dfrac{1-x^k}{1-x}$

Now S_{k+1} is assumed to be $S_{k+1} = \dfrac{1-x^{k+1}}{1-x}$

Now we know $S_{k+1} - S_k = x^k$ as $k+1$ th term in series is $x^{k+1-1} = x^k$.

Hence $S_{k+1} - S_k = \dfrac{1-x^{k+1} - 1 + x^k}{1-x} = \dfrac{x^k(-x+1)}{1-x}$

$= x^k$ ✓ $\quad$ so $S_N = \dfrac{1-x^N}{1-x}$ is true.

Now $\lim\limits_{N\to\infty} S_N = \dfrac{1}{1-x}$ for $|x|<1$. So Geometric series is convergent for $|x|<1$.

(2) Ratio test $\quad S$ converges if $\lim\limits_{n\to\infty} \left| \dfrac{u_{n+1}}{u_n} \right| < 1$

(3) Cauchy's test $\quad S \quad$ " $\quad$ " $\quad \lim\limits_{n\to\infty} |u_n|^{1/n} < 1$

Note these tests test for absolute convergence.

Taylor series. $\left(\text{For } f(z) = \sum_{n=0}^{\infty} a_n z^n \right)$.

If $\sum_{n=0}^{\infty} a_n z^n$ converges for $z = c$ $[c \in \mathbb{C}]$ then it absolutely converges for all z s.t. $|z| < |c|$.

i.e, within some 'radius of convergence' the Taylor series will converge. Let this $= R$. R can be determined from Ratio and Cauchy tests.

(1) Ratio : $$\frac{1}{R} = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$ if limit exists.

(2) Cauchy : $$\frac{1}{R} = \lim_{n\to\infty} |a_n|^{1/n}$$ " " ".

Proof results from convergence properties of tests and desired properties of R being satisfied when the above are made as an hypothesis.

Note: A function $f(z)$ is ANALYTIC at $z = z_0$ if it has a Taylor expansion about $z = z_0$ with a non-zero radius of convergence.

i.e, for $f(z) = \sum_{n=1}^{\infty} a_n (z - z_0)^n$; $\dfrac{1}{\lim_{n\to\infty} |a_n|^{1/n}} \neq 0$.

or use ratio test.

2nd order ODE's. (linear)

The general form of a 2nd order linear ODE is: (in $y(x)$):

$$y'' + p(x) y' + q(x) y = f(x)$$

It is homogenous if $f(x) = 0$ and inhomogenous if $f(x) \neq 0$

The general solution will be of the form

$$y(x) = y_c(x) + y_p(x)$$

where $y_c(x)$ is the solution to the homogenous case and $y_p(x)$ is another solution to the ODE. $y_c(x)$ must be composed of two linearly independent functions s.t.

$$y_c = c_1 y_1 + c_2 y_2$$

where c_1, c_2 are arbitrary constants defined by the b.c's of the ODE (if appropriate). linear independence is found by computation of wronskian:

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} \neq 0 \text{ if } y_1 ; y_2 \text{ linearly independent.}$$

Solution methods: Homogenous case.

(1) Series solution. Try solution $y(x) = \sum_{n=0}^{\infty} a_n x^n$
This will work if ODE is analytic at $x=0$ i.e $x=0$ is an ORDINARY POINT. → can try any Taylor series about ordinary point $x = x_0$ by modifying $x \to x - x_0$

If $x=0$ is not analytic ⇒ singular point (i.e ∞ crops up). If $p(x)\,x$ and $q(x)\,x^2$ are analytic at $x=0$ then $x=0$ is a regular singular point. Otherwise it is an irregular or essential singularity and it is not that helpful.

For regular singular points try a power series

(Frobenius series) $y(x) = x^\sigma \sum_{n=0}^{\infty} a_n x^n$

$\rightsquigarrow$ comparing coefficients of x^0 yields a quadratic for σ.* Note some solutions of this (when σ differs by an integer) will NOT work. ONE solution will always work though $\Rightarrow$ use Wronskian to find the second solution.

* indicial equation.

Idea is to substitute power series into ODE and change indexes of x to yield a recurrence relation for a_n. Use this and linear independence considerations to find $y_c = c_1 y_1 + c_2 y_2$. (i.e find y_1 and y_2 then find c_1, c_2 via b.c's if any).

For values of σ which don't work: Find 1 solution then use Wronskian.

$$W(x) = y_1 y_2' - y_2 y_1' \Rightarrow \frac{W}{y_1^2} = \frac{y_2'}{y_1} - \frac{y_2 y_1'}{y_1^2}$$

$$= \frac{d}{dx}\left(\frac{y_2}{y_1}\right) \quad \left[\text{check: } \frac{d}{dx}\left(\frac{y_2}{y_1}\right) = \frac{y_1 y_2' - y_2 y_1'}{y_1^2} = \frac{y_2'}{y_1} - \frac{y_2 y_1'}{y_1^2} \checkmark\right]$$

$W' = y_1 y_2'' - y_1'' y_2$

↑ use fact that y_2, y_1 satisfy ODE to yield $W' = -pW$ $\Rightarrow W = Ce^{\int -p(x')dx'}$

$$\Rightarrow \text{as } W = C\exp\left[\int^x - p(x')dx'\right]$$

(we can make $C = 1$ as it can be 'absorbed' by c_1, c_2)

$$= \exp\left[\int -p(x')dx'\right]$$

↑ dummy variable

$$\Rightarrow \int^x W/y_1^2 \, dx' = \frac{y_2}{y_1} \Rightarrow y_2 = y_1(x) \int^x \frac{1}{y_1^2(x')} \exp\left[\int^{x'} p(x')dx'\right] dx'$$

So if y_1 is known we can find y_2.

OTHERWISE Try a substitution and try to find y_c analytically (without series method).

For Inhomogenous cases : $f(x) \neq 0$.

(2) Method of variation of parameters. For $y_c = c_1 y_1(x) + c_2 y_2(x)$ constants

Known.

let $y_p(x) = \underbrace{c_1(x) \; y_1(x) + c_2(x)}_{\text{functions}} y_2(x)$.

and restrict :
$$\begin{aligned} c_1' y_1' + c_2' y_2' &= f \\ c_1' y + c_2' y_2 &= 0 \end{aligned}$$

Allows c_1', c_2' to be found. Integration ∴ yields y_p.

(3) Green's functions.

2nd order, linear

Define $G(x,z)$ s.t for a given ODE of b.c $y(\alpha)=0 \quad y(\beta)=0$ as:

$$y(x) = \int_\alpha^\beta G(x,z) \, f(z) \, dz \qquad (*)$$

By substitution into the ODE is appears $\mathcal{L}G(x,z)$ has properties of $\delta(x-z)$. $\mathcal{L}$ is the operation of the ODE on $G(x,z)$

Hence : $$G'' + pG' + qG = \delta(x-z).$$

So for $x > z$, $x < z$ we have a homogenous ODE to solve for G which we can integrate up using $(*)$

$$y(x) = \underbrace{\int_\alpha^x G(x,z) f(z) dz}_{x>z} + \underbrace{\int_x^\beta G(x,z) f(z) dz}_{x<z}$$

Note that the coefficients of the solutions of the homogeneous equation will be function of z. These can be found by noting:

$G(x,z)$ must satisfy b.c's of ODE
~~Derivatives up to~~ First derivative of G at $x = z$ has unit discontinuity whereas lower derivatives are continuous.

i.e/ if $G(x,z) = A(z)\,y(x) + B(z)\,p(x)$

(1) $A(z)\,y(z) - B(z)\,p(z) = 0$
(2) $A(z)\,y'(z) - B(z)\,p'(z) = 1$

These conditions should allow $G(x,z)$ to be found.

Summary:

$y'' + py' + qy = f(x)$
2ND ORDER LINEAR ODE.

- $f(x) = 0$
  - p, q constant: Try $y = Ae^{px}$ etc...
  - $p, q : (x)$: Substitution OR Series Solution. Note singular points. Taylor or Frobenius series. May need Wronskian.
- $f(x) \neq 0$
  - p, q constant: Solve homo then try y_p of same form to $f(x)$. i.e $ax + bx^2 + cx^3 \ldots$ OR USE LAPLACE TRANSFORM.
  - $p, q : (x)$: Variation of parameters. Solve homo.
  - Green's functions. Need to solve homo for $G(x,z)$.

Note/ b.c's must be homogeneous. If not substitute polynomial that fits b.c's s.t 'y' $= y - P(x)$ i.e so 'y'$(x = \text{b.c}) = 0$.

Eigenfunctions and Sturm Liouville Formalism

Analogous to a vector space we may consider a set of functions defined on an interval $[\alpha, \beta]$ which, since the interval is continuous, form an infinite set of 'basis functions' which all other functions on that interval may be expressed in terms of. [Assume basis functions are linearly independent]. ~~let these~~

i.e. $$f(x) = \sum_{n=0}^{\infty} c_n y_n(x)$$ where $y_n(x)$ = basis function.

Define inner product (aka scalar product with metric)

$$\langle f(x) | g(x) \rangle = \int_{\alpha}^{\beta} f^*(x)\, g(x)\, w(x)\, dx$$

$w(x)$ is some weight function which we will define later. (It must be real and defined on $[\alpha, \beta]$).

orthogonality results when $\langle f | g \rangle = 0$

Norm of function f is $\|f\| = \langle f | f \rangle^{\frac{1}{2}} = \left[\int_{\alpha}^{\beta} f(x) f^*(x)\, w(x)\, dx \right]^{\frac{1}{2}}$

We can define an orthonormal set of basis functions y_n by satisfying $\langle y_i | y_j \rangle = \delta_{ij}$ where $\|y_i\| = 1$.

For an orthonormal basis $y_n(x)$ a function $f(x)$ can be written

$$f(x) = \sum_{n=0}^{\infty} a_n y_n(x) \quad \text{where } a_n = \langle y_n | f(x) \rangle$$

This restriction on a_n ⇒ orthonormality of basis functions a useful property.

$f(x)$ defined on $[\alpha, \beta]$ – satisfies Dirichlet conditions

Useful inequalities For an infinite set of functions on $[\alpha, \beta]$ with non zero inner product (A Hilbert space):

(i) Schwartz inequality $|\langle f|g\rangle| \le \langle f|f\rangle^{\frac{1}{2}} \langle g|g\rangle^{\frac{1}{2}}$
~~when~~ when $f = \lambda g$, $\lambda \in \mathbb{C}$ equality holds.

(ii) Triangle inequality $\|f+g\| \le \|f\| + \|g\|$
If $f = \lambda g$ equality holds. $(\lambda \in \mathbb{C})$.

(iii) Bessel's inequality: For coefficients of an orthonormal basis c_n: $\langle f|f\rangle = \|f\|^2 \ge \sum_{n=0}^{\infty} |c_n|^2$

Self Adjoint or Hermitian operators.

For a linear operator $\mathcal{L}$ acting on a function $y(x)$ $\mathcal{L}$ is self adjoint or Hermitian if

$$\langle u|\mathcal{L}v\rangle = \langle v|\mathcal{L}u\rangle^*$$ for some functions $u(x)$, $v(x)$.

Operators of the form $\mathcal{L} = -\frac{\partial}{\partial x}\left(p(x)\frac{\partial}{\partial x}\right) + \sigma(x)$

are Hermitian. $\rightarrow$ S.L operator. We can extend matrix analogy by considering:

$$\mathcal{L}y = \lambda y \qquad \lambda \in \mathbb{C}.$$ postulate.

$y =$ 'eigenfunctions' λ - eigenvalue.

It turns out for Hermitian operators $\{\lambda\}$ are real and eigenfunctions are orthogonal if λ are distinct.

$w(x)$ is defined to convert a non self adjoint | Hermitian $\mathcal{L}$ into one which is.

Hence $\mathcal{L}y = \lambda w y$ will be a general eigenfunction equation where $\mathcal{L}$ is Hermitian $= w\mathcal{L}'$ where $\mathcal{L}$ may not be Hermitian.

Or...... recognise the $\lambda w(x)$ from the equation of $\mathcal{L}y = \lambda w y$ and multiply by integrating factor

$$F(x) = \exp\left[\int^x \frac{r(z) - p'(z)}{p(z)}\, dz\right]$$

For general $\mathcal{L}y = \lambda w y$ as

$$p(x)y'' + r(x)y' + q(x)y + \lambda w(x)y = 0$$

For self adjoint prm $p'(x) = r(x)$.
So if not then multiply by $F(x)$.

Use of S.L prm : ① Eigenfunction expansions
② Solving Inhomo. ODE's via....

$$G(x,z) = \sum_{n=0}^{\infty} \frac{1}{\lambda_n} y_n(x) y_n^*(z)$$

orthonormal.

λ_n = nth eigenvalue, y_n = nth eigenfunction.

So solving $\mathcal{L}y = w(x)\lambda y$ for y_n, λ_n can be used to solve $\mathcal{L}y = f(x)$.

Note

what is the purpose of Eigenfunctions and S.L theory?

with analogy to Hermitian matrix $\underline{\underline{H}}$, Hermitian operators $\mathcal{L}$ can be present in eigenvalue equations:

$$\underline{\underline{H}}\,\underline{y} = \lambda \underline{y} \longleftrightarrow \mathcal{L} y = \lambda y$$

or weight function if $\mathcal{L}$ was not originally S. adj.

Now if $\underline{\underline{H}}$ is known $\{\lambda\}$ and $\{\underline{y}\}$ can be found using

$\left|\underline{\underline{H}} - \underline{\underline{I}}\lambda\right| = 0$ to find $\{\lambda\}$ and then back substitution to find $\{\underline{y}\}$.

This does not appear to be possible for $\mathcal{L}y = \lambda y$.
(Because order of $\mathcal{L}$ is $\infty \times \infty$?)

$\therefore$ Since for given $\mathcal{L}$, y and λ cannot be explicitly determined what is the point?
→ guess λ and solve $\mathcal{L}y = \lambda w y$ for y. or guess y and use Raleigh Ritz to find λ,
→ For eigenfunction expansions (orthonormal eigenfunctions)

$$f(x) = \sum_{i=0}^{\infty} y_i(x)\, c_i \qquad c_i = \langle y_i | f \rangle$$

where $f(x)$ satisfies Dirichlet conditions on boundary
→ How does one determine y_i? → seems like one must invert a S.L operator and solve for y_i λ_i subject to b.cs. Wouldn't it be easier to write down a Taylor series for $f(x)$. Seems far less ambiguous!

(Though I suppose if orthogonality of 'basis' is required a S.L approach may be helpful – However we can always use a Gram Schmidt procedure if we know our expansion to orthogonalise....)

If S.L approach is used to solve ODE's.

→ For homogeneous cases the S.L approach seems to be: "Solve the homogeneous equation (with a λy ambiguously placed) and then ~~recog~~ recognise eigenfunctions". → Surely we have to already solve the ODE for this method to work!

→ For inhomogeneous cases if we solve $\mathcal{L}y = w\lambda y$ (i.e for homo-case) we can recover a form of $G(x,z)$. Are we not effectively solving the homo case to find $G(x,z)$ $x<z$, $x>z$ anyway? Again what is the point with ~~the~~ introducing the S.L concept?

Calculus of Variations

Define **FUNCTIONAL** $F[y]$ as the integral:

$$F[y] = \int_\alpha^\beta dx\, f(y, y'; x) \qquad y' \equiv \frac{dy}{dx}$$

i.e, f is a function of y and y' which are in turn functions of independent variable x.

For a problem: "what are the functions $y(x)$ s.t $F[y]$ is stationary" we can consider the effect on F of changing y by a small amount δy.

$$\text{i.e}\quad \delta F = F[y+\delta y] - F[y]$$

$$= \int_\alpha^\beta dx\, f(y+\delta y, y' + (\delta y)'; x) - \int_\alpha^\beta dx\, f(y, y'; x)$$

$$= \int_\alpha^\beta dx \left[\delta y \frac{\partial f}{\partial y} + (\delta y)' \frac{\partial f}{\partial y'} \right] + O(\delta y^2)$$

As by two variable Taylor series:

$$f(y+\delta y, y' + (\delta y)'; x) \approx \delta y \frac{\partial f}{\partial y} + (\delta y)' \frac{\partial f}{\partial y'} + f(y, y')$$

This term cancels.

So for stationary $F[y]$, $\delta F = 0$

$$\Rightarrow \quad 0 = \int_\alpha^\beta dx \left(\delta y \frac{\partial f}{\partial y} + (\delta y)' \frac{\partial f}{\partial y'} \right) \quad \text{neglecting } \delta y^2 \text{ terms.}$$

$$\Rightarrow \quad 0 = \int_\alpha^\beta dx\, \delta y \frac{\partial f}{\partial y} + \left[\delta y \frac{\partial f}{\partial y'} \right]_\alpha^\beta - \int_\alpha^\beta \delta y \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) dx$$

$$\Rightarrow \quad \left[\delta y \frac{\partial f}{\partial y'} \right]_\alpha^\beta + \int_\alpha^\beta dx \left(\delta y \frac{\partial f}{\partial y} - \delta y \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right).$$

restrict $\delta y(\beta) = \delta y(\alpha) = 0 \Rightarrow \dfrac{\partial f}{\partial y} - \dfrac{d}{dx}\left(\dfrac{\partial f}{\partial y'} \right) = 0$ as δy is arbitrary within $[\alpha, \beta]$ for equality to hold.

yields Euler equation for F, y, y'

$$\frac{\partial F}{\partial y} = \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)$$

So if F does not contain y explicitly:
$\frac{\partial F}{\partial y} = 0 \Rightarrow \frac{\partial F}{\partial y'} = \text{constant}.$

If F does not contain x explicitly:
→ multiply E eq by y' and note:

$$\frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) = y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \frac{\partial F}{\partial y'}y''$$

$$\rightarrow \quad y'\frac{\partial F}{\partial y} = y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \Rightarrow y''\frac{\partial F}{\partial y'} + y'\frac{\partial F}{\partial y} = \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right)$$

So if no explicit x dependence: → LHS $= \frac{d}{dx}F$

i.e. $dF = dy\frac{\partial F}{\partial y} + dy'\frac{\partial F}{\partial y'}$ by chain rule

$(dy' = d^2y)$. Hence $F = y'\frac{\partial F}{\partial y'} + C.$

Use these relations to solve 'variational' problems.

Solution method: Express problem in Functional form.

$$F[y] = \int_{\alpha}^{\beta} dx\, F(y, y', x)$$

on some interval $[\alpha, \beta]$. Apply Euler eq and use simplifications where appropriate. (i.e. inspect integrand of functional).

If Functional is constrained by another functional = constant

i.e/ $$J = \int_\alpha^\beta g(x,y,y')\,dx = \text{constant}$$

Define new functional $K = F + \lambda J$ and extremise this. (i.e find λ via constraint – method of Lagrange multipliers).

→ i.e, Lagrange's equations of motion
Fermat's principle in optics.

use this method.

S.L problems and calculus of variations

Considering the functional $F[y] = \int_\alpha^\beta [p(x)y'^2 + \sigma(x)y^2]\,dx$
for $p \neq 0$ for $\alpha < x < \beta$

constrained by $G[y] = 1 = \int_\alpha^\beta w(x)\,y^2\,dx$ $\quad w \neq 0$ for $\alpha < x < \beta$

→ Minimising ⇒ $\delta F - \lambda\,\delta G = 0$ (L. multipliers).

↓ Applying Euler equation to new functional $F - \lambda G$ yields S.L equation

$$-(py')' + \sigma y = \lambda w y. \iff \mathcal{L}y = \lambda w y.$$

i.e λ is the eigenvalue of the S.L equation.

let $\Lambda = \frac{F}{G}$ ⇒ $\delta\Lambda = \dfrac{G\,\delta F - F\,\delta G}{G^2} = \dfrac{1}{G}(\delta F - \Lambda\,\delta G)$

So for stationary Λ, $\delta F - \Lambda\,\delta G = 0$ i.e $\Lambda = \lambda$.

(we can do this as G = constant so extremising Λ is equivalent to extremising F).

So when Λ is stationary:

$$\lambda = \frac{\int_\alpha^\beta (p y'^2 + \sigma y^2) dx}{\int_\alpha^\beta w y^2 dx}$$

where y = eigenfunction i.e. solution to S.L equation. If we guess these we can ∴ guess λ by this method.

If we guess with an eigenfunction of 'adjustable' parameters we can find λ and then minimise w.r.t this parameter to find minimum or stationary λ. If we include more than one adjustable parameter minimise our eigenvalue by considering $\nabla \lambda = 0$. ~~(Thus 'vector space'~~ ~~of~~ Imagine λ as a function of parameters ....

Fin!