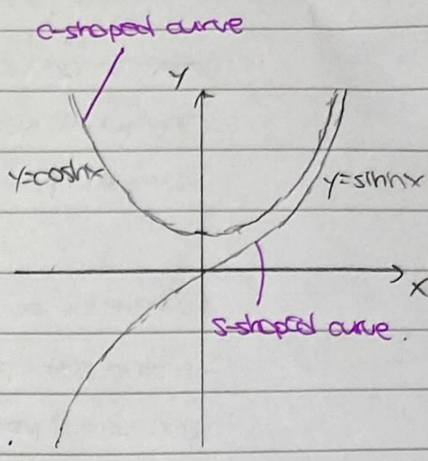
Functions and series

Hyperbolic functions

Definitions and properties.



- The hyperbolic functions are defined as:
  $\boxed{\sinh(x) = \frac{e^x - e^{-x}}{2}} = sh(x)$ // $\boxed{\cosh(x) = \frac{e^x + e^{-x}}{2}} = ch(x)$.
- $\sinh x$ is an odd function: $\sinh(-x) = -\sinh(x)$
  $\cosh x$ is an even function: $\cosh(-x) = \cosh(x)$.
- As $x \to \infty$, $\sinh(x) \approx \frac{1}{2}e^x$; $\cosh(x) \approx \frac{1}{2}e^x$
  As $x \to -\infty$, $\sinh(x) \approx -\frac{1}{2}e^{x}$; $\cosh(x) \approx \frac{1}{2}e^{x}$.
- $\sinh x$ is strictly increasing ; $\cosh x$ has a min. value of 1 at $x=0$.
- $\sinh x$ has a zero at $x=0$; $\cosh x$ is always +ve.
- $\cosh x + \sinh x = e^x$; $\cosh x - \sinh x = e^{-x}$.

Similarities with trigonometric functions.

- Osborne's rule states that any trigonometric identity gives rise to a corresponding hyperbolic identity when reversing the sign for each product of 2 $\sinh x$ terms.
- This is true because $\boxed{i\sinh x = \sin ix}$, $\boxed{\cosh x = \cos ix}$
  ↳ e.g.: $\cos^2 z + \sin^2 z = 1$
  $\cos^2(ix) + \sin^2(ix) = 1$
  $(\cosh x)^2 + (i\sinh x)^2 = 1$
  $\cosh^2 x - \sinh^2 x = 1$.
  ↳ e.g.: $\sin 2z = 2\sin z\cos z$
  $\sin(2ix) = 2\sin(ix)\cos(ix)$
  $\cancel{i}\sinh 2x = 2(\cancel{i}\sinh x)(\cosh x)$
  $\sinh 2x = 2\sinh x\cosh x$.
- other functions are defined by analogy w/ trigonometric functions.
  ↳ $\tanh x = \frac{\sinh x}{\cosh x}$
  ↳ $\coth x = \frac{\cosh x}{\sinh x}$
  ↳ $\operatorname{sech} x = \frac{1}{\sinh x}$
  ↳ $\operatorname{csch} x = \frac{1}{\cosh x}$

The graph is the original one reflected about $y=x$.

Inverse functions.

calculator returns the +ve (principal) value only.

- $\boxed{\sinh^{-1} x = \ln(x + \sqrt{x^2+1})}$; $\boxed{\cosh^{-1} x = \ln(x \pm \sqrt{x^2-1})}$
- Derivation as follows:

$y = \sinh^{-1} x \rightarrow x = \sinh y = \frac{e^y - e^{-y}}{2}$.
$2x = e^y - e^{-y}$
$e^{2y} - 2xe^y - 1 = 0$
$e^y = \frac{2x \pm \sqrt{4x^2+4}}{2}$
$= x \pm \sqrt{x^2+1}$.
As $\sqrt{x^2+1} > |x|$ and $e^y > 0$,
$\therefore y = \sinh^{-1} x = \ln(x + \sqrt{x^2+1})$.

$y = \cosh^{-1} x \rightarrow x = \cosh y = \frac{e^y + e^{-y}}{2}$
$2x = e^y + e^{-y}$
$e^{2y} - 2xe^y + 1 = 0$
$e^y = \frac{2x \pm \sqrt{4x^2-4}}{2}$
$= x \pm \sqrt{x^2-1}$.
$\therefore y = \cosh^{-1} x = \ln(x \pm \sqrt{x^2-1})$.
[Actually, $-\ln(x + \sqrt{x^2-1}) = \ln(x - \sqrt{x^2-1})$.]

Power series expansions.

Definition and convergence.

- The power series expansion for a function $f(x)$ about the pt. $x=0$ is the infinite sum of powers of x:

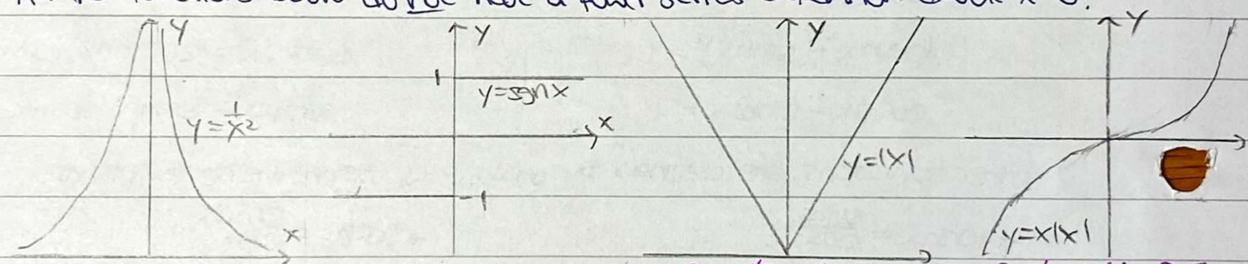
$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots$$

where $a_0, a_1, a_2 \ldots$ are constants

- A truncated power series is a good approximation for small x. ($x \approx 0$). The error can be estimated by the size of the first term ignored (error term).
- The approximation is only useful when the power series converges. Often power series only converge for small enough values of x, $|x|<r$, where r is the radius of convergence of the series.
- Whether a series converges or diverges can be determined using various tests (refer to supplementary notes).

Existence of a power series.

- For a valid power series to exist, the function should be continuous in the region of $x=0$, and all the derivatives $f'(x), f''(x), f'''(x)$ etc. must also be continuous at $x=0$.
- All the functions below do not have a power series expansion about $x=0$.



singularity ($f \to \infty$) | f discontinuous ($x=0$) | f continuous, f' discontinuous ($x=0$) | f, f' continuous, f'' discontinuous ($x=0$)

Taylor's Theorem

- If we have a function $f(x)$ and assume that it does have a power series expansion about $x=0$, i.e. $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots$, [#].

Taylor's theorem states that $\boxed{a_n = \frac{1}{n!}\frac{d^n f}{dx^n}\bigg|_{x=0} = \frac{1}{n!} f^{(n)}(0)}$

so $\boxed{f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots}$

- To find the nth coefficient a_n, simply differentiate eqn [#] n times:

$$f^{(n)}(x) = n(n-1)\ldots(2)(1)a_n + O(x).$$

so $\quad f^{(n)}(0) = n!\,a_n + 0 \quad \to \quad a_n = \frac{f^{(n)}(0)}{n!}$

* This only works if $f(x)$ is infinitely differentiable (smooth).

Important specialization of Taylor's Theorem: The Binomial Expansion.

- $(1+x)^\alpha = 1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\dots+\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n+\dots$

↳ If $\alpha \in \mathbb{Z}^+$, the series is finite and valid for all x.

↳ If α is fractional ($\mathbb{Z}^-$, the series is infinite and valid for $|x|<1$.

Generally, $\boxed{(a+bx)^n = a^n + na^{n-1}b+\frac{n(n-1)}{2!}a^{n-2}b^2+\dots}$, valid for $|bx|<a$).

- A particularly important and useful case of the binomial expansion is when $\alpha=-1$.

$(1+x)^{-1} = 1-x+x^2-x^3+x^4+\dots$ [geometric progression w/ $a=1$, $r=-x$].

Taylor's Theorem at other points (with shifted origin).

- To generate a power series of some function $f(x)$ about $x=a$ using Taylor's Theorem, we simply shift the origin to a. i.e. set $y=x-a \rightarrow x=a+y$., then

$$f(x) = f(a+y) = g(y).$$

Applying Taylor's Theorem to $g(y)$, to expand it in powers of y,

$$g(y) = g(0)+g'(0)y+\frac{g''(0)}{2!}y^2+\dots+\frac{g^{(n)}(0)}{n!}y^n+\dots \quad \text{small } y$$

Now, $g(0)=f(a)$, $g'(0)=f'(a)$, $g''(0)=f''(a)$ etc.

$\therefore g(y) = f(a+y) = f(a)+f'(a)y+\frac{f''(a)}{2!}y^2+\dots+\frac{f^{(n)}(a)}{n!}y^n+\dots$ small y.

Substituting back $y=x-a$ (or $x=a+y$).

$$\boxed{f(x) = f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\dots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\dots} \quad x\approx a.$$

- As $f(a+y) = f(a)+f'(a)y+\frac{f''(a)}{2!}y^2+\dots+\frac{f^{(n)}(a)}{n!}y^n$, we can rename the variables: $a\rightarrow x$, $y\rightarrow \Delta x$. Then,

$$\boxed{f(x+\Delta x) = f(x)+f'(x)\Delta x+\frac{f''(x)}{2!}\Delta x^2+\dots+\frac{f^{(n)}(x)}{n!}\Delta x^n+\dots} \quad \text{small } \Delta x.$$

(useful for approximating small changes of a function).

* For $\sin x$ or $\cos x$, we can directly put $x=a+y$ and expand using the compound angle identity: $\sin(a+y)=\sin a\cos y+\cos a\sin y$; $\cos(a+y)=\cos a\cos y-\sin a\sin y$., before using the standard series of $\sin y$ / $\cos y$ and substituting back $y=a-x$.

Standard series (in databook). [valid for real and complex no.].

- $e^z = 1+z+\frac{z^2}{2!}+\frac{z^3}{3!}+\dots = \sum_{n=0}^{\infty}\frac{z^n}{n!}$
- $\sin z = z-\frac{z^3}{3!}+\frac{z^5}{5!}+\dots = \sum_{n=0}^{\infty}(-1)^n\frac{z^{2n+1}}{(2n+1)!}$; $\cos z = 1-\frac{z^2}{2!}+\frac{z^4}{4!}+\dots+\sum_{n=0}^{\infty}(-1)^n\frac{z^{2n}}{(2n)!}$
- $\sinh z = z+\frac{z^3}{3!}+\frac{z^5}{5!}+\dots = \sum_{n=0}^{\infty}\frac{z^{2n+1}}{(2n+1)!}$; $\cosh z = 1+\frac{z^2}{2!}+\frac{z^4}{4!}+\dots+\sum_{n=0}^{\infty}\frac{z^{2n}}{(2n)!}$

convergent for all z

- $\tan z = z+\frac{1}{3}z^3+\frac{2}{15}z^5+\frac{17}{315}z^7+\dots$ convergent for all $|z|<\frac{\pi}{2}$.
- $\ln(1+z) = z-\frac{z^2}{2}+\frac{z^3}{3}+\dots = \sum_{n=0}^{\infty}(-1)^n\frac{z^{n+1}}{n+1}$ converges on/within circle $|z|=1$ except $z=-1$
- $(1+z)^\alpha = 1+\alpha z+\frac{\alpha(\alpha-1)}{2!}z^2+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}z^3+\dots$ converges for all $|z|<1$.

Manipulating power series

- If 2 power series converge for the same range, then they can be manipulated in a variety of ways w/o problems and the series obtained converges for at least that range.
- Suppose 2 series converge for $|x|<r$, then for x in this range:

① Add them term by term.

↳ $(a_0+a_1x+a_2x^2+\ldots)+(b_0+b_1x+b_2x^2+\ldots)$

$= (a_0+b_0)+(a_1+b_1)x+(a_2+b_2)x^2+\ldots$

↳ eg: $\frac{1}{1+x}+\frac{1}{1-x}=\frac{2}{1-x^2}$

$\frac{1}{1+x}=1-x+x^2-x^3+\ldots+(-1)^nx^n+\ldots$

$\frac{1}{1-x}=1+x+x^2+x^3+\ldots+x^n+\ldots$

$\rightarrow \frac{2}{1-x^2}=2(1+x^2+x^4+\ldots)$.

② Rearrange them at will

↳ $a_0+a_1x+a_2x^2+a_3x^3+\ldots$ even terms odd terms.

$=(a_0+a_2x^2+a_4x^4+\ldots)+(a_1x+a_3x^3+a_5x^5+\ldots)$

↳ e.g: $e^x=1+x+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+\ldots$

$=\left(1+\frac{x^2}{2!}+\frac{x^4}{4!}+\ldots\right)+\left(x+\frac{x^3}{3!}+\frac{x^5}{5!}+\ldots\right)$

$=\cosh x+\sinh x$.

③ Multiply them term by term.

↳ $(a_0+a_1x+a_2x^2+\ldots)(b_0+b_1x+b_2x^2+\ldots)$

$=a_0b_0+(a_0b_1+b_0a_1)x+(a_0b_2+a_1b_1+a_2b_0)x^2+\ldots$

↳ eg: $\sin x\cos x$.

$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\ldots$ and $\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\ldots$

$\sin x\cos x=\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}-\ldots\right)\left(1-\frac{x^2}{2!}+\frac{x^4}{4!}-\ldots\right)$

$=x+x^3\left(-\frac{1}{2!}-\frac{1}{3!}\right)+x^5\left(\frac{1}{4!}+\frac{1}{3!}\frac{1}{2!}+\frac{1}{5!}\right)+\ldots$

$=x+x^3\left(-\frac{3+1}{3!}\right)+x^5\left(\frac{5+10+1}{5!}\right)+\ldots$

$=\frac{1}{2}\left(2x-\frac{1}{3!}(2x)^3+\frac{1}{5!}(2x)^5+\ldots\right)=\frac{1}{2}\sin 2x$

④ Differentiate or integrate them term by term.

↳ $\frac{d}{dx}(a_0+a_1x+a_2x^2+\ldots)$

$=a_1+2a_2x+3a_3x^2+4a_4x^3+\ldots$

↳ e.g: $\int\frac{1}{1+x}dx=\ln(1+x)$.

$\frac{1}{1+x}=1-x+x^2-x^3+\ldots+(-1)^nx^n+\ldots$

$\ln(1+x)=\int(1-x+x^2-x^3+\ldots+(-1)^nx^n+\ldots)dx$

$=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4!}+\ldots+C$.

As $\ln(1+0)=0$ $\quad C=0$

$\rightarrow \ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots$

Big O notation and limits

Big O notation.

- When approximating using a truncated Taylor's series, we should keep track of the size of the terms we are ignoring by using big O notation.
- $O(x^n)$ denotes all the terms that are of order n or above.
- Properties of big O notation:
  - ↳ $x\,O(x^n) = O(x^{n+1})$ or more generally, $\boxed{O(x^m)\,O(x^n) = O(x^{m+n})}$
  - ↳ $\boxed{O(x^n) + c\,O(x^n) = O(x^n)}$ where c is any constant $[O(x^n) \overset{c=-1}{-} O(x^n) = O(x^n)]$.
  - ↳ $\boxed{O(x^n) + O(x^m) = O(x^m)}$ for +ve integers $m \le n$.

- e.g: show that $\frac{1}{\sin x} = \frac{1}{x} + O(x)$

$$\frac{1}{\sin x} = \frac{1}{x + O(x^3)} = \frac{1}{x(1+O(x^2))}$$
$$= [x(1+O(x^2))]^{-1}$$
$$= x^{-1}[1 - O(x^2) + [O(x^2)]^2].$$
$$= \frac{1}{x}(1 + O(x^2))$$
$$= \frac{1}{x} + O(x).$$

Limits.

- Rough defn: If as x gets closer to a, a function f(x) get closer and closer to a value α, we say that α is the limit of the function f(x) as x approaches a.

$$\lim_{x \to a} f(x) = \alpha.$$

- e.g: Find $\lim_{x\to 0}\left(\frac{1}{x^n}\left(\sin x - \frac{x}{1-ax^2}\right)\right)$ where $a \neq -1/6$.

$\sin x = x - \frac{x^3}{6} + O(x^5).$

$$-\frac{x}{1-ax^2} = -x \cdot \frac{1}{1-ax^2} = -x(1+ax^2+a^2x^4+\ldots)$$
$$= -x(1+ax^2+O(x^4))$$

$$\text{so } \sin x - \frac{x}{1-ax^2} = \cancel{x} - \frac{x^3}{6} + O(x^5) - \cancel{x} - ax^3 + O(x^5)$$
$$= -x^3\left(a+\frac{1}{6}\right) + O(x^5).$$

and $\frac{1}{x^n}\left(\sin x - \frac{x}{1-ax^2}\right) = -x^{3-n}\left(a+\frac{1}{6}\right) + O(x^{5-n})$

case (i) : $n<3$, $\lim_{x\to 0}\left(\frac{1}{x^n}\left(\sin x - \frac{x}{1-ax^2}\right)\right) = 0.$

case (ii) : $n=3$, $\lim_{x\to 0}\left(\frac{1}{x^n}\left(\sin x - \frac{x}{1-ax^2}\right)\right) = -\left(a+\frac{1}{6}\right)$

case (iii) : $n>3$, $\lim_{x\to 0}\left(\frac{1}{x^n}\left(\sin x - \frac{x}{1-ax^2}\right)\right) = \pm\infty$ (sign of $-\left(a+\frac{1}{6}\right)$).

*Note if $a=-1/6$, $\lim_{x\to 0}\left(\frac{1}{x^n}\left(\sin x - \frac{x}{1-ax^2}\right)\right) = 0$ for all n. this assumes $\lim_{x\to 0^+}$.

Using de L'Hôpital's rule to evaluate limits.

- de L'Hôpital's rule states that for indeterminate $\frac{f(x)}{g(x)}$,

$$\boxed{\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)}}.$$

where indeterminate forms include $\boxed{\frac{0}{0}, \frac{\infty}{\infty}}$ $0\cdot\infty$, $\infty-\infty$, 0^0, 1^∞, ∞^0.

- Proof (special case where $a=0$, $f(0)=0$, $g(0)=0$, $g'(0)\neq 0$).

$$\frac{f(x)}{g(x)} = \frac{\cancel{f(0)}^{0} + xf'(0) + O(x^2)}{\cancel{g(0)}^{0} + xg'(0) + O(x^2)} = \frac{xf'(0)+O(x^2)}{xg'(0)+O(x^2)}.$$

$$= \frac{f'(0)+O(x)}{g'(0)+O(x)}.$$

$$\lim_{x\to 0}\frac{f(x)}{g(x)} = \frac{f'(0)}{g'(0)}. \qquad \text{since } \lim_{x\to 0}[O(x)] = 0.$$

Miscellaneous limits. (in databook).

- $\lim_{n\to\infty} n^s x^n = 0$. * if $|x|<1$ and $s\in\mathbb{R}$.
- $\lim_{n\to\infty} \frac{x^n}{n!} = 0$ for all x.
- $\lim_{n\to\infty}(1+\frac{x}{n})^n = e^x$ for all x. (take ln on both sides then find the limit).
- $\lim_{x\to 0} x^s \ln x = 0$ * for $s>0$.
- In general, "exponentials always win over powers" and "everything wins over logs".

- e.g.: $\lim_{x\to 2\pi}\frac{(x-2\pi)^2}{\sin^2 x}$. [change of variable so new variable $\to 0$].

We can generate a power series for $\sin x$ about $x=2\pi$, but it is easier to set $y = x-2\pi$, where y is small.

$$\lim_{x\to 2\pi}\frac{(x-2\pi)^2}{\sin^2 x} = \lim_{y\to 0}\frac{y^2}{\sin^2(y+2\pi)}$$

$$= \lim_{y\to 0}\frac{y^2}{\sin^2 y} = 1^2 = 1.$$

- e.g.: Find an approximation in powers of x for $\exp[\sqrt{1+x}-1]$, neglecting $O(x^3)$.

$\sqrt{1+x} = (1+x)^{1/2} = 1+\frac{1}{2}x-\frac{1}{8}x^2+O(x^3)$.

$\exp[\sqrt{1+x}-1] = \exp[\frac{1}{2}x-\frac{1}{8}x^2+O(x^3)]$. $\to$ set $y = \frac{1}{2}x-\frac{1}{8}x^2+O(x^3)$.

For small x, y is small $\to$ req. for using Taylor expansion of e^y.

$e^y = 1+y+\frac{y^2}{2!}+O(y^3)$.

$$e^{\sqrt{1+x}-1} = 1+(\tfrac{1}{2}x-\tfrac{1}{8}x^2+O(x^3)) + \tfrac{1}{2}(\tfrac{1}{2}x-\tfrac{1}{8}x^2+O(x^3))^2 + O[(\tfrac{1}{2}x-\tfrac{1}{8}x^2+O(x^3))^3]$$

$$= 1+\tfrac{1}{2}x\cancel{-\tfrac{1}{8}x^2}+O(x^3)\cancel{+\tfrac{1}{8}x^2}+O(x^3)+O(x^3).$$

$$= 1+\tfrac{1}{2}x+O(x^3).$$

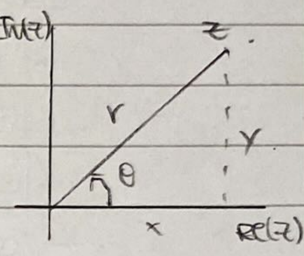
Complex numbers

Fundamentals of complex numbers

Basic properties.

- A general complex no. z can be represented as $x+iy$, where $i=\sqrt{-1}$.
- Addition: $z_1+z_2=(x_1+iy_1)+(x_2+iy_2)$
  $= (x_1+x_2)+i(y_1+y_2)$
- Subtraction: $z_1-z_2=(x_1+iy_1)-(x_2+iy_2)$
  $=(x_1-x_2)+i(y_1-y_2)$

  same as vectors
- Multiplication: $z_1z_2=(x_1+iy_1)(x_2+iy_2)$
  $=(x_1x_2-y_1y_2)+i(x_1y_2+x_2y_1)$
- Division: $\frac{z_1}{z_2}=\frac{x_1+iy_1}{x_2+iy_2}=\frac{x_1+iy_1}{x_2+iy_2}\cdot\frac{x_2-iy_2}{x_2-iy_2}$
  $=\frac{(x_1x_2+y_1y_2)+i(y_1x_2-x_1y_2)}{x_2^2+y_2^2}$



- We can represent z graphically on an Argand diagram:
  - ↳ In cartesian form: $z=x+iy$
    where x is the real part $Re(z)$ and y is the imaginary part $Im(z)$
  - ↳ In polar form: $z=r(\cos\theta+i\sin\theta)$
    where r is the modulus $|z|$ and θ is the argument $\arg(z)$.
- Converting between cartesian (x,y) and polar (r,θ) form of z:
  - ↳ $x=r\cos\theta$, $y=r\sin\theta$ // $r=\sqrt{x^2+y^2}$, $\theta=\arctan\left(\frac{y}{x}\right)\ (+\pi)$.

Complex conjugation.

- The complex conjugate of $z=x+iy$ is $\bar{z}=x-iy$, and is denoted by $\bar{z}$ or z^*.
- On the Argand diagram, complex conjugation is equivalent to a reflection in the real axis.
- Useful properties of complex conjugation:
  - ↳ $z\bar{z}=|z|^2$, which is a real no. ↳ If $z=\bar{z}$, z is real
  - ↳ $\boxed{Re(z)=\frac{z+\bar{z}}{2},\ Im(z)=\frac{z-\bar{z}}{2}}$ ↳ $|\bar{z}|=|z|$
  - ↳ $\boxed{\overline{z_1\pm z_2}=\bar{z}_1\pm\bar{z}_2,\ \overline{z_1z_2}=\bar{z}_1\bar{z}_2,\ \overline{\left(\frac{z_1}{z_2}\right)}=\frac{\bar{z}_1}{\bar{z}_2}}$

Euler's formula.

- Euler's formula states that $\boxed{e^{i\theta}=\cos\theta+i\sin\theta}$.
- Proof: $e^{i\theta}=1+i\theta+\frac{(i\theta)^2}{2!}+\frac{(i\theta)^3}{3!}+\frac{(i\theta)^4}{4!}+\dots$
  $=\left(1-\frac{\theta^2}{2!}+\frac{\theta^4}{4!}+\dots\right)+i\left(\theta-\frac{\theta^3}{3!}+\frac{\theta^5}{5!}+\dots\right)$
  $=\cos\theta+i\sin\theta$

Exponential form

- As we can express z in polar form as $z = r(\cos\theta + i\sin\theta)$, using Euler's formula, we can express z in exponential form $z = re^{i\theta}$.
- Multiplication: $z_1 z_2 = (r_1 e^{i\theta_1})(r_2 e^{i\theta_2})$ $= r_1 r_2 e^{i(\theta_1+\theta_2)}$] $|z_1 z_2| = |z_1||z_2|$, $\arg(z_1 z_2) = \arg(z_1) + \arg(z_2)$
- Division: $\frac{z_1}{z_2} = \frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}}$ $= \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}$] $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$, $\arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2)$.
- The geometrical interpretation of multiplying a complex no. w by another one $z = re^{i\theta}$ has the effect of stretching w by a factor of r and rotating w by θ on the Argand diagram.
- Useful identities:
  - ↳ $\boxed{\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}}$, $\boxed{\sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}}$
  - ↳ $\cos\theta = \mathrm{Re}(e^{i\theta})$, $\sin\theta = \mathrm{Im}(e^{i\theta})$.

de Moivre's Theorem

- As $e^{i(n\theta)} = (e^{i\theta})^n$,

$$\boxed{\cos(n\theta) + i\sin(n\theta) = (\cos\theta + i\sin\theta)^n}$$

(Also can be seen easily using a geometric interpretation).

Principal value of θ.

- For any integer n,

$$e^{2\pi i n} = \cos(\overset{1}{\cancel{2\pi n}}) + i\sin(\overset{0}{\cancel{2\pi n}}) = 1$$

→ $z = re^{i\theta} = re^{i\theta}e^{2\pi i n} = re^{i(\theta+2n\pi)}$ [we can add $2n\pi$ to θ w/o affecting the value of z].

- To avoid ambiguity, we use the principal value of θ, w/ θ restricted to

$$-\pi < \theta \leq \pi$$

Locus.

- In general, to find the locus, use $z = x+iy$ and $|z| = \sqrt{x^2+y^2}$.
- Common loci
  - ↳ $|z-a| = |z-b|$: $\perp$ bisector between a and b.
  - ↳ $\arg(z-a) = \theta$: Half line from a w/ angle θ from the real axis (not inc. a).
  - ↳ $|z-a| = r$: circle w/ radius r and centre a.
  - ↳ $z = a + re^{i\theta}$: circle w/ radius r and centre a.

  ] other forms of circle: $|z| = 2|z-3|$, $z\bar{z} - 4(z+\bar{z}) + 12 = 0$.

Applications of complex numbers.

Evaluating nth roots.

- To find the nth roots of a complex no. $z = re^{i\theta}$:

$$\boxed{z^{1/n} = (re^{i\theta} \cdot e^{i2k\pi})^{1/n} = r^{1/n} e^{i\left(\frac{\theta + 2k\pi}{n}\right)}} \quad , \quad k = 0, 1, \ldots n-1$$

so we have n roots.

- The roots are the vertices of a regular n-gon.
- The sum of roots is 0, the product of roots is ± 1 (odd +, even −).
  - ↳ proof: $z^n - e^{i\theta} = 0$ → apply sum/product of root formula.

Polynomials with real coefficients.

- If $p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$, where $a_i \in \mathbb{R}$.

  Then the roots of $p(z) = 0$ are either real or complex conjugate pairs.

- Proof: Assume z_1 is a root of $p(z) = 0$.

$$a_n z_1^n + a_{n-1} z_1^{n-1} + \ldots + a_1 z_1 + a_0 = 0.$$

$$\overline{a_n z_1^n + a_{n-1} z_1^{n-1} + \ldots + a_1 z_1 + a_0} = \bar{0} \qquad \text{Taking complex conjugate on both sides.}$$

$$\overline{a_n z_1^n} + \overline{a_{n-1} z_1^{n-1}} + \ldots + \overline{a_1 z_1} + \bar{a}_0 = 0 \qquad \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$$

$$\bar{a}_n \overline{z_1^n} + \bar{a}_{n-1} \overline{z_1^{n-1}} + \ldots + \bar{a}_1 \bar{z}_1 + \bar{a}_0 = 0 \qquad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$$

$$a_n (\bar{z}_1)^n + a_{n-1} (\bar{z}_1)^{n-1} + \ldots + a_1 \bar{z}_1 + a_0 = 0 \qquad \bar{a}_i = a_i \text{ as } a \in \mathbb{R},\ \overline{z^n} = (\bar{z})^n.$$

→ $\bar{z}_1$ is a root.

so either $z_1 = \bar{z}_1$ ($z \in \mathbb{R}$) or z_1 and $\bar{z}_1$ are a complex conjugate pair.

Complex trigonometric and hyperbolic functions.

- $\cosh z = \dfrac{e^z + e^{-z}}{2}$, setting $z = iw$, $\cosh(iw) = \dfrac{e^{iw} + e^{-iw}}{2} = \cos(w)$.

  $\sinh z = \dfrac{e^z - e^{-z}}{2}$. setting $z = iw$, $\sinh(iw) = \dfrac{e^{iw} - e^{-iw}}{2} = i\sin(w)$

- $\cos z = \dfrac{e^{iz} + e^{-iz}}{2}$, setting $z = iw$, $\cos(iw) = \dfrac{e^{-w} + e^{w}}{2} = \cosh(w)$

  $\sin z = \dfrac{e^{iz} - e^{-iz}}{2i}$. setting $z = iw$, $\sin(iw) = \dfrac{e^{-w} - e^{w}}{2i} = i\sinh(w)$

→ $\boxed{\cosh(iw) = \cos(w)\ ;\ \cos(iw) = \cosh(w)}$ // $\boxed{\sinh(iw) = i\sin(w)\ ;\ \sin(iw) = i\sinh(w)}$

- e.g.: What are the real and imaginary parts of $\sin(x+iy)$.

$$\sin(x+iy) = \sin x \cos iy + \cos x \sin iy$$

$$= \sin x \cosh y + \cos x\, (i \sinh y)$$

$$= \sin x \cosh y + i \cos x \sinh y.$$

→ real part: $\sin x \cosh y$; imaginary part: $\cos x \sinh y$.

- eg: Find $\sin^{-1}(2)$.

$\sin^{-1}(2)$ must be complex as $2 \notin [-1,1]$. → Let $\sin^{-1}(2) = x+iy$.

→ $2 = \sin(x+iy) = \sin x \cosh y + i\cos x \sinh y$.

Comparing real and imaginary parts,

Re: $\sin x \cosh y = 2$ [1]

Im: $\cos x \sinh y = 0$. [2].

From [2]; $\cos x = 0$ or $\sinh y = 0$

- case (a): $\sinh y = 0$

then $y=0 \to \cosh y = 1 \to$ eqn [1] becomes $\sin x = 2$ ⇒ reject $\sinh y = 0$.

- case (b): $\cos x = 0$.

then $x = \pm\frac{\pi}{2} + 2k\pi \to \sin x = \pm 1 \to$ eqn [1] becomes $\pm\cosh y = 2$.

We now have $\cosh y = \pm 2$, but $\cosh y > 0$ for $y \in \mathbb{R}$.

so $\cosh y = 2 \Rightarrow \sin x = 1$

∴ $x = \frac{\pi}{2} + 2k\pi$, $y = \pm \operatorname{arcosh}(2)$

and $\sin^{-1}(2) = (\frac{\pi}{2} + 2k\pi) \pm \operatorname{arcosh}(2)$.

Complex natural logarithms.

- $\ln z = \ln(re^{i\theta}) = \ln r + \ln e^{i\theta} = \ln r + i\theta$.

More generally, $\boxed{\ln z = \ln r + i(\theta + 2k\pi)}$ $k \in \mathbb{Z}$.

- For complex w and z,

$$z^w = (e^{\ln z})^w = e^{w \ln z}.$$

Trigonometric identities.

① Multiple angles. → useful for solving trigonometric eqn.

- e.g: Express $\cos(3\theta)$ and $\sin(3\theta)$ in terms of $\cos\theta$ and $\sin\theta$.

$\cos(3\theta) + i\sin(3\theta) = (\cos\theta + i\sin\theta)^3$

$= \cos^3\theta + 3\cos^2\theta\sin\theta\, i - 3\cos\theta\sin^2\theta - \sin^3\theta\, i$

→ Comparing Re/Im parts : $\cos(3\theta) = \cos^3\theta - 3\cos\theta\sin^2\theta$; $\sin(3\theta) = 3\cos^2\theta\sin\theta - \sin^3\theta$.

② Powers of $\sin\theta$ and $\cos\theta$ → useful for integration.

- eg: Express $\cos^3\theta$ in terms of multiple angles.

Let $z = \cos\theta + i\sin\theta$. $\cos^3\theta = (\cos\theta)^3 = \left(\frac{z+z^{-1}}{2}\right)^3$

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

$= \frac{1}{4}\left(\frac{z^3+z^{-3}}{2}\right) + \frac{3}{4}\left(\frac{z+z^{-1}}{2}\right)$

$= \frac{1}{4}\cos(3\theta) + \frac{3}{4}\cos\theta$.

Trigonometric integrals.

- As $\cos\theta = Re(e^{i\theta})$ and $\sin\theta = Im(e^{i\theta})$, we can solve integrals in the form $\int e^{\alpha\theta}\cos\beta\theta d\theta$ or $\int e^{\alpha\theta}\sin\beta\theta d\theta$ using a shortcut:

$$\int e^{\alpha\theta}\cos\beta\theta d\theta = Re\left(\int e^{\alpha\theta + i\beta\theta}d\theta\right) \quad ; \quad \int e^{\alpha\theta}\sin\beta\theta d\theta = Im\left(\int e^{\alpha\theta+i\beta\theta}d\theta\right)$$

- eg: Find $\int_0^\pi e^{-2x}\cos x dx$ and $\int_0^\pi e^{-2x}\sin x dx$.

Consider $\tilde{I} = \int_0^\pi e^{-2x}e^{ix}dx$. $Re(\tilde{I}) = \int_0^\pi e^{-2x}\cos x dx$ and $Im(\tilde{I}) = \int_0^\pi e^{-2x}\sin x dx$.

$$\begin{aligned}
\tilde{I} &= \int_0^\pi e^{x(i-2)}dx \\
&= \left[\frac{1}{i-2}e^{x(i-2)}\right]_0^\pi \\
&= \frac{1}{i-2}\left[e^{\pi(i-2)} - 1\right] \\
&= \frac{1}{i-2}\left[\cancel{e^{i\pi}}^{-1}\cdot e^{-2\pi} - 1\right] \\
&= \frac{e^{-2\pi}+1}{2-i} \\
&= (e^{-2\pi}+1)\cdot\frac{2+i}{2^2+1^2}
\end{aligned}$$

$\rightarrow Re(\tilde{I}) = \frac{2}{5}(e^{-2\pi}+1)$, $Im(\tilde{I}) = \frac{1}{5}(e^{-2\pi}+1)$.

Series.

- To sum a series, we can generally consider $C+iS$ and take the real part to find C and the imaginary part to find S.

- eg: Find $C = \cos\theta + \cos3\theta + \cos5\theta + \ldots + \cos(2N+1)\theta$.

Consider $S = \sin\theta + \sin3\theta + \sin5\theta + \ldots + \sin(2N+1)\theta$.

$$\begin{aligned}
C+iS &= e^{i\theta} + e^{i3\theta} + e^{i5\theta} + \ldots + e^{i(2N+1)\theta} \\
&= \frac{e^{i\theta}(1-(e^{i2\theta})^{N+1})}{1-e^{i2\theta}} \\
&= \frac{e^{i\theta}(1-e^{i2(N+1)\theta})}{e^{i\theta}(e^{-i\theta}-e^{i\theta})} \\
&= \frac{1-e^{i2(N+1)\theta}}{-2i\sin\theta} \\
&= \frac{i(1-e^{i2(N+1)\theta})}{2\sin\theta}
\end{aligned}$$

$1-e^{2i\theta} = e^{i\theta}(e^{-i\theta}-e^{i\theta})$

$\frac{e^{i\theta}-e^{-i\theta}}{2i} = \sin\theta$

$\rightarrow C = Re\left(\frac{i(1-e^{i2(N+1)\theta})}{2\sin\theta}\right) = -\frac{i(i\sin(2(N+1)\theta))}{2\sin\theta} = \frac{\sin 2(N+1)\theta}{2\sin\theta}$.

* Instead of using $1-e^{2i\theta} = e^{i\theta}(e^{-i\theta}-e^{i\theta})$, we can multiply the top and bottom by the complex conjugate of $1-e^{2i\theta}$, $1-e^{-2i\theta}$.

$\Rightarrow$ In general, setting $\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}$ or $\sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2}$ works as well, but $\cos\theta = Re(e^{i\theta})$ and $\sin\theta = Im(e^{i\theta})$ is much faster.

Complex impedance.

- To avoid confusion w/ i which represents the current, we use $j = \sqrt{-1}$.
- Consider a complex voltage $V(t) = V_0 e^{j\omega t}$ and complex current $i(t) = Ie^{j(\omega t + \phi)} = I_0 e^{j\omega t}$

For a resistor, $V = iR$. $\rightarrow \frac{V}{i} = Z_R = R$.

For a capacitor, $i = C\frac{dV}{dt} = j\omega C V_0 e^{j\omega t} = j\omega C V \rightarrow \frac{V}{i} = Z_C = \frac{1}{j\omega C}$

For an inductor, $V = L\frac{di}{dt} = j\omega L I_0 e^{j\omega t} = j\omega L i \rightarrow \frac{V}{i} = Z_L = j\omega L$.

- If we want to find the physical voltage $V_{phys}(t)$ and physical current $i_{phys}(t)$, we simply take the real part of the complex voltage $V(t)$ and complex current $i(t)$.

i.e. $V_{phys}(t) = Re(V(t))$; $i_{phys}(t) = Re(i(t))$

- Using complex no, we can solve AC circuit problems w/o explicitly setting up differential eqns $\rightarrow$ they become algebraic eqns.

(For a RLC circuit, we would have a 2nd order ODE in Q).

* Note $\frac{d}{dt}(Re(\,)) = Re(\frac{d}{dt}(\,))$.

Differential equations

Differential equations

Definitions and notation.

- An ordinary differential eqn (ODE) has a dependent variable as a function of only 1 independent variable (e.g. $y(x)$). The general form of an ODE is:

$$f(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \frac{d^3y}{dx^3}, \ldots) = 0.$$

- A partial differential eqn (PDE) has a dependent variable as a function of more than 1 independent variable (e.g. $f(x,y,z,t)$).
- Whenever any of the y terms ($y, \frac{dy}{dx}, \frac{d^2y}{dx^2}$ etc.) are multiplied together, squared, or functioned (e.g. $\sin y$), the eqn. is non-linear.
- If the variable is a function of space, $y' = \frac{dy}{dx}$, $y'' = \frac{d^2y}{dx^2}$.
- If the variable is a function of time, $\dot{x} = \frac{dx}{dt}$, $\ddot{x} = \frac{d^2x}{dt^2}$.

First-order differential equations

First-order equations.

- The general form for first-order ODEs is:

$$g(x, y, \frac{dy}{dx}) = 0 \quad \text{or} \quad \frac{dy}{dx} = f(x,y).$$

- Irrespective of the form of the first-order ODE, 1 boundary condition is always req. since solving the eqn. is equivalent to integrating once. (to find the PS).

Direct integration.

- We can use the direct integration method for ODEs in the form.

$$\frac{dy}{dx} = f(x).$$

- To solve ODEs in this form:

$$y = \int f(x)dx.$$

Separable equations.

- We can use the separation of variables method for ODEs in the form

$$\frac{dy}{dx} = \frac{g(x)}{h(y)}.$$

- To solve ODEs in this form:

$$\int h(y)dy = \int g(x)dx$$

Integrating factor.

- We can use the integrating factor method for ODEs in the form.

$$\frac{dy}{dx} + P(x)y = q(x).$$

- To solve ODEs in this form:

$$\boxed{e^{\int P(x)dx}}\left(\frac{dy}{dx} + P(x)y\right) = e^{\int P(x)dx} q(x)$$

multiply both sides by integrating factor $e^{\int P(x)dx}$

$$\frac{d}{dx}\left(e^{\int P(x)dx} y\right) = e^{\int P(x)dx} q(x).$$

DB $$e^{\int P(x)dx} y = \int e^{\int P(x)dx} q(x)dx.$$

Reducible form / substitution.

- We can use the substitution method for ODEs in the form.

$$\frac{dy}{dx} = f\left(\frac{y}{x}\right).$$

we could use other substitutions, only do this if given in the Q.

- To solve ODEs in this form:

use the substitution $u = \frac{y}{x}$. (so $y = ux$, $\frac{dy}{dx} = x\frac{du}{dx} + u$)

→ we get separable eqn: $x\frac{du}{dx} = f(u) - u$

After solving the separable ODE, substitute back $u = \frac{y}{x}$.

- eg: $\frac{dy}{dx} = \frac{x^2 - xy + y^2}{x^2}$. use $u = \frac{y}{x}$. Given $y(1) = 0$.

$$\frac{dy}{dx} = \frac{x^2 - xy + y^2}{x^2}$$

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

Using $u = \frac{y}{x}$, $y = ux$, $\frac{dy}{dx} = x\frac{du}{dx} + u$

$$x\frac{du}{dx} + u = 1 - u + u^2$$

$$x\frac{du}{dx} = 1 - 2u + u^2$$

$$\int \frac{1}{(1-u)^2} du = \int \frac{1}{x} dx$$

$$\frac{1}{1-u} = \ln x + C.$$

Substituting back $u = \frac{y}{x}$. $$\frac{1}{1-\frac{y}{x}} = \ln x + C.$$

Using $y(1) = 0$. $$\frac{1}{1-\frac{0}{1}} = \ln 1 + C \quad \rightarrow C = 1.$$

$$\frac{1}{1-\frac{y}{x}} = \ln x + 1$$

$$y = x\left(1 - \frac{1}{1 + \ln x}\right).$$

* We can use the specific substitution when x and y appear on the RHS as a ratio of terms which only involves sums of powers. Also the sums of orders wrt x and y for each term in the numerator and denominator.

Second-order differential equation

Second-order equation.

- The general form for second-order ODE is :

$$g\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}\right) = 0 \quad \text{or} \quad f\left(y, \frac{dy}{dx}, \frac{d^2y}{dx^2}\right) = g(x).$$

- If $g(x) = 0$, then the ODE is homogeneous.
- Irrespective of the form of the second-order ODE, 2 boundary conditions are always req. since solving the eqn. is equivalent to integrating twice. (to find the PS)

Linear with constant coefficients.

- We can solve ODEs in the form:

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x) \qquad \text{where } a, b, c \text{ are constants.}$$

- The general sol'n (GS) is made up of 2 parts – the complementary function (CF) and the particular integral (PI) [GS = CF + PI].
  - ↳ substituting CF into the LHS gives 0. (CF is the GS when $f(x) = 0$) homogeneous
  - ↳ substituting PI into the LHS gives $f(x)$

Finding the complementary function.

- Initially, consider the homogeneous case : $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$.

Using the trial sol'n $y = Ae^{\lambda x}$, → $\frac{dy}{dx} = \lambda Ae^{\lambda x}$, $\frac{d^2y}{dx^2} = \lambda^2 Ae^{\lambda x}$.

so $\quad a(\lambda^2 Ae^{\lambda x}) + b(\lambda Ae^{\lambda x}) + c(Ae^{\lambda x}) = 0$

$$Ae^{\lambda x}(a\lambda^2 + b\lambda + c) = 0.$$

As $Ae^{\lambda x} \neq 0$, $\qquad a\lambda^2 + b\lambda + c = 0$. (Auxiliary equation AE).

- A property of linear, homogeneous ODEs is that sol'ns may be superposed. i.e. if y_1 and y_2 are both sol'ns to the eqn, then a linear combination of y_1 and y_2 is also a sol'n. [$\alpha y_1 + \beta y_2$ is a sol'n].

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = a\frac{d^2}{dx^2}(\alpha y_1 + \beta y_2) + b\frac{d}{dx}(\alpha y_1 + \beta y_2) + c(\alpha y_1 + \beta y_2)$$

$$= \alpha\left(a\frac{d^2y_1}{dx^2} + b\frac{dy_1}{dx} + cy_1\right) + \beta\left(a\frac{d^2y_2}{dx^2} + b\frac{dy_2}{dx} + cy_2\right) = 0.$$

- The ODE is 2nd order → superpose 2 sol'ns : $y = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$.
- There are 3 cases for the auxiliary eqn:
  - ↳ 2 distinct real roots λ_1, λ_2 → $\boxed{y = Ae^{\lambda_1 x} + Be^{\lambda_2 x}}$
  - ↳ 2 repeated real roots λ → $\boxed{y = Ae^{\lambda x} + xBe^{\lambda x}}$
  - ↳ 2 complex conjugate roots $\lambda \pm \omega i$ → $\boxed{y = e^{\lambda x}(A\sin\omega t + B\cos\omega t)}$

based on the original form $y = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$

Finding the particular integral

- Now, consider the non-homogeneous case: $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$.
- To solve for the PI, set up a trial function then substitute into the DE to solve for any unknown constants.

$f(x)$	Trial function
constant	α
x^n $(n\in\mathbb{Z}^+)$	$\alpha x^n+\beta x^{n+1}+\dots$
e^{kx}	αe^{kx}
$x^n e^{kx}$	$(\alpha x^n+\beta x^{n+1}+\dots)e^{kx}$
$\sin px$ / $\cos px$	$\alpha\sin px+\beta\cos px$
$e^{kx}\sin px$ / $e^{kx}\cos px$	$e^{kx}(\alpha\sin px+\beta\cos px)$

- Sum/product of diff. functions → sum/product of corresponding trial functions
- If the trial function appears in the CF, multiply the trial function by x (repeat if req.)

Systems of ODEs.

- To solve a pair of first-order linear ODEs, eliminate 1 of the variables to obtain a second order linear ODEs w/ just 1 variable.
- Systems of DEs in the following form can be solved:

$$\begin{cases}\frac{dx}{dt}+ax+by=f(t)\\ \frac{dy}{dt}+cx+dy=g(t)\end{cases}$$ where a,b,c,d are constants.

- In general, rearrange to express y in terms of $x,\frac{dx}{dt},t$.
  Differentiate to obtain an expression for $\frac{dy}{dt}$ in terms of $x,\frac{dx}{dt},\frac{d^2x}{dt^2},t$.
  Substitute into the other eqn. to obtain a second-order ODE in x.
  After solving for the GS for x, use the expression for y to find the GS for y.

Applications of ODEs.

Linear system concept.

- The non-homogeneous linear ODE.

$$a\ddot{x}+b\dot{x}+cx=f(t)$$

may be thought of as representing a linear system whose input is $f(t)$ and output is $x(t)$.

INPUT $f(t)$ → LINEAR SYSTEM → OUTPUT $x(t)$.

- An important feature of linear systems is that the inputs can be superposed.

input $f_1(t)$ produces output $x_1(t)$
input $f_2(t)$ produces output $x_2(t)$] input $\alpha f_1(t)+\beta f_2(t)$ produces output $\alpha x_1(t)+\beta x_2(t)$.

Modelling with differential elements

- We can model physical problems that have a spatial coordinate as the independent variable using ODEs by applying elemental analysis.
- To model using differential elements,
  - ↳ Draw a diagram of an element.
  - ↳ Resolve forces / moments or balance energy.
  - ↳ Take the limit of $\delta x \to 0$ to obtain an ODE.
  - ↳ Solve the ODE.

* $F(x+\delta x) = F(x) + \delta F$

where F is a physical quantity that depends on the spatial independent variable x.

Difference equations

Difference equations

Definitions and notation.

- The general form of a difference eqn is:

$$f(n, u_n, u_{n+1}, u_{n+2}, \ldots) = 0.$$

- whenever any of the u_n terms (u_n, u_{n+1}, u_{n+2} etc) are multiplied together, divided or functioned (e.g. $\sin u_n$), the eqn. is non-linear.

Linear difference equations with constant coefficients

Linear with constant coefficients.

- we can solve difference eqns in the form

$a u_{n+1} + b u_n = f(n)$ where a, b are constants [first-order]

$a u_{n+2} + b u_{n+1} + c u_n = f(n)$ where a, b, c are constants [second-order]

(Also works for higher order difference eqns).

- The general soln (GS) is made up of 2 parts — the complementary function (CF) and the particular function [GS = CF + PF].
  - ↳ substituting CF into the LHS gives 0 (CF is the GS when homogeneous $f(n)=0$)
  - ↳ substituting PF into the LHS gives $f(n)$.

Finding the complementary function.

- Initially, consider the homogeneous case: $a u_{n+1} + b u_n = 0$

or $a u_{n+2} + b u_{n+1} + c u_n = 0$.

using the trial soln $u_n = A \cdot \lambda^n$. → $u_{n+1} = A \cdot \lambda^{n+1}$, $u_{n+2} = A \cdot \lambda^{n+2}$

so $a(A \cdot \lambda^{n+1}) + b(A \cdot \lambda^n) = 0$ or $a(A \cdot \lambda^{n+2}) + b(A \cdot \lambda^{n+1}) + c(A \lambda^n) = 0$.

$A \lambda^n (a\lambda + b) = 0$. $A \lambda^n (a\lambda^2 + b\lambda + c) = 0$.

As $A \cdot \lambda^n \neq 0$, $a\lambda + b$ or $a\lambda^2 + b\lambda + c = 0$ (Auxiliary equation AE)

For 1st order difference eqn: $\boxed{u_n = A \cdot \lambda^n}$ similar logic to ODE.

For 2nd order difference eqn → superpose 2 soln: $u_n = A \cdot \lambda_1^n + B \cdot \lambda_2^n$

* There are 2 cases for the auxiliary eqn (second-order).
  - ↳ 2 distinct roots λ_1, λ_2 → $u_n = A \cdot \lambda_1^n + B \cdot \lambda_2^n$
  - ↳ 2 repeated real roots λ → $u_n = A \cdot \lambda^n + nB \cdot \lambda^n$

Finding the particular function

- Now, consider the non-homogeneous case; $au_{n+1} + bu_n = f(n)$

or $au_{n+2} + bu_{n+1} + cu_n = f(n)$.

- To solve for the PF, set up a trial function then substitute into the difference eqn. to solve for any unknown constants.

f(n)	Trial function
constant	α
n^k $(k \in \mathbb{Z}^+)$.	$\alpha n^k + \beta n^{k-1} + \dots$
k^n	αk^n

- sum of different functions → sum of corresponding trial functions.
- If the trial function appears in the CF, multiply the trial function by n (repeat if req.)

other difference equations

First-order difference equation in the form $u_{n+1} = f(n)u_n$.

- To solve a difference eqn. in the form $u_{n+1} = f(n)u_n$.

$$u_{n+1} = f(n)u_n$$
$$\frac{u_{n+1}}{u_n} = f(n)$$
$$\frac{u_n}{u_{n-1}} \cdot \frac{u_{n-1}}{u_{n-2}} \dots \frac{u_3}{u_2} \cdot \frac{u_2}{u_1} = \prod_{k=1}^{n-1} f(n)$$
$$u_n = \prod_{k=1}^{n-1} f(n)\, u_1$$

-e.g: Find u_n, given $u_{n+1} = 2(n+1)5^n u_n$, $u_1 = 3$.

$$u_{n+1} = 2(n+1)5^n u_n$$
$$\frac{u_{n+1}}{u_n} = 2(n+1)5^n$$
$$\frac{u_n}{u_1} = \prod_{k=1}^{n-1} 2(n+1)5^n$$
$$\frac{1}{3}u_n = \prod_{k=1}^{n-1} 2 \cdot \prod_{k=1}^{n-1} (n+1) \cdot \prod_{k=1}^{n-1} 5^n$$
$$u_n = 3 \cdot 2^{n-1} \cdot n! \cdot 5^{\frac{n(n-1)}{2}}$$

Vectors

Fundamentals, lines and planes

Definitions and notation.

- Scalar quantities have size but NO direction; vector quantities have size AND direction.

  (Magnitude is non-negative but size can be either +ve or −ve).

- Vectors are often written in terms of components:

$$\underline{F} = F_x \underline{i} + F_y \underline{j} + F_z \underline{k}$$

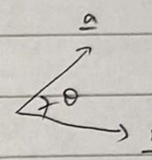
  and the components can be written as (F_x, F_y, F_z) or $\begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix}$.

- The magnitude of a vector is written $|\underline{F}|$ and is given by

$$|\underline{F}| = \sqrt{F_x^2 + F_y^2 + F_z^2}.$$

- A unit vector $\hat{\underline{F}}$ has a magnitude of 1, so $\hat{\underline{F}} = \frac{\underline{F}}{|\underline{F}|}$.

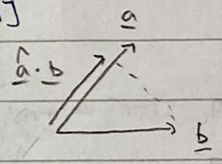
Scalar product (dot product).



- For any 2 vectors $\underline{a}$ and $\underline{b}$, we define their scalar product by:

$$\boxed{\underline{a} \cdot \underline{b} = |\underline{a}||\underline{b}|\cos\theta} \rightarrow \text{gives scalar.}$$

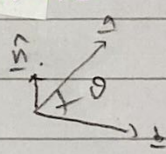
- Properties of the scalar product:
  - ↳ $\underline{a} \cdot \underline{b} = \sum_{\forall i} a_i b_i$ (In 3D, $\underline{a} \cdot \underline{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$). $\rightarrow$ $\underline{a} \cdot \underline{a} = |\underline{a}|^2$.
  - ↳ $\underline{a} \cdot (k\underline{b}) = (k\underline{a}) \cdot \underline{b} = k(\underline{a} \cdot \underline{b})$. [scalar multiplication]
  - ↳ $\underline{a} \cdot \underline{b} = \underline{b} \cdot \underline{a}$ [commutative]
  - ↳ $\underline{a} \cdot (\underline{b} + \underline{c}) = \underline{a} \cdot \underline{b} + \underline{a} \cdot \underline{c}$. [distributive across addition]
  - ↳ $\underline{a} \cdot \underline{b} = 0 \iff \underline{a} \perp \underline{b}$, for non-zero vectors $\underline{a}, \underline{b}$.



- Geometrically, $\hat{\underline{a}} \cdot \underline{b}$ gives the projection of $\underline{b}$ onto $\underline{a}$.

both have scalar multiplication, are distributive but neither is associative

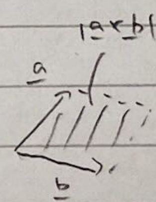
Vector product (cross product).



- For any 2 vectors $\underline{a}$ and $\underline{b}$, we define their vector product by:

$$\boxed{\underline{a} \times \underline{b} = |\underline{a}||\underline{b}|\sin\theta\,\hat{\underline{n}}} \rightarrow \text{gives vector}$$

- Properties of the vector product:
  - ↳ $\underline{a} \times \underline{b} = \begin{vmatrix} \underline{i} & a_1 & b_1 \\ \underline{j} & a_2 & b_2 \\ \underline{k} & a_3 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2)\underline{i} - (a_1 b_3 - a_3 b_1)\underline{j} + (a_1 b_2 - a_2 b_1)\underline{k}$.
  - ↳ $\underline{a} \times (k\underline{b}) = (k\underline{a}) \times \underline{b} = k(\underline{a} \times \underline{b})$ [scalar multiplication]
  - ↳ $\underline{a} \times \underline{b} = -\underline{b} \times \underline{a}$ [anti-commutative]
  - ↳ $\underline{a} \times (\underline{b} + \underline{c}) = \underline{a} \times \underline{b} + \underline{a} \times \underline{c}$ [distributive across addition]
  - ↳ $\underline{a} \times \underline{b} = 0 \iff \underline{a} \parallel \underline{b}$, for non-zero vectors $\underline{a}, \underline{b}$. $\rightarrow \underline{a} \times \underline{a} = 0$.



- Geometrically, $|\underline{a} \times \underline{b}|$ is the area of the //gram formed by $\underline{a}$ and $\underline{b}$

Equation of line

- Vector equation : $\boxed{\underline{r} = \underline{a} + \lambda \underline{d}}$
  - ↳ $\underline{a}$: position vector of pt. A ↳ $\underline{d}$: direction vector of the line.
- Cartesian equation : $\boxed{\frac{x-a_x}{d_x} = \frac{y-a_y}{d_y} = \frac{z-a_z}{d_z}}$
  - ↳ a_i : ith component of position vector ↳ d_i : ith component of direction vector.

* To convert from vector form to cartesian form, make λ the subject.

If $d_i = 0$, then $i = 0$. , where $i = x/y/z$.

Equation of plane.

- Scalar equation : $\boxed{(\underline{r}-\underline{a}) \cdot \underline{n} = 0 \quad \text{or} \quad \underline{r} \cdot \hat{\underline{n}} = \underline{a} \cdot \hat{\underline{n}} = p}$
  - ↳ $\underline{a}$: position vector of pt. A ↳ $\underline{n}$: normal vector of the plane.
  - ↳ p : minimum distance between the plane and the origin.
- Vector equation : $\boxed{\underline{r} = \underline{a} + \lambda \underline{u} + \mu \underline{v}}$
  - ↳ $\underline{a}$: position vector of pt. A ↳ $\underline{u}, \underline{v}$: basis vectors of the plane.
- Cartesian equation : $\boxed{n_x x + n_y y + n_z z = d}$ $(d = \underline{a} \cdot \underline{n})$.
  - ↳ n_i : ith component of normal vector.

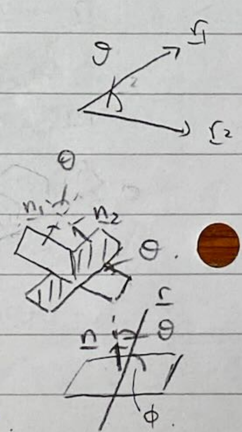
* By defn, $\hat{\underline{n}} = \frac{\underline{u} \times \underline{v}}{|\underline{u} \times \underline{v}|}$. To convert from vector form to other forms, dot both sides with $\underline{n}$ as $\underline{u} \cdot \underline{n} = \underline{v} \cdot \underline{n} = 0$.

Finding the equation of plane.

- Given 3 pts that lie on the plane A, B, C, where $\vec{OA} = \underline{a}$, $\vec{OB} = \underline{b}$, $\vec{OC} = \underline{c}$.
  - ↳ $\underline{n} = \vec{AB} \times \vec{BC}$
  $= (\underline{b} - \underline{a}) \times (\underline{c} - \underline{b})$

  Substitute a_i into the cartesian eqn. to find d.
- Given 1 pt. on the plane A and 2 basis vectors $\underline{u}, \underline{v}$, where $\vec{OA} = \underline{a}$.
  - ↳ $\underline{r} = \underline{a} + \lambda \underline{u} + \mu \underline{v}$.

Finding the angle between lines and planes.



- Angle between 2 lines $\underline{r}_1, \underline{r}_2$.
  - ↳ $\underline{d}_1 \cdot \underline{d}_2 = |\underline{d}_1||\underline{d}_2|\cos\theta \rightarrow \theta = \arccos\left(\frac{\underline{d}_1 \cdot \underline{d}_2}{|\underline{d}_1||\underline{d}_2|}\right)$.
- Angle between 2 planes Π_1, Π_2.
  - ↳ $\underline{n}_1 \cdot \underline{n}_2 = |\underline{n}_1||\underline{n}_2|\cos\theta \rightarrow \theta = \arccos\left(\frac{\underline{n}_1 \cdot \underline{n}_2}{|\underline{n}_1||\underline{n}_2|}\right)$
- Angle between a line $\underline{r}$ and a plane Π
  - ↳ $\underline{d} \cdot \underline{n} = |\underline{d}||\underline{n}|\cos\theta \rightarrow \theta = \arccos\left(\frac{\underline{d} \cdot \underline{n}}{|\underline{d}||\underline{n}|}\right)$, $\phi = 90° - \theta$.

Finding the intersection between lines and planes

- Intersection between 2 lines $\underline{r}_1, \underline{r}_2$, given $\underline{d}_1$ not // $\underline{d}_2$ (lines not //)
  - ↳ set the x, y components to each other and solve for λ, μ.

    Check that $z_1 = z_2$ using the newly solved values of λ, μ.

    * If $z_1 \neq z_2$, the lines $\underline{r}_1, \underline{r}_2$ do not intersect.
- Intersection between 2 planes Π_1, Π_2, given $\underline{n}_1$ not // $\underline{n}_2$ (planes not //)
  - ↳ Let $\underline{r} = \underline{a} + \lambda \underline{d}$ be the intersection line.

$$\underline{d} = \underline{n}_1 \times \underline{n}_2 .$$

Now set the x-coordinates of the 2 planes Π_1, Π_2 to be 0 and solve for the corresponding y/z-coordinates, y_0, z_0

If there is a unique sol'n for y_0, z_0, the line crosses $x=0$.

$$\underline{r} = \begin{pmatrix} 0 \\ y_0 \\ z_0 \end{pmatrix} + \lambda (\underline{n}_1 \times \underline{n}_2).$$

If there is no unique sol'n for y_0, z_0, the line does not cross $x=0$, i.e. $d_x = 0$.

Retry by setting $y=0$ or $z=0$.

- Intersection between a line $\underline{r}$ and a plane Π.
  - ↳ From $\underline{r} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} + \lambda \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix}$, $i = a_i + \lambda d_i$, where $i = x/y/z$.

    Substitute each component into the cartesian eqn. of plane and solve for λ.

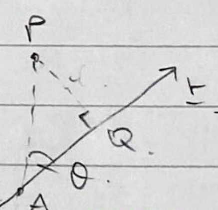
$$\sum_{\forall i} n_i (a_i + \lambda d_i) = d.$$

Substitute the newly solved value of λ into the vector eqn. of line.

* If λ cannot be solved, the line $\underline{r}$ does not intersect the plane Π.

Shortest distance between points, lines and planes.

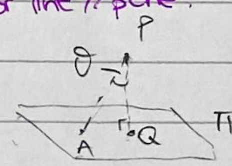
same for // lines



- Shortest distance between a point P and a line $\underline{r}$. ($\vec{OP} = \underline{p}$).
  - ↳ Shortest distance PQ is given by

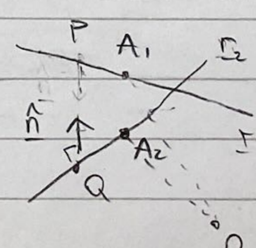
$$PQ = AP\sin\theta = \frac{AP|\underline{d}|\sin\theta}{|\underline{d}|} = \left|\frac{\vec{AP} \times \underline{d}}{|\underline{d}|}\right| = \left|(\underline{p}-\underline{a}) \times \hat{\underline{d}}\right|$$

same for line // plane.



- Shortest distance between a point P and a plane Π ($\vec{OP} = \underline{p}$).
  - ↳ Shortest distance PQ is given by

$$PQ = AP\cos\theta = \frac{AP|\underline{n}|\cos\theta}{|\underline{n}|} = \left|\frac{\vec{AP} \cdot \underline{n}}{|\underline{n}|}\right| = \left|(\underline{p}-\underline{a}) \cdot \hat{\underline{n}}\right|$$



- Shortest distance between 2 (skew) lines $\underline{r}_1, \underline{r}_2$
  - ↳ $\underline{n}$ is along the common normal. $\hat{\underline{n}} = \frac{\underline{d}_1 \times \underline{d}_2}{|\underline{d}_1||\underline{d}_2|}$

$$\vec{OP} = \underbrace{\underline{a}_1}_{\vec{OA_1}} + \underbrace{\lambda_0 \underline{d}_1}_{\vec{A_1P}} = \underbrace{\underline{a}_2}_{\vec{OA_2}} + \underbrace{\mu_0 \underline{d}_2}_{\vec{A_2Q}} + \underbrace{k\hat{\underline{n}}}_{\vec{QP}} \quad \text{(for some } \lambda_0, \mu_0\text{)}.$$

By defn, $\hat{\underline{n}} \perp \underline{d}_1$ and $\underline{d}_2$, so dotting both sides w/ $\hat{\underline{n}}$

$$\underline{a}_1 \cdot \hat{\underline{n}} + \lambda_0 \overset{0}{\cancel{\underline{d}_1 \cdot \hat{\underline{n}}}} = \underline{a}_2 \cdot \hat{\underline{n}} + \mu_0 \overset{0}{\cancel{\underline{d}_2 \cdot \hat{\underline{n}}}} + k \overset{1}{\cancel{\hat{\underline{n}} \cdot \hat{\underline{n}}}}$$

$$\boxed{k = |(\underline{a}_2 - \underline{a}_1) \cdot \hat{\underline{n}}| = \left|(\underline{a}_2 - \underline{a}_1) \cdot \frac{\underline{d}_1 \times \underline{d}_2}{|\underline{d}_1||\underline{d}_2|}\right|}$$

Triple product

Scalar triple product

- The scalar triple product of $\underline{a}, \underline{b}, \underline{c}$ is given by

there is only 1 order of evaluation which makes sense

$$\underline{a}\cdot(\underline{b}\times\underline{c}), \quad (\text{or simply } \underline{a}\cdot\underline{b}\times\underline{c}). \rightarrow \text{gives scalar}$$

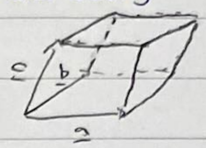
- Properties of scalar triple product:

$$\boxed{\begin{aligned} &\hookrightarrow \underline{a}\cdot\underline{b}\times\underline{c} = \underline{b}\cdot\underline{c}\times\underline{a} = \underline{c}\cdot\underline{a}\times\underline{b} \\ &\hookrightarrow \underline{a}\cdot\underline{b}\times\underline{c} = \underline{a}\times\underline{b}\cdot\underline{c} \\ &\hookrightarrow \underline{a}\cdot\underline{b}\times\underline{c} = -\underline{a}\cdot\underline{c}\times\underline{b} \\ &\hookrightarrow \underline{a}\cdot\underline{a}\times\underline{c} = \underline{a}\cdot\underline{c}\times\underline{a} = 0. \end{aligned}}$$

[cyclic order]

[interchanging order of dot and cross]

[anticyclic order]



- Geometrically, $\underline{a}\cdot\underline{b}\times\underline{c}$ gives volume of the parallelepiped w/ edges formed by $\underline{a}, \underline{b}, \underline{c}$.

Vector triple product

- The vector triple product of $\underline{a}, \underline{b}, \underline{c}$ is given by

never write $\underline{a}\times\underline{b}\times\underline{c}$ as it is ambiguous.

$$\boxed{\begin{aligned} \text{DB} \quad & \underline{a}\times(\overset{\text{mid}}{\underline{b}}\times\overset{\text{other}}{\underline{c}}) = \overset{\text{mid}}{\underline{b}}(\underline{a}\cdot\underline{c}) - \overset{\text{other}}{\underline{c}}(\underline{a}\cdot\underline{b}) \\ \text{or} \quad \text{DB} \quad & (\overset{\text{other}}{\underline{a}}\times\overset{\text{mid}}{\underline{b}})\times\underline{c} = \overset{\text{mid}}{\underline{b}}(\underline{a}\cdot\underline{c}) - \overset{\text{other}}{\underline{a}}(\underline{a}\cdot\underline{c}) \end{aligned}}$$

] gives vector

- e.g: Prove that $(\underline{a}\times\underline{b})\cdot(\underline{c}\times\underline{d}) = \underline{a}\cdot\underline{c}\,\underline{b}\cdot\underline{d} - \underline{a}\cdot\underline{d}\,\underline{b}\cdot\underline{c}$

For simplicity, let $\underline{e} = \underline{c}\times\underline{d}$.

$$\begin{aligned} \underline{a}\times\underline{b}\cdot\underline{e} &= \underline{a}\cdot\underline{b}\times\underline{e} && \text{swapping } \cdot \text{ and } \times \\ &= \underline{a}\cdot\underline{b}\times(\underline{c}\times\underline{d}) && \text{so we can expand vector triple prod.} \\ &= \underline{a}\cdot(\underline{c}(\underline{b}\cdot\underline{d}) - \underline{d}(\underline{b}\cdot\underline{c})) \\ &= \underline{a}\cdot\underline{c}\,\underline{b}\cdot\underline{d} - \underline{a}\cdot\underline{d}\,\underline{b}\cdot\underline{c} && \text{as req.} \end{aligned}$$

strategy: group as $\underline{e}$ → expand (maybe after $\cdot$ $\times$ swap).

- eg: Simplify $(\underline{a}\times\underline{b})\cdot(\underline{a}\times\underline{c})\times(\underline{a}\times\underline{d})$

For simplicity, let $\underline{e} = \underline{a}\times\underline{c}$.

$$\begin{aligned} (\underline{a}\times\underline{b})\cdot\underline{e}\times(\underline{a}\times\underline{d}) &= (\underline{a}\times\underline{b})\cdot(\underline{a}(\underline{e}\cdot\underline{d}) - \underline{d}(\underline{a}\cdot\underline{e})) \\ &= (\underline{a}\times\underline{b})\cdot(\underline{a}(\underline{a}\times\underline{c}\cdot\underline{d}) - \underline{d}(\cancelto{0}{\underline{a}\cdot\underline{a}\times\underline{c}})) \\ &= (\underline{a}\times\underline{c}\cdot\underline{d})(\cancelto{0}{\underline{a}\times\underline{b}\cdot\underline{a}}) = 0. \end{aligned}$$

Geometrically, $(\underline{a}\times\underline{b}), (\underline{a}\times\underline{c}), (\underline{a}\times\underline{d})$ all contain $\underline{a}$

→ All 3 vectors are ⊥ to $\underline{a}$ so they all lie on a plane w/ $\underline{n} = k\underline{a}$.

As the 3 vectors lie on a plane, the volume of the parallelepiped is 0.

$\therefore (\underline{a}\times\underline{b})\cdot(\underline{a}\times\underline{c})\times(\underline{a}\times\underline{d}) = 0$.

Basis.

- The basis of a vector space is a set of linearly independent vectors (basis vectors) that span the full space.
- n vectors $v_1, v_2 \ldots v_n$ are said to be linearly independent if each vector v_i cannot be represented by a linear combination of the other $n-1$ vectors.

$$\underline{v_i} \neq a_1\underline{v_1} + a_2\underline{v_2} + \ldots + a_n\underline{v_n} \quad \text{for any real } a_i$$

- The span of n vectors $v_1, v_2 \ldots v_n$ is the set of all their linear combinations.

$$a_1\underline{v_1} + a_2\underline{v_2} + \ldots + a_n\underline{v_n} \quad \text{for all real } a_i$$

- For an n-dimensional space, there exists a unique linear combination of n basis vectors for each pt. in space.

$$\boxed{\underline{x} = a_1\underline{v_1} + a_2\underline{v_2} + \ldots + a_n\underline{v_n}}$$

- e.g: Find a sol'n for $\underline{x}$ of the equation

$$\underline{x} + \underline{x} \times \underline{b} = \underline{c}$$

given $\underline{b}$ not $//$ $\underline{c}$.

As $\underline{b}$ is not $//$ $\underline{c}$, $\underline{b}, \underline{c}, \underline{b}\times\underline{c}$ are linearly independent $\rightarrow$ set of 3 basis vectors.

$\therefore$ Any general vector $\underline{x}$ in 3D can be expressed as a linear combination of $\underline{b}, \underline{c}, \underline{b}\times\underline{c}$.

$$\boxed{\underline{x} = \alpha\underline{b} + \beta\underline{c} + \gamma\,\underline{b}\times\underline{c}}$$

$$\underline{x}\times\underline{b} = \alpha\,\cancelto{0}{\underline{b}\times\underline{b}} + \beta\,\underline{c}\times\underline{b} + \gamma(\underline{b}\times\underline{c})\times\underline{b}$$

$$= \beta\,\underline{c}\times\underline{b} + \gamma(\underline{c}(\underline{b}\cdot\underline{b}) - \underline{b}(\underline{b}\cdot\underline{c}))$$

$$= -(\underline{b}\cdot\underline{c})\gamma\,\underline{b} + (\underline{b}\cdot\underline{b})\gamma\,\underline{c} - \beta\,\underline{b}\times\underline{c}.$$

substituting into the vector eqn:

$$(\alpha\underline{b} + \beta\underline{c} + \gamma\,\underline{b}\times\underline{c}) + (-(\underline{b}\cdot\underline{c})\gamma\,\underline{b} + (\underline{b}\cdot\underline{b})\gamma\,\underline{c} - \beta\,\underline{b}\times\underline{c}) = \underline{c}.$$

since $\underline{b}, \underline{c}, \underline{b}\times\underline{c}$ are linearly independent $\rightarrow$ we can compare coefficients.

$$\left.\begin{aligned} \underline{b}:&\quad \alpha - (\underline{b}\cdot\underline{c})\gamma = 0 \\ \underline{c}:&\quad \beta + (\underline{b}\cdot\underline{b})\gamma = 1 \\ \underline{b}\times\underline{c}:&\quad \gamma - \beta = 0 \end{aligned}\right] \qquad \begin{aligned} \alpha &= \frac{\underline{b}\cdot\underline{c}}{1+|\underline{b}|^2} \\ \beta &= \frac{1}{1+|\underline{b}|^2} \\ \gamma &= \frac{1}{1+|\underline{b}|^2} \end{aligned}$$

$$\rightarrow \underline{x} = \alpha\underline{b} + \beta\underline{c} + \gamma\,\underline{b}\times\underline{c}$$

$$= \frac{\underline{b}\cdot\underline{c}}{1+|\underline{b}|^2}\underline{b} + \frac{1}{1+|\underline{b}|^2}\underline{c} + \frac{1}{1+|\underline{b}|^2}\underline{b}\times\underline{c}$$

$$= \frac{(\underline{b}\cdot\underline{c})\underline{b} + \underline{c} + \underline{b}\times\underline{c}}{1+|\underline{b}|^2}.$$

Matrices, transformations, eigenvalues and eigenvectors

Matrix algebra

Notation.

- A typical vector is written as

$$\underline{a} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \text{ or } \underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \text{ and position vectors } \underline{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ or } \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

- The most commonly used basis $(\underline{i}, \underline{j}, \underline{k})$ is implied.
- A typical 3x3 matrix is written as

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ or } \underline{\underline{A}} = \begin{bmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{bmatrix}$$

where a_{ij} is the element in the ith row and jth column.

Matrix multiplication.

- Matrix multiplication is implied by juxtaposition.
- The ij element a_{ij} of the product of the matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$ is:

$$(\underline{\underline{A}}\underline{\underline{B}})_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

or equivalently, as scalar products,

$$\boxed{(\underline{\underline{A}}\underline{\underline{B}})_{ij} = \underline{a}_i \cdot \underline{b}_j}$$

where $\underline{a}_i$ and $\underline{b}_j$ are the ith row vector of $\underline{\underline{A}}$ and the jth column vector of $\underline{\underline{B}}$ respectively.

- A $m \times n$ matrix multiplied by a $p \times q$ matrix is only defined when $n = p$ and the product is a matrix of order $m \times q$.
- Vectors are simply treated as non-square matrices.
- Properties of matrix multiplication
  - ↳ $\underline{\underline{A}}\underline{\underline{B}} \neq \underline{\underline{B}}\underline{\underline{A}}$ [not commutative]
  - ↳ $\underline{\underline{A}}(\underline{\underline{B}}+\underline{\underline{C}}) = \underline{\underline{A}}\underline{\underline{B}} + \underline{\underline{A}}\underline{\underline{C}}$ [distributive over addition]
  - ↳ $(\underline{\underline{A}}\underline{\underline{B}})\underline{\underline{C}} = \underline{\underline{A}}(\underline{\underline{B}}\underline{\underline{C}}) = \underline{\underline{A}}\underline{\underline{B}}\underline{\underline{C}}$ [associative]
- Scalar multiplication / scalar product of vectors can be expressed as matrix multiplication
  - ↳ $\boxed{\lambda \underline{a} = (\lambda \underline{a}^T)^T = \underline{a}\underline{\lambda}}$, where λ is a 1x1 matrix.
  - ↳ $\boxed{\underline{a} \cdot \underline{b} = \underline{a}^T \underline{b}}$ $[\lambda]\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ is not defined, but $[\lambda][a_1, a_2 \dots a_n]$ is defined.

Identity matrix

- The identity matrix $\underline{\underline{I}}$ is defined s.t. for a square matrix $\underline{\underline{A}}$

$$\boxed{\underline{\underline{I}}\underline{\underline{A}} = \underline{\underline{A}}\underline{\underline{I}} = \underline{\underline{A}}}$$

- For 3x3 matrices, the identity matrix is

$$\underline{\underline{I}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Transpose.

- The transpose $\underline{A}^T$ is formed by interchanging the rows and columns of the matrix.

$$\underline{A} = \begin{bmatrix} \leftarrow a_1 \rightarrow \\ \leftarrow a_2 \rightarrow \\ \leftarrow a_3 \rightarrow \end{bmatrix} \qquad \underline{A}^T = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ a_1 & a_2 & a_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}.$$

or equivalently, in index notation

$$\boxed{(\underline{A}^T)_{ij} = a_{ji}}$$

- Properties of the transpose
  - ↳ $(\underline{A}^T)^T = \underline{A}$.
  - ↳ $\boxed{(\underline{A}\underline{B}\underline{C}\ldots\underline{F})^T = \underline{F}^T\ldots\underline{C}^T\underline{B}^T\underline{A}^T}$
- The inner/outer products of 2 column vectors can be expressed using the transpose.
  - ↳ Inner product : $\underset{\text{inside}}{\underline{a}^T\underline{b}} = [a_1\, a_2 \ldots a_n]\begin{bmatrix} b_1 \\ b_2 \\ b_n \end{bmatrix} = \sum_{i=1}^{n} a_i b_i = \underline{a}\cdot\underline{b}$
  - ↳ Outer product : $\underset{\text{outside}}{\underline{a}\,\underline{b}^T} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}[b_1, b_2 \ldots b_n] = \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & & \ddots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{bmatrix}$ i.e. $(\underline{a}\underline{b}^T)_{ij} = a_ib_j$

Determinants

- For a 2x2 matrix $\underline{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant $\det(\underline{A})$ or $|\underline{A}|$ is given by

$$\boxed{\det(\underline{A}) = ad - bc}$$

- For a 3x3 matrix $\underline{B} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, the determinant $\det(\underline{B})$ or $|\underline{B}|$ is given by.

$$\det(\underline{B}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

↳ compute using Sarrus' rule or expand wrt a row/column.

$(-1)^{i+j}$: $\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$

① Sarrus' rule

- Consider the matrix $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

$$\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

The determinant is given by

$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$

$- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$

② Expand wrt row/column.

- Sum each element multiplied by its cofactor across a row/column.
- The cofactor of the ijth element is given by:

$$C_{ij} = (-1)^{i+j} M_{ij}.$$

where M_{ij} is the determinant of the submatrix after deleting the ith row + jth column (minor matrix)

- When a matrix is used to represent a transformation, the determinant is the scale factor for the transformation. (2x2: ratio of areas, 3x3: ratio of volumes).
- Properties of determinants.
  - ↳ $\det(\underline{A}^T) = \det(\underline{A})$
  - ↳ $\det(\underline{A}\underline{B}) = \det(\underline{A}) \times \det(\underline{B})$.
  - ↳ $\begin{vmatrix} \leftarrow \alpha a_1 \rightarrow \\ \leftarrow a_2 \rightarrow \\ \leftarrow a_3 \rightarrow \end{vmatrix} = \alpha \begin{vmatrix} \leftarrow a_1 \rightarrow \\ \leftarrow a_2 \rightarrow \\ \leftarrow a_3 \rightarrow \end{vmatrix}$
  - ↳ $\begin{vmatrix} \leftarrow a_1 \rightarrow \\ \leftarrow a_2 \rightarrow \\ \leftarrow a_3 \rightarrow \end{vmatrix} = -\begin{vmatrix} \leftarrow a_2 \rightarrow \\ \leftarrow a_1 \rightarrow \\ \leftarrow a_3 \rightarrow \end{vmatrix} \rightarrow \begin{vmatrix} \leftarrow a_1 \rightarrow \\ \leftarrow a_1 \rightarrow \\ \leftarrow a_3 \rightarrow \end{vmatrix} = 0.$
  - ↳ $\begin{vmatrix} \leftarrow a_1 \rightarrow \\ \leftarrow a_2 \rightarrow \\ \leftarrow a_3 \rightarrow \end{vmatrix} = \begin{vmatrix} \leftarrow a_1 + m a_2 \rightarrow \\ \leftarrow a_2 \rightarrow \\ \leftarrow a_3 \rightarrow \end{vmatrix}$

Inverse matrix

- For a square matrix $\underline{\underline{A}}$, the inverse $\underline{\underline{A}}^{-1}$ is defined s.t.

$$\boxed{\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{I}}}$$

- The inverse of a matrix $\underline{\underline{A}}^{-1}$ exists iff $\det(\underline{\underline{A}}) \neq 0$. → $\underline{\underline{A}}$ is non-singular.
- The inverse of a matrix $\underline{\underline{A}}^{-1}$ is given by.

$$\boxed{\underline{\underline{A}}^{-1} = \frac{\operatorname{adj}(\underline{\underline{A}})}{\det(\underline{\underline{A}})}}$$

where the adjoint / adjugate matrix $\operatorname{adj}(\underline{\underline{A}})$ is given by the transpose of the cofactor matrix $\operatorname{cof}(\underline{\underline{A}})$.

$$\boxed{\operatorname{adj}(\underline{\underline{A}}) = \operatorname{cof}(\underline{\underline{A}})^T}$$

Intersection of 3 planes.

- There are 4 ways in which 3 distinct planes Π_1, Π_2, Π_3 can intersect in 3D space.
  - ↳ They intersect at a single pt. (unique sol'n)
  - ↳ They form a sheaf (infinitely many sol'n)
  - ↳ They form a triangular prism. (no sol'n)
  - ↳ 2 or more planes are //. (no sol'n).
- The 3 planes can be represented as a system of 3 simultaneous linear eqns.

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases} \xrightarrow{\text{in matrix form}} \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ a & b & c \\ \downarrow & \downarrow & \downarrow \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \uparrow \\ d \\ \downarrow \end{bmatrix}.$$

→ $\left|\begin{matrix} \uparrow & \uparrow & \uparrow \\ a & b & c \\ \downarrow & \downarrow & \downarrow \end{matrix}\right| = 0$ → unique sol'n → intersect at a single pt.

↳ $\left|\begin{matrix} \uparrow & \uparrow & \uparrow \\ a & b & c \\ \downarrow & \downarrow & \downarrow \end{matrix}\right| \neq 0$ → eqn.s are consistent → infinitely many sol'n → sheaf

↳ eqn.s not consistent → no sol'n → 2 or more planes //

↳ triangular prism.

Mappings and transformations

Transformations

- Matrices are used to map one vector to another. If $\underline{x}' = \underline{\underline{A}}\,\underline{x}$, we say $\underline{\underline{A}}$ maps $\underline{x}$ to $\underline{x}'$.
- Let $\underline{a_1}, \underline{a_2}, \underline{a_3}$ be the 3 column vectors of $\underline{\underline{A}}$. i.e. $\underline{\underline{A}} = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ a_1 & a_2 & a_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}$.

  $\underline{\underline{A}}$ maps $\underline{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\underline{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\underline{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ to $\underline{a_1}, \underline{a_2}, \underline{a_3}$ respectively.

  → To determine the matrix representing a transformation, we need to determine where the 3 coordinate vectors $\underline{i}, \underline{j}, \underline{k}$, map to under the transformation.

Orthogonal matrices.

- An orthogonal matrix $\underline{\underline{Q}}$ is made up of a set of orthonormal vectors. (orthonormal set)
- The vectors in an orthonormal set are mutually perpendicular and unit vectors.
- Properties of orthogonal matrices.

$q_i \cdot q_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

↳ $\boxed{\underline{\underline{Q}}^T = \underline{\underline{Q}}^{-1}}$: $\underline{\underline{Q}}\,\underline{\underline{Q}}^T = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ q_1 & q_2 & q_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}\begin{bmatrix} \leftarrow q_1 \rightarrow \\ \leftarrow q_2 \rightarrow \\ \leftarrow q_3 \rightarrow \end{bmatrix} = \begin{bmatrix} q_1\cdot q_1 & q_1\cdot q_2 & q_1\cdot q_3 \\ q_2\cdot q_1 & q_2\cdot q_2 & q_2\cdot q_3 \\ q_3\cdot q_1 & q_3\cdot q_2 & q_3\cdot q_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underline{\underline{I}}$.

↳ Transformation preserves length and angle

$$|\underline{\underline{Q}}\underline{x}|^2 = (\underline{\underline{Q}}\underline{x})^T \underline{\underline{Q}}\underline{x} = \underline{x}^T(\underline{\underline{Q}}^T\underline{\underline{Q}})\underline{x} = \underline{x}^T\underline{\underline{I}}\,\underline{x} = \underline{x}^T\underline{x} = |\underline{x}|^2 \quad \text{[length invariance]}$$

$$(\underline{\underline{Q}}\underline{x})\cdot(\underline{\underline{Q}}\underline{y}) = (\underline{\underline{Q}}\underline{x})^T\underline{\underline{Q}}\underline{y} = \underline{x}^T(\underline{\underline{Q}}^T\underline{\underline{Q}})\underline{y} = \underline{x}^T\underline{\underline{I}}\,\underline{y} = \underline{x}^T\underline{y} = \underline{x}\cdot\underline{y}. \quad \text{[scalar product preserved]}$$

- If $\det(\underline{\underline{Q}}) = 1$, $\underline{\underline{Q}}$ represents a rotation. (Proper orthogonal matrix)

  If $\det(\underline{\underline{Q}}) = -1$, $\underline{\underline{Q}}$ represents a reflection (+rotation).

Rotation matrix.

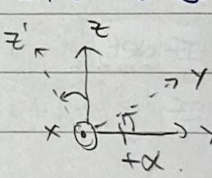
- Rotation matrices are proper orthogonal matrices, i.e. $\underline{\underline{Q}}\,\underline{\underline{Q}}^T = \underline{\underline{I}}$ and $\det(\underline{\underline{Q}}) = 1$.
- In 2D, the matrix $\underline{\underline{Q}}$ that represents a rotation by an angle $+\alpha$ is given by

  ↖ anticlockwise direction

$$\underline{\underline{Q}} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

- In 3D, we can rotate about any arbitrary axis. Consider the special case where we rotate by angle $+\alpha$ about the x-axis. (similar method for y-axis / z-axis).

  When looking from +ve x to the origin, $+\alpha$ represents an anticlockwise rotation of the y-z plane by an angle α.



Change of basis (coordinate system).

Change of basis for a vector

- Consider a new set of basis vectors $\underline{m}_1, \underline{m}_2, \underline{m}_3$. In this new basis, the coordinates (α, β, γ) represents the position vector $\underline{x} = \alpha\underline{m}_1 + \beta\underline{m}_2 + \gamma\underline{m}_3$.

  This can be conveniently written as $\underline{x} = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ m_1 & m_2 & m_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$.

- $\underline{\underline{M}} = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ m_1 & m_2 & m_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}$ is the change of basis vector, where the column vectors $\underline{m}_1, \underline{m}_2, \underline{m}_3$ represent the unit vectors in the new basis, expressed using the coords in the standard basis.
- If we let the position vector that uses the coords in the new basis to be $\underline{x}'$.

$$\boxed{\underline{x} = \underline{\underline{M}}\,\underline{x}'}$$

⚠ The 2 position vectors $\underline{x}, \underline{x}'$ represent the <u>same</u> vector in space. It is only the basis (coordinate system) that is different!

change of basis of a matrix

- let's say we have a matrix $\underline{\underline{A}}$ that maps the vector $\underline{x}$ to $\underline{y}$ in the standard basis. Now we want to find a new matrix $\underline{\underline{A}}'$ that maps the vector $\underline{x}'$ to $\underline{y}'$ in the new basis.

$$\underline{y} = \underline{\underline{A}}\,\underline{x} \qquad \text{and} \qquad \underline{y}' = \underline{\underline{A}}'\underline{x}'$$

Recall $\underline{x} = \underline{\underline{M}}\,\underline{x}'$, where $\underline{\underline{M}}$ is the change of basis matrix. Similarly, $\underline{y} = \underline{\underline{M}}\,\underline{y}'$

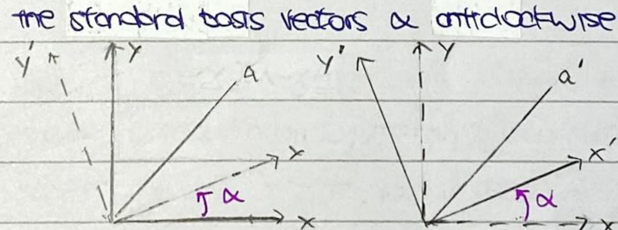
$$\underline{y} = \underline{\underline{A}}\,\underline{x}$$
$$\underline{\underline{M}}\,\underline{y}' = \underline{\underline{A}}\,\underline{\underline{M}}\,\underline{x}'$$
$$\underline{y}' = \underline{\underline{M}}^{-1}\underline{\underline{A}}\,\underline{\underline{M}}\,\underline{x}'$$
$$= \underline{\underline{A}}'\underline{x}' \rightarrow \boxed{\underline{\underline{A}}' = \underline{\underline{M}}^{-1}\underline{\underline{A}}\,\underline{\underline{M}}}$$

* Note $\underline{\underline{A}}$ and $\underline{\underline{A}}'$ represent the <u>same</u> transformation.
$\underline{x}$ and $\underline{x}'$ // $\underline{y}$ and $\underline{y}'$ represent the <u>same</u> vector in space.

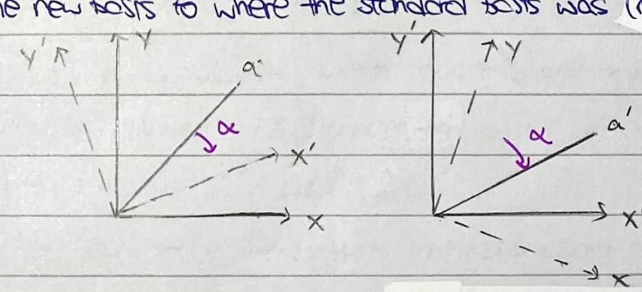
- $\underline{\underline{A}}' = \underline{\underline{M}}^{-1}\underline{\underline{A}}\,\underline{\underline{M}}$ takes in a vector $\underline{x}'$ and returns a new vector $\underline{y}'$, where $\underline{x}'$ and $\underline{y}'$ are in the new basis. Geometrically, we are doing: $\underline{y}' = \underline{\underline{A}}'\underline{x}' = \underline{\underline{M}}^{-1}\underline{\underline{A}}\,\underline{\underline{M}}\,\underline{x}'$
  - ↳ ① Take in the vector $\underline{x}'$ and convert it into the standard basis to give $\underline{x}$. [$\underline{x} = \underline{\underline{M}}\,\underline{x}'$]
  - ↳ ② Map the vector $\underline{x}$ to the vector $\underline{y}$ in the standard basis. [$\underline{y} = \underline{\underline{A}}\,\underline{x}$]
  - ↳ ③ Take the vector $\underline{y}$ and convert it back to the new basis to give $\underline{y}'$ [$\underline{y}' = \underline{\underline{M}}^{-1}\underline{y}$].

Special case of change of basis — rotations

- Rotation matrices are orthogonal matrices so $\underline{\underline{Q}}^{-1} = \underline{\underline{Q}}^T$.
- Rotation matrices are easy to visualise — if $\underline{\underline{Q}}$ represents a rotation by an angle α in a certain direction, $\underline{\underline{Q}}^{-1}/\underline{\underline{Q}}^T$ represents a rotation by the same angle α in the opposite direction (about the original axis of rotation).
- Let's say we rotate the standard basis vectors α anticlockwise to give our new basis,



and now we set the new basis to where the standard basis was (rotate clockwise by α).



we can see this change of basis matrix $\underline{\underline{Q}}$ which rotates the basis (standard → new) α anticlockwise effectively maps the vector $\underline{a}$ to $\underline{a}'$ (by rotating α clockwise), which can be represented by $\boxed{\underline{\underline{R}} = \underline{\underline{Q}}^{-1} = \underline{\underline{Q}}^T}$

- For a vector, $\underline{x} = \underline{\underline{Q}}\,\underline{x}' \rightarrow \boxed{\underline{x}' = \underline{\underline{R}}\,\underline{x}}$; For a matrix, $\underline{\underline{A}}' = \underline{\underline{Q}}^{-1}\underline{\underline{A}}\,\underline{\underline{Q}} \rightarrow \boxed{\underline{\underline{A}}' = \underline{\underline{R}}\,\underline{\underline{A}}\,\underline{\underline{R}}^{-1} = \underline{\underline{R}}\,\underline{\underline{A}}\,\underline{\underline{R}}^T}$

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors

- A non-zero vector $\underline{x}$ that obeys the relation $\boxed{\underline{\underline{A}}\,\underline{x} = \lambda \underline{x}}$ is said to be the eigenvector of $\underline{\underline{A}}$, and the scalar λ is the corresponding eigenvalue.
- If $\underline{\underline{A}}$ represents a transformation, then the eigenvector(s) $\underline{x}$ is the direction which is unaltered before and after $\underline{\underline{A}}$ is applied and the eigenvalue(s) λ is the scale factor applied to its magnitude.
- Although the eigenvector $\underline{x}$ must be nonzero, the eigenvalue λ can be 0. (Matrix $\underline{\underline{A}}$ is singular)
- If $\underline{x}$ is an eigenvector of $\underline{\underline{A}}$, then so is $k\underline{x}$ for any non-zero scalar k → only the dir. matters.
- For a matrix w/ real entries, the eigenvalues and their eigenvectors come in complex conjugate pairs.
- Properties of eigenvalues
  - ↳ $\boxed{\sum_{i=1}^{n} \lambda_i = Tr(\underline{\underline{A}}) = \sum_{i=1}^{n} a_{ii}}$
  - ↳ $\boxed{\prod_{i=1}^{n} \lambda_i = \det(\underline{\underline{A}})}$. Volume of parallelepiped
  - ↳ The eigenvalues of $\underline{\underline{A}}^T$ are λ_i
  - ↳ The eigenvalues of $\underline{\underline{A}}^{-1}$ are $\frac{1}{\lambda_i}$ (provided $\underline{\underline{A}}^{-1}$ exists). Reverse scale factor
  - ↳ The eigenvalues of $k\underline{\underline{A}}$ are $k\lambda_i$
  - ↳ The eigenvalues of $\underline{\underline{A}}^k$ are λ_i^k

Finding the eigenvalues

- We start by finding the eigenvalues. Consider the general $n \times n$ case.

$$\underline{\underline{A}}\,\underline{x} = \lambda \underline{x}$$

$$\underline{\underline{A}}\,\underline{x} = \lambda \underline{\underline{I}}\,\underline{x}$$

$$(\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{x} = 0$$

We can only have a non-trivial soln when the matrix $(\underline{\underline{A}} - \lambda\underline{\underline{I}})$ maps a non-zero vector $\underline{x} = 0$ to the zero vector. This is only true if $\det(\underline{\underline{A}} - \lambda\underline{\underline{I}})$ / scale factor is 0.

$$\boxed{\det(\underline{\underline{A}} - \lambda\underline{\underline{I}}) = 0}$$

Expanding the determinant, we will always get a polynomial of order n in the variable λ. This is the characteristic equation of the matrix.

$$b_n\lambda^n + b_{n-1}\lambda^{n-1} + \dots + b_1\lambda + b_0 = 0.$$

From the fundamental thm. of algebra, there must be n soln's / roots.

$$\lambda_1, \lambda_2, \dots \lambda_n.$$

These are the n eigenvalues of the matrix $\underline{\underline{A}}$.

Finding the eigenvalues of a 2x2 matrix (shortcut).

- For a 2x2 matrix $\underline{A}$, the eigenvalues are given by.

$$\boxed{\lambda = m \pm \sqrt{m^2 - p}}$$

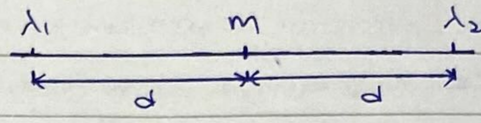
where m is the half of the trace of $\underline{A}$.

and p is the determinant of $\underline{A}$.

- Actually m is the mean of the eigenvalues $m = \frac{\lambda_1 + \lambda_2}{2}$

and p is the product of the eigenvalues. $p = \lambda_1 \lambda_2$.

If we draw the no. line,



$\lambda_1 = m - d$, $\lambda_2 = m + d$. $\rightarrow$ $\lambda_1 \lambda_2 = (m+d)(m-d) = p$

$m^2 - d^2 = p \rightarrow d = \sqrt{m^2 - p}$.

therefore $\lambda = m \pm d = m \pm \sqrt{m^2 - p}$.

Finding the eigenvectors.

- For each eigenvalue λ_i, we need to find a corresponding eigenvector $\underline{x}_i$ s.t.

$$\boxed{(\underline{A} - \lambda_i \underline{\underline{I}})\underline{x}_i = 0}$$

Consider finding the eigenvector $\underline{x}_1$ corresponding to the eigenvalue λ_1

$$\begin{bmatrix} A_{11}-\lambda_1 & A_{12} & \cdots & A_{1(n-1)} & A_{1n} \\ A_{21} & A_{22}-\lambda_1 & \cdots & A_{2(n-1)} & A_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ A_{(n-1)1} & A_{(n-1)2} & \cdots & A_{(n-1)(n-1)}-\lambda_1 & A_{(n-1)n} \\ A_{n1} & A_{n2} & \cdots & A_{n(n-1)} & A_{nn}-\lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \underline{0}$$

By construction, the matrix $[\underline{A} - \lambda \underline{\underline{I}})$ has zero determinant $\rightarrow$ the n eqns are not independent.

$\rightarrow$ 1 of the eqn. is a combination of the other n-1 eqns and so can be neglected.

(usually, it does not matter which eqn. we remove. Here, we ignore the last eqn.).

$$\begin{bmatrix} A_{11}-\lambda_1 & A_{12} & \cdots & A_{1(n-1)} & A_{1n} \\ A_{21} & A_{22}-\lambda_1 & \cdots & A_{2(n-1)} & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{(n-1)1} & A_{(n-1)2} & \cdots & A_{(n-1)(n-1)}-\lambda_1 & A_{(n-1)n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \underline{0}$$

As the magnitude of the eigenvector is arbitrary, we can set any (nonzero) component of the eigenvector to unit size (or any convenient size). Here, we set $x_n = 1$.

$$\begin{bmatrix} A_{11}-\lambda_1 & A_{12} & \cdots & A_{1(n-1)} & A_{1n} \\ A_{21} & A_{22}-\lambda_1 & \cdots & A_{2(n-1)} & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{(n-1)1} & A_{(n-1)2} & \cdots & A_{(n-1)(n-1)}-\lambda_1 & A_{(n-1)n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ 1 \end{bmatrix} = \underline{0}$$

If $x_n = 0$ (i.e. $x_n \neq 1$), then we choose a different $x_i = 1$ and the new RHS would be $\begin{bmatrix} -A_{1i} \\ -A_{2i} \\ \vdots \\ -A_{(n-1)i} \end{bmatrix}$ after rearranging.

Rearranging,

$$\begin{bmatrix} A_{11}-\lambda_1 & A_{12} & \cdots & A_{1(n-1)} \\ A_{21} & A_{22}-\lambda_1 & \cdots & A_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ A_{(n-1)1} & A_{(n-1)2} & \cdots & A_{(n-1)(n-1)}-\lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} -A_{1n} \\ -A_{2n} \\ \vdots \\ -A_{(n-1)n} \end{bmatrix}$$

Now we have a system of n-1 eqns in n-1 unknowns. Hence we can compute the remaining n-1 components of the eigenvector. $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix}$.

↳ if this cannot be solved, the n-1 eqns are not independent → we need to neglect a different eqn. instead.

Therefore the eigenvector $\underline{x}_1$ is $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix}$ (corresponding to the eigenvalue λ_1).

- It may be useful to normalise the eigenvectors: (so they have unit length) $\underline{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \frac{1}{\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Symmetric matrices

- A matrix $\underline{\underline{S}}$ is said to be symmetric if

$$\boxed{\underline{\underline{S}}^T = \underline{\underline{S}}}$$

or equivalently in index notation,

$$\boxed{S_{ij} = S_{ji}}$$

for an anti-symmetric matrix $\underline{\underline{A}}$,

$$\underline{\underline{A}}^T = -\underline{\underline{A}}$$

and $A_{ij} = -A_{ji}$

- Properties of symmetric matrices.

↳ If $\underline{\underline{S}}$ is a real, symmetric matrix, the eigenvalues and eigenvectors of $\underline{\underline{S}}$ are real.

Proof by contradiction — we assume the eigenvalues and eigenvectors are complex.

Let $\lambda = \alpha + i\beta$, $\underline{x} = \underline{u} + i\underline{v}$. (note $\bar{\underline{x}} = \underline{u} - i\underline{v}$).

By defn $\underline{\underline{S}}\underline{x} = \lambda \underline{x}$.

Premultiplying by $\bar{\underline{x}}^T$ $\bar{\underline{x}}^T \underline{\underline{S}}\, \underline{x} = \lambda \bar{\underline{x}}^T \underline{x}$

$$= \lambda(\underline{u}^T - i\underline{v}^T)(\underline{u} + i\underline{v})$$

Matrix multiplication is distributive. $= \lambda(\underline{u}^T\underline{u} - \cancel{i\underline{v}^T\underline{u}} + \cancel{\underline{u}^T i\underline{v}} + \underline{v}^T\underline{v})$.

$$= \lambda(\underline{u}\cdot\underline{u} + \underline{v}\cdot\underline{v}) \quad [1]$$

Taking the conjugate $\overline{\bar{\underline{x}}^T \underline{\underline{S}}\, \underline{x}} = \overline{\lambda(\underline{u}\cdot\underline{u} + \underline{v}\cdot\underline{v})}$

$$\bar{\bar{\underline{x}}}^T \bar{\underline{\underline{S}}}\, \bar{\underline{x}} = \bar{\lambda}(\overline{\underline{u}\cdot\underline{u}} + \overline{\underline{v}\cdot\underline{v}})$$

$\underline{\underline{S}}^T = \underline{\underline{S}}$ and $\underline{u}, \underline{v} \in \mathbb{R}$. $\underline{x}^T \underline{\underline{S}}\, \bar{\underline{x}} = \bar{\lambda}(\underline{u}\cdot\underline{u} + \underline{v}\cdot\underline{v})$.

Taking the transpose $(\underline{x}^T \underline{\underline{S}}\bar{\underline{x}})^T = \bar{\underline{x}}^T \underline{\underline{S}}\, \underline{x} = \bar{\lambda}(\underline{u}\cdot\underline{u} + \underline{v}\cdot\underline{v})$ [2].

Comparing eqns [1], [2], $\lambda = \bar{\lambda} \rightarrow \lambda$ must be real $\rightarrow$ eigenvectors $\underline{x}$ must be real.

↳ If $\underline{\underline{S}}$ is a real, symmetric matrix, then the eigenvectors of $\underline{\underline{S}}$ are orthogonal.

In this proof, we assume the eigenvalues are distinct.

Let λ and μ be 2 distinct eigenvalues of a real symmetric matrix $\underline{\underline{S}}$ and $\underline{u}$ and $\underline{v}$ be the corresponding eigenvectors.

i.e. $\underline{\underline{S}}\underline{u} = \lambda\underline{u}$ $\qquad$ $\underline{\underline{S}}\underline{v} = \mu\underline{v}$

$\underline{v}^T \underline{\underline{S}}\, \underline{u} = \lambda \underline{v}^T \underline{u}$ $\qquad$ $\underline{u}^T \underline{\underline{S}}\, \underline{v} = \mu \underline{u}^T \underline{v}$

$= \lambda \underline{v}\cdot\underline{u}$ [3] $\qquad$ $= \mu \underline{u}\cdot\underline{v}$. [4]

Taking the transpose of [4], $(\underline{u}^T \underline{\underline{S}}\, \underline{v})^T = (\mu\, \underline{u}\cdot\underline{v})^T$

$\underline{\underline{S}}^T = \underline{\underline{S}}$ $\qquad$ $\underline{v}^T \underline{\underline{S}}\, \underline{u} = \mu \underline{u}\cdot\underline{v}$ [5]

Comparing [3], [5]. $\lambda \underline{v}\cdot\underline{u} = \mu \underline{u}\cdot\underline{v}$

$$(\lambda - \mu)\, \underline{u}\cdot\underline{v} = 0.$$

Since we assumed $\lambda \neq \mu$ (distinct eigenvalues), $\lambda - \mu \neq 0$.

Therefore $\underline{u}\cdot\underline{v} = 0 \rightarrow \underline{u}, \underline{v}$ are orthogonal.

(This means $\underline{u}_j^T \underline{\underline{S}}\, \underline{u}_i = \lambda_i \underline{u}_j^T \underline{u}_i = 0$ for $i \neq j \rightarrow$ eigenvectors are orthogonal wrt $\underline{\underline{S}}$).

* The matrix of normalised eigenvectors $\underline{\underline{U}}$ of a symmetric matrix $\underline{\underline{S}}$ is orthogonal.

Defective matrices

- A defective matrix is a square matrix that does not have a complete basis of eigenvectors, and therefore is not diagonalisable.
- In particular, a $n \times n$ matrix is defective iff it does not have n linearly independent eigenvectors.
- If the algebraic multiplicity is equal to the geometric multiplicity, it is possible to form a set of n linearly independent eigenvectors.
- The algebraic multiplicity is the no. of times an eigenvalue is repeated.
  The geometric multiplicity is the no. of linearly independent eigenvectors associated w/ the repeated eigenvalue

Diagonalising matrices

- Let $\underline{A}$ be a 3×3 matrix w/ eigenvalues $\lambda_1, \lambda_2, \lambda_3$ and corresponding normalised eigenvectors $\underline{u}_1, \underline{u}_2, \underline{u}_3$.

  By defn, $\underline{\underline{A}}\underline{u}_1 = \lambda_1 \underline{u}_1$; $\underline{\underline{A}}\underline{u}_2 = \lambda_2 \underline{u}_2$; $\underline{\underline{A}}\underline{u}_3 = \lambda_3 \underline{u}_3$.

  This can be combined into a single matrix: $\underline{\underline{A}}\begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{u}_1 & \underline{u}_2 & \underline{u}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \lambda_1\underline{u}_1 & \lambda_2\underline{u}_2 & \lambda_3\underline{u}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}$

  $= \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{u}_1 & \underline{u}_2 & \underline{u}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$

  Setting $\underline{\underline{U}} = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{u}_1 & \underline{u}_2 & \underline{u}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}$ and $\underline{\underline{\Lambda}} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$, we have $\underline{\underline{A}}\,\underline{\underline{U}} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}$.

  Provided that $\underline{\underline{U}}^{-1}$ exists, $\boxed{\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^{-1}}$

  * For $\underline{\underline{U}}^{-1}$ to exist, $\det(\underline{\underline{U}}) \neq 0$, which is only true iff $\underline{u}_1, \underline{u}_2, \underline{u}_3$ are linearly independent.
  (we don't have linearly independent eigenvectors $\underline{u}_i$ for defective matrices → cannot diagonalise).

- In the special case of a real symmetric matrix $\underline{\underline{S}}$, it has orthogonal eigenvectors.
  → eigenvectors are linearly independent → $\det(\underline{\underline{U}}) \neq 0$ → $\underline{\underline{U}}^{-1}$ exists → must be diagonalisable.
  In addition, $\underline{\underline{U}}$ is an orthogonal matrix, so $\underline{\underline{U}}^{-1} = \underline{\underline{U}}^T$, i.e. $\boxed{\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^T}$

- $\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^{-1}$ takes a vector $\underline{x}$ and returns a new vector $\underline{y}$, where $\underline{x}$ and $\underline{y}$ are in the standard basis. Geometrically, we are doing: $\underline{y} = \underline{\underline{A}}\underline{x} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^{-1}\underline{x}$
  ↳ ① Take the vector $\underline{x}$ and convert it into the new basis to give $\underline{x}'$ $[\underline{x}' = \underline{\underline{U}}^{-1}\underline{x}]$
  ↳ ② Map the vector $\underline{x}'$ to the vector $\underline{y}'$ in the new basis $[\underline{y}' = \underline{\underline{\Lambda}}\underline{x}']$
  ↳ ③ Take the vector $\underline{y}'$ and convert it back to the standard basis to give $\underline{y}$. $[\underline{y} = \underline{\underline{U}}\underline{y}']$.
- We choose a special new basis s.t. the new basis vectors are the eigenvectors of the transformation so in the new basis, the transformation is purely stretching/compressing along the (basis) axes.

Repeated multiplication by a matrix

recurrence relations $\underline{x}_{n+1} = \underline{\underline{A}}\,\underline{x}_n$
($\rightarrow \underline{x}_n = \underline{\underline{A}}^n \underline{x}_0$

- Diagonalising is useful for dealing w/ repeated multiplication by a matrix.
- Let $\underline{\underline{A}}$ be a 3x3 matrix whose eigenvectors are linearly independent (i.e. nondefective matrix).

$$\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^{-1}$$

$$\underline{\underline{A}}^n = (\underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^{-1})(\underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^{-1})\dots(\underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^{-1})$$

As $\underline{\underline{U}}^{-1}\underline{\underline{U}} = \underline{\underline{I}}$, $\boxed{\underline{\underline{A}}^n = \underline{\underline{U}}\,\underline{\underline{\Lambda}}^n\,\underline{\underline{U}}^{-1}}$ where $\underline{\underline{\Lambda}}^n = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix}$

(This proves that $\underline{\underline{A}}^n$ has the same eigenvectors as $\underline{\underline{A}}$ but eigenvalues are raised to power n).

By construction, the eigenvectors of $\underline{\underline{A}}$ u_i are linearly independent, so any vector in 3D space can be expressed as a linear combination of $\underline{u}_1, \underline{u}_2, \underline{u}_3$.

$$\underline{x} = \alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 + \alpha_3 \underline{u}_3 .$$

When mapped by $\underline{\underline{A}}$,

$$\underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}(\alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 + \alpha_3 \underline{u}_3)$$
$$= \alpha_1 \underline{\underline{A}}\,\underline{u}_1 + \alpha_2 \underline{\underline{A}}\,\underline{u}_2 + \alpha_3 \underline{\underline{A}}\,\underline{u}_3$$
$$= \alpha_1 \lambda_1 \underline{u}_1 + \alpha_2 \lambda_2 \underline{u}_2 + \alpha_3 \lambda_3 \underline{u}_3$$

once again we see the eigenvalues of $\underline{\underline{A}}^n$ are λ_i^n.

Repeated multiplication by $\underline{\underline{A}}$, $\boxed{\underline{\underline{A}}^n \underline{x} = \alpha_1 \lambda_1^n \underline{u}_1 + \alpha_2 \lambda_2^n \underline{u}_2 + \alpha_3 \lambda_3^n \underline{u}_3}$

If we have ordered the eigenvalues s.t. $|\lambda_1| > |\lambda_2| > |\lambda_3|$,

consider $\frac{1}{\lambda_1^n} \underline{\underline{A}}^n \underline{x}$. $\quad \frac{1}{\lambda_1^n} \underline{\underline{A}}^n \underline{x} = \alpha_1 \underline{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^n \underline{u}_2 + \alpha_3 \left(\frac{\lambda_3}{\lambda_1}\right)^n \underline{u}_3$

As $n \to \infty$, because $\left|\frac{\lambda_2}{\lambda_1}\right| < 1$ and $\left|\frac{\lambda_3}{\lambda_1}\right| < 1$, $\left(\frac{\lambda_2}{\lambda_1}\right)^n \to 0$ and $\left(\frac{\lambda_3}{\lambda_1}\right)^n \to 0$.

so as $n \to \infty$, $\quad \frac{1}{\lambda_1^n} \underline{\underline{A}}^n \underline{x} \to \alpha_1 \underline{u}_1$ i.e. $\boxed{\underline{\underline{A}}^n \underline{x} \to \alpha_1 \lambda_1^n \underline{x}}$

- When we repeatedly multiply a vector by a matrix,
  - ↳ this eventually becomes simply multiplying by the largest eigenvalue and,
  - ↳ this eventually picks the part of the vector // to the corresponding eigenvector
- We can estimate the largest eigenvalue (and the corresponding eigenvector) by repeatedly multiplying a random vector by the matrix.
- If a vector represents a population of diff species, the largest eigenvalue represents the pop. growth per time interval, and the corresponding normalised eigenvector gives the pop. distribution.

Non-symmetric matrices.

- Non-symmetric matrices, even w/ real entries, generally have complex eigenvalues and eigenvectors. However, they must come in complex conjugate pairs. (real entries).
- Rotation matrices are orthogonal matrices, but not symmetric matrices, so they generally have complex eigenvalues/eigenvectors (also, its eigenvector matrix is not an orthogonal matrix)
- It can be found for the 2D rotation matrix $\underline{\underline{R}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$,

$$\lambda = \cos\theta \pm i\sin\theta \qquad \underline{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \pm \begin{bmatrix} 0 \\ 1 \end{bmatrix} i$$

Generalised eigenvalue problems

Generalised eigenvalue problem

- A non-zero vector $\underline{x}$ that obeys the relation $\boxed{\underline{\underline{A}}\,\underline{x} = \lambda \underline{\underline{M}}\,\underline{x}}$ is an eigenvector of the generalised problem, and λ is the corresponding eigenvalue.
- If $\underline{\underline{M}}$ is invertible, this can be phrased as a regular eigenvalue problem: $\boxed{\underline{\underline{M}}^{-1}\underline{\underline{A}}\,\underline{x} = \lambda \underline{x}}$
- For the jth eigenpair,

$$\underline{\underline{A}}\,\underline{x}_j = \lambda_j \underline{\underline{M}}\,\underline{x}_j$$

Taking the dot product w/ the eigen vector $\underline{x}_i$

$$\underline{x}_i^T \underline{\underline{A}}\,\underline{x}_j = \lambda_j \underline{x}_i^T \underline{\underline{M}}\,\underline{x}_j$$

If the matrices $\underline{\underline{A}}$ and $\underline{\underline{M}}$ are symmetric,

$$\underline{x}_i^T \underline{\underline{A}}\,\underline{x}_j = \underline{x}_j^T \underline{\underline{A}}\,\underline{x}_i \quad [1] \qquad \underline{x}_i^T \underline{\underline{M}}\,\underline{x}_j = \underline{x}_j^T \underline{\underline{M}}\,\underline{x}_i \quad [2]$$

From [1], using the defn of the generalised eigenvalue problem,

$$\underline{x}_i^T \underline{\underline{A}}\,\underline{x}_j = \underline{x}_j^T \underline{\underline{A}}\,\underline{x}_i \quad \rightarrow \quad \lambda_j \underline{x}_i^T \underline{\underline{M}}\,\underline{x}_j = \lambda_i \underline{x}_j^T \underline{\underline{M}}\,\underline{x}_i$$

Now using [2],

$$\lambda_j \underline{x}_j^T \underline{\underline{M}}\,\underline{x}_i = \lambda_i \underline{x}_j^T \underline{\underline{M}}\,\underline{x}_i$$

$$(\lambda_j - \lambda_i)\, \underline{x}_j^T \underline{\underline{M}}\,\underline{x}_i = 0$$

For distinct eigenvalues $\lambda_j \neq \lambda_i$ (when $i \neq j$),

$$\boxed{\underline{x}_j^T \underline{\underline{M}}\,\underline{x}_i = 0} \quad \text{when } i \neq j$$

The eigenvectors $\underline{x}_i$ and $\underline{x}_j$ are said to be orthogonal wrt. $\underline{\underline{A}}$ and $\underline{\underline{M}}$

- For the case $i=j$,

$$\boxed{\underline{x}_i^T \underline{\underline{A}}\,\underline{x}_i = \lambda_i \underline{x}_i^T \underline{\underline{M}}\,\underline{x}_i}$$

$\underline{\underline{M}}$ does not flip any vectors, eigenvalues are +ve

- If the matrix $\underline{\underline{M}}$ is symmetric positive definite (SPD), i.e. $\underline{y}^T \underline{\underline{M}}\,\underline{y} > 0$ for all non-zero $\underline{y}$, we can normalise the eigenvector $\underline{x}_i$ to $\underline{u}_i$ using

$$\boxed{\underline{u}_i = \frac{\underline{x}_i}{\sqrt{\underline{x}_i^T \underline{\underline{M}}\,\underline{x}_i}}}$$

(we req $\underline{\underline{M}}$ is SPD to ensure $\sqrt{\underline{x}_i^T \underline{\underline{M}}\,\underline{x}_i}$ is real)

- Under this normalisation,

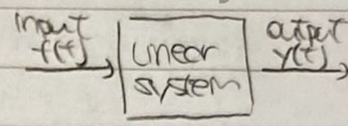
$$\boxed{\underline{u}_i^T \underline{\underline{M}}\,\underline{u}_i = 1} \quad \text{and therefore} \quad \boxed{\underline{u}_i^T \underline{\underline{A}}\,\underline{u}_i = \underline{u}_i^T \lambda_i \underline{\underline{M}}\,\underline{u}_i = \lambda_i}$$

- In summary, given $\underline{\underline{A}}$ and $\underline{\underline{M}}$ are symmetric and w/ normalisation,

$$\boxed{\underline{u}_i^T \underline{\underline{M}}\,\underline{u}_i = \begin{bmatrix} 1 & i=j \\ 0 & i \neq j \end{bmatrix}} \quad \text{and} \quad \boxed{\underline{u}_i^T \underline{\underline{A}}\,\underline{u}_i = \begin{bmatrix} \lambda_i & i=j \\ 0 & i \neq j \end{bmatrix}}$$

Convolution

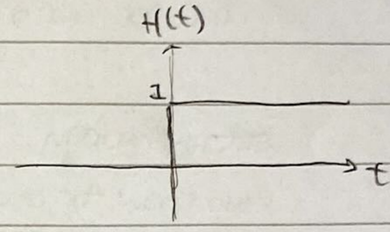
Linear system and impulse response



Linear systems

- A linear time-invariant system satisfies the principle of superposition.
  - ↳ If input $f_1(t)$ gives output $y_1(t)$ and input $f_2(t)$ gives output $y_2(t)$, then input $\alpha f_1(t) + \beta f_2(t)$ gives output $\alpha y_1(t) + \beta y_2(t)$, where α, β are constants.
- For a sinusoidal wave input, a linear system gives a sinusoidal output of the same frequency (but possibly different amplitude / phase).

Step function (Heaviside step function)

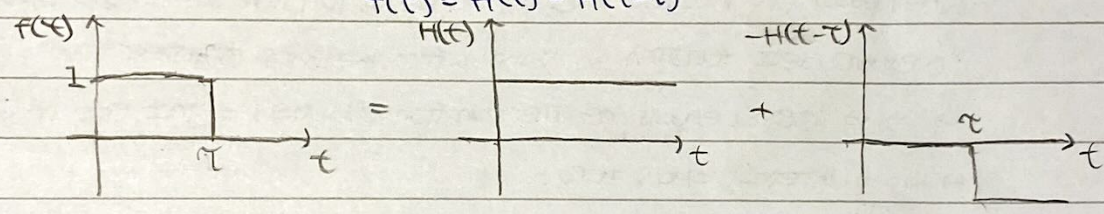


- The step function $H(t)$ is defined as

$$\boxed{H(t) = \begin{cases} 0 & t<0 \\ 1 & t>0 \end{cases}}$$

- A linear system outputs a step response $r(t)$ when given an input step function $H(t)$
- We can express a pulse in terms of step functions via superposition.

$$f(t) = H(t) - H(t-\tau)$$



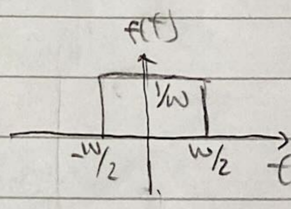
Similarly, we can express the output in terms of step responses via superposition.

$$y(t) = r(t) - r(t-\tau).$$

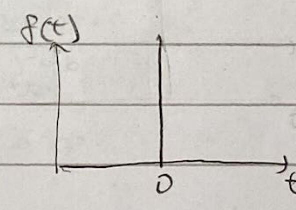
more precisely, $y(t) = r(t)H(t) - r(t-\tau)H(t-\tau)$

* $H(t)$ acts like an on/off switch — we can use it to "combine" piecewise functions.

Delta function (Dirac delta function)



- Consider a pulse of width w and height $1/w$.
  In the limit as $w \to 0$, the pulse becomes a delta function $\delta(t)$
- The delta function is a spike w/ unit area. It goes "bang" when its argument is 0.



$$\boxed{\delta(t) = 0 \quad \text{except at } t=0}$$

$$\boxed{\int_a^b \delta(t)\,dt = 1 \quad \text{provided } 0 \in (a,b)}$$

- A linear system outputs an impulse response $g(t)$ when given an input delta function $\delta(t)$.
- We can express a general input in terms of delta functions via superposition.

$$f(t) = \int_{-\infty}^{t} \delta(t-\tau) f(\tau)\,d\tau$$

Similarly, we can express the output in terms of impulse responses via superposition

$$y(t) = \int_{-\infty}^{t} g(t-\tau) f(\tau)\,d\tau$$

(Refer to convolution for more details)

Relation between the step function and delta function

- $\int_{-\infty}^{t} \delta(t)\,dt = H(t)$ - $\boxed{\frac{d}{dt}[H(t)] = \delta(t)}$

Delta function $\delta(t)$ $\underset{\text{differentiate}}{\overset{\text{integrate}}{\rightleftarrows}}$ Step function $H(t)$

- $\int_{-\infty}^{t} g(t)\,dt = r(t)$ - $\boxed{\frac{d}{dt}[r(t)] = g(t)}$

Impulse response $g(t)$ $\underset{\text{differentiate}}{\overset{\text{integrate}}{\rightleftarrows}}$ step response $r(t)$

where zeros are

* Take care when integrating or differentiating. Double check the boundary conditions and look out for discontinuities if we decompose a function into parts.

Sifting theorem.

- We know for a delta function, $\int_a^c \delta(t-b)\,dt = 1$, provided $b \in (a,c)$.
  Sifting theorem states that $\boxed{\int_a^c \delta(t-b) f(t)\,dt = f(b), \text{ provided } b \in (a,c)}$
  (If $b \notin (a,c)$, then the integral evaluates to 0).
- When evaluating integrals involving the delta function, we should answer the question "Does the delta function go "bang" within the limits of integration".
  ↳ Yes → integral equals the other function evaluated at that time it goes "bang".
  ↳ No → integral equals zero.
- e.g: $\int_{-\pi}^{\pi} \cos(2t)\,\delta(t-\frac{\pi}{2})\,dt$.
  "Bang" at $t=\frac{\pi}{2}$, $\in(-\pi,\pi)$ → $I = \cos(2\cdot\frac{\pi}{2}) = 0$.
- e.g: $\int_0^{\pi} t\,\delta(t+\frac{\pi}{2})\,dt$
  "Bang" at $t=-\frac{\pi}{2}$, $\notin(0,\pi)$ → $I = 0$.

Convolution

Solving differential equations using convolution.

- Linear systems are often described using differential equations, for example
  $a\frac{d^2y}{dt^2} + b\frac{dy}{dt} + cy = f(t)$, where $f(t)$ is the input and $y(t)$ is the output.
- For a general input $f(t)$, we can use convolution to solve for the output as follows:

Differential equation $\xrightarrow{\text{solve}}$ Step response $r(t)$

↓ differentiate

Impulse response $g(t)$

General input $f(t)$ $\xrightarrow{\text{convolution}}$ corresponding output $y(t)$

Finding the impulse response. (and step response)

only proved for step response — not true for impulse response

- To solve for the step response, we set $f(t) = 1$ and solve for just $t \geq 0$. Then, we set the boundary condition $y(0)=0$ ($\dot{y}(0)=0$...) to imply that $f(t)=0$ for $t<0$.
- We can justify that $y(0)=0$ etc. by contradiction → assume $y(0) \neq 0$.
  - ↳ As we are dealing w/ causal systems, $y(t)=0$ for $t<0$. So for $y(0) \neq 0$, there must be a step discontinuity in y at $t=0$.
  - ↳ Assume that y has a step height h at $t=0$ → $\frac{dy}{dt}$ must have a delta function at $t=0$. ($\frac{d^2y}{dt^2}$ is a $\pm$ delta function).
  - ↳ Rearranging the original DE, $a\frac{d^2y}{dt^2} + b\frac{dy}{dt} = -cy + f(t)$. At $t=0$, LHS is infinite and RHS is finite → contradiction. i.e. $y, \frac{dy}{dt}$... must be continuous at $t=0$.

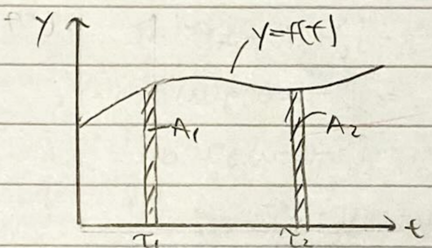
  here, $f(t)$ is finite

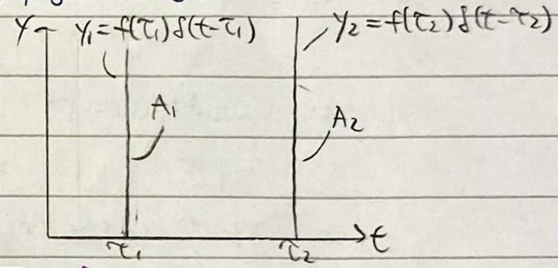
- We can then find the impulse response by differentiating the step response.

* 1st order system: $r(0)$ continuous ; $g(0)$ discontinuous
2nd order system: $r(0), g(0)$ continuous ; $\frac{dg}{dt}(0)$ discontinuous

Convolution integral

- We can break down a general input $f(t)$ into many scaled and delayed impulses that has the same area under the curve for any given interval.





so $f(t) = \int_{-\infty}^{t} f(\tau)\delta(t-\tau)\,d\tau$. (we use for $\int d\tau$ for uncountable no. of strips)

- Therefore the corresponding output would be

$$\boxed{y(t) = \int_{-\infty}^{t} f(\tau) g(t-\tau)\,d\tau}$$

- t is time that relates to the output of the system $y(t)$.
- τ is time that relates to the input of the system $f(\tau)$.

Evaluating convolution integrals — splitting integrals.

- For functions w/ discontinuities, we need to split the convolution integral into several parts.
- Consider the function $f(t) = \begin{cases} f_1(t) & t<t_1 \\ f_2(t) & t_1<t<t_2 \\ f_3(t) & t>t_2 \end{cases}$ as the input
- Given an impulse response $g(t)$, the output $y(t)$ is given by

$$y(t) = \int_{-\infty}^{t} f(\tau) g(t-\tau)\,d\tau = \int_{-\infty}^{t_1} f_1(\tau) g(t-\tau)\,d\tau + \int_{t_1}^{t_2} f_2(\tau) g(t-\tau)\,d\tau + \int_{t_2}^{t} f_3(\tau) g(t-\tau)\,d\tau.$$

(more precisely, $y(t) = \int_{-\infty}^{t_1} f_1(\tau) g(t-\tau)\,d\tau\, H(t) + \int_{t_1}^{t_2} f_2(\tau) g(t-\tau)\,d\tau\, H(t-t_1) + \int_{t_2}^{t} f_3(\tau) g(t-\tau)\,d\tau\, H(t-t_2)$)

- e.g: Given a system has an impulse response of $g(t) = \frac{1}{3}\sin(3t)$. $(t \geq 0)$
  what is the output $y(t)$ when the input is $f(t) = H(t) \cdot t$?
  we can rewrite $f(t)$ as $f(t) = \begin{cases} 0 & t \leq 0 \\ t & t \geq 0 \end{cases}$

so $y(t) = \int_{-\infty}^{t} f(\tau) g(t-\tau) d\tau$
$= \int_{-\infty}^{0} 0 \cdot g(t-\tau) d\tau + \int_{0}^{t} t \cdot g(t-\tau) d\tau$
$= \int_{0}^{t} \frac{1}{3} \cdot \tau \cdot \sin(3(t-\tau))\, d\tau$
$= \frac{1}{9}t - \frac{1}{27}\sin(3t)$

Evaluating convolution integrals — alternate convolution integral

- The normal convolution integral $y(t) = \int_{-\infty}^{t} f(\tau) g(t-\tau) d\tau$ can be inconvenient to compute when we have a complicated expression for $g(t)$.
- If given all signals are zero for $t<0$, we can use an alternate convolution integral that has a term of the form $g(t)$ rather than $g(t-\tau)$.
- Given all signals are zero for $t<0$,

$$y(t) = \int_{0}^{t} f(\tau) g(t-\tau) d\tau.$$

sub $u = t-\tau$, $du = -d\tau$

u	τ
t	0
0	t

$$= \int_{t}^{0} f(t-u) g(u) \cdot (-du)$$
$$= \int_{0}^{t} f(t-u) g(u) du.$$

changing the dummy variable u back to τ,

$$\boxed{y(t) = \int_{0}^{t} f(t-\tau) g(\tau) d\tau}$$ given all signals are zero for $t<0$.

- e.g: Given a system has an impulse response of $g(t) = 3t^2 - 4t + 7$ $(t \geq 0)$
  what is the output $y(t)$ when the input is $f(t) = t$ $(t \geq 0)$
  Note that $g(t) = 0$ and $f(t) = 0$ → we can use the alternate convolution integral
  so $y(t) = \int_{0}^{t} f(t-\tau) g(\tau) d\tau$
  $= \int_{0}^{t} (t-\tau) \cdot (3\tau^2 - 4\tau + 7) d\tau$
  $= \frac{1}{4}t^4 - \frac{2}{3}t^3 + \frac{7}{2}t^2$

Avoiding convolution integrals — special cases.

- If the input $f(t)$ only consists of step functions, the output $y(t)$ is obtained by replacing the step functions w/ the step response. more precisely,
- e.g. $f(t) = vH(t) - vH(t-k)$. $y(t) = vr(t) - vr(t-k)$. / $y(t) = vr(t)H(t) - vr(t-k)H(t-k)$

* Similar idea applies to delta function inputs, but such inputs are uncommon.

Spatial convolution

- So far, we only looked at systems that have inputs and outputs that vary as a function of time. These systems are causal – there is no output before the input that causes it.

  i.e. $g(t) = 0$ for $t<0$

- However, systems that have inputs and outputs that vary as a function of spatial location can have $g(x) \neq 0$ for any x. → use $\boxed{y(x) = \int_{-\infty}^{\infty} g(x-\tau) f(\tau) d\tau}$

  $g(x)$ is the displacement at position x for a unit load at $x=0$

  different limits of integration than temporal / causal systems

- e.g: Consider a 1D strip of material that deforms linearly according to $g(x) = \frac{1}{\cosh x}$ when subject to a unit force at $x=0$.

  Calculate the deformation of the strip in response to a uniform load of $f(x) = 1$ applied from $x=0$ to $x=2$.

$$y(x) = \int_{-\infty}^{\infty} g(x-\tau) f(\tau) d\tau$$
$$= \int_{-\infty}^{0} \cancel{\frac{1}{\cosh(x-\tau)}} \cdot 0 \, d\tau + \int_{0}^{2} \frac{1}{\cosh(x-\tau)} \cdot 1 \, d\tau + \int_{2}^{\infty} \cancel{\frac{1}{\cosh(x-\tau)}} \cdot 0 \, d\tau$$
$$= 2[\arctan(e^{2-x}) - \arctan(e^{-x})]$$

Variable impulse response.

- In a real world example, the spatial impulse response $g(x)$ would also depend on the position of the pt. load a. → we use a function $g(x,a)$.

  ($g(x,a)$ is the displacement at position x for a unit load at position a).

- Adding up all the contributions from the pt loads, we get

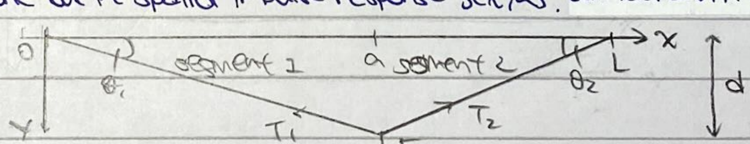
  so we don't need to shift g using "$x-a$".

$$\boxed{y(x) = \int_{-\infty}^{\infty} g(x,a) f(a) da}$$

* If $g(x,a)$ is a piecewise function, be careful when choosing which part of $g(x,a)$ to use for each individual integral.

- e.g: Consider a taut string suspended between 2 pts a distance L apart. It is subject to a uniform load K per unit length which results in a small displacement

First, we need to work out the spatial impulse response $g(x,a)$. Consider a unit pt. load at $x=a$.



For small displacements, $\theta_1 \approx \theta_2 \approx 0$ → $\cos\theta_1 \approx \cos\theta_2 \approx 1$

N2L ⇔ $\quad T_1 \cos\theta_1 = T_2 \cos\theta_2 \quad → \quad T_1 = T_2 = T$.

N2L ↕ $\quad T_1 \sin\theta_1 + T_2 \sin\theta_2 = 1 \quad → \quad T(\tan\theta_1 + \tan\theta_2) = 1$

$$\therefore \left(\frac{d}{a} + \frac{d}{L-a}\right) = \frac{1}{T} \quad → \quad d = \frac{a(L-a)}{TL}$$

After finding the max displacement d for a unit load at $x=a$, we can work out $g(x,a)$ using geometry.

Segment 1, $x<a$ $\quad g(x,a)_1 = \frac{x}{a} d = \frac{x(L-a)}{TL}$

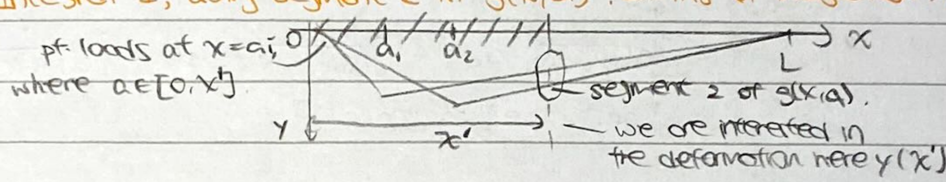
Segment 2, $x>a$. $\quad g(x,a)_2 = \left(\frac{L-x}{L-a}\right) d = \frac{a(L-x)}{TL}$

Now we can apply the variable spatial convolution integral.

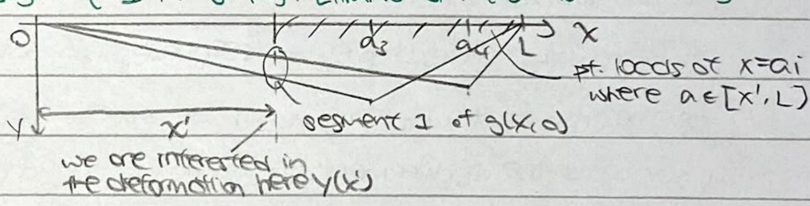
$$
\begin{aligned}
y(x) &= \int_{-\infty}^{\infty} g(x,a) f(a)\, da \\
&= \int_{-\infty}^{0} \cancel{g(x,a)}^{\,0} \cdot 0\, da + \int_0^L g(x,a) \cdot K \cdot da + \int_L^{\infty} \cancel{g(x,a)}^{\,0} \cdot 0\, da \\
&= \int_0^x g(x,a)_2 \cdot K da + \int_x^L g(x,a)_1 \cdot K da \\
&= \int_0^x \frac{a(L-x)}{TL} K\, da + \int_x^L \frac{x(L-a)}{TL} K\, da \\
&= \frac{K}{2T} x(L-x)
\end{aligned}
$$

* A closer look of the integrals :

Integral 1, using segment 2 in $g(x,a)$. [limits of integration : $0, x$].



Integral 2, using segment 1 in $g(x,a)$. [limits of integration : x, L]

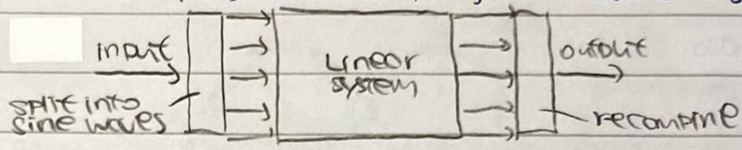


Fourier series

Background

Motivation

- Sinusoids are easy to deal w/ when considering linear systems — For a linear system, a sinusoidal input gives a sinusoidal output of the same frequency. (The output sinusoid may have a different amplitude or phase).
- It would be useful to express an arbitrary signal in terms of a sum of sine waves.



Analogy between vectors and functions

- Vectors and functions are analogous in many ways:
- 2 vectors $\underline{u}, \underline{v}$ are orthogonal if $\underline{u} \cdot \underline{v} = 0$. ← inner product $= 0$.

  2 functions f, g are orthogonal on the interval $[a,b]$ if $\langle f, g \rangle = 0$.

  where $\langle f, g \rangle = \int_a^b f(x) g^*(x)\,dx$ (g^* is the complex conjugate of g).
- We can represent an arbitrary vector $\underline{r}$ in terms of a linear combination of (orthogonal) basis vectors $\underline{e}_i$.

$$\underline{r} = \sum_i a_i \underline{e}_i, \quad \text{where} \quad a_i = \frac{\underline{r} \cdot \underline{e}_i}{\underline{e}_i \cdot \underline{e}_i}$$

We can represent an arbitrary function f in terms of a linear combination of orthogonal basis functions ϕ_i

$$f(x) = \sum_i a_i \phi_i(x), \quad \text{where} \quad a_i = \frac{\langle f, \phi_i \rangle}{\langle \phi_i, \phi_i \rangle}$$

* To have a set of orthogonal basis functions ϕ_i, any 2 distinct functions in the set are orthogonal to each other,

  i.e. $\langle \phi_m, \phi_n \rangle = 0$ when $n \neq m$.

Real Fourier series

Real Fourier series

- The basis functions for the real Fourier series are

$$1, \ \cos(nt), \ \sin(nt) \qquad \text{where } n \in \mathbb{Z}^+$$

actually, any interval w/ width of 2π.

- These basis functions form an orthogonal set under the interval $[-\pi, \pi]$.
- We can represent $f(t)$ between $-\pi$ and $+\pi$ using a linear combination of $1, \cos(nt), \sin(nt)$:

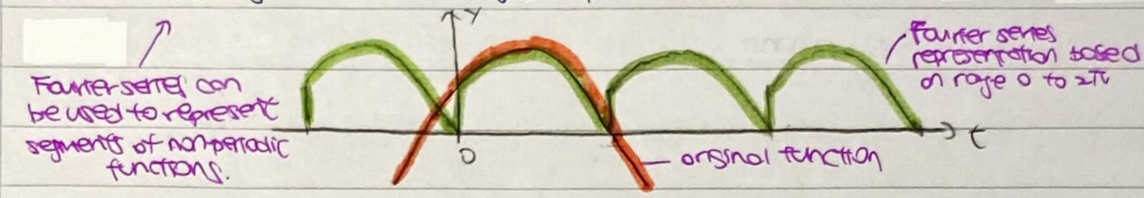
$$\boxed{f(t) = d \cdot 1 + \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt)}$$

where

$$\boxed{\begin{aligned} d &= \frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \cdot 1\,dt \\ a_n &= \frac{\langle f, \cos nt \rangle}{\langle \cos nt, \cos nt \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cdot \cos(nt)\,dt \\ b_n &= \frac{\langle f, \sin nt \rangle}{\langle \sin nt, \sin nt \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cdot \sin(nt)\,dt \end{aligned}}$$

Fourier series properties

- We can use any range of length 2π instead of $-\pi \le t \le \pi$ in the Fourier formulae.
  - e.g. We can use $0 \le t \le 2\pi$ and the limits of integration of the coefficient integrals would be $0, 2\pi$ instead of $-\pi, +\pi$.
- We are only modelling the function $f(t)$ in the specified range (e.g. $-\pi$ to π, 0 to 2π). Outside this range, the model will just repeat w/ a period of 2π.



Real general range Fourier series

- If we want to model a periodic signal w/ period other than 2π, or a section of a non-periodic signal of length other than 2π, we need a more general formula.
- To model a function $f(x)$ over the range 0 to L, substitute $t = \frac{2\pi x}{L}$ $\left[\frac{t}{2\pi} = \frac{x}{L}\right]$

Now, our basis functions are

$$1, \cos\left(\frac{2\pi nx}{L}\right), \sin\left(\frac{2\pi nx}{L}\right), \quad \text{where } n \in \mathbb{Z}^+$$

actually, any interval w/ width of L

- these basis functions form an orthogonal set under the interval $[0, L]$
- We can represent $f(x)$ between 0 and L using a linear combination of $1, \cos\left(\frac{2\pi nx}{L}\right), \sin\left(\frac{2\pi nx}{L}\right)$:

$$\boxed{f(x) = a \cdot 1 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi nx}{L}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi nx}{L}\right)}$$

where

$$\boxed{\begin{aligned} a &= \frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{L}\int_0^L f(x) \cdot 1 \, dx \\ a_n &= \frac{\langle f, \cos\frac{2\pi nx}{L} \rangle}{\langle \cos\frac{2\pi nx}{L}, \cos\frac{2\pi nx}{L} \rangle} = \frac{2}{L}\int_0^L f(x) \cdot \cos\left(\frac{2\pi nx}{L}\right) dx \\ b_n &= \frac{\langle f, \sin\frac{2\pi nx}{L} \rangle}{\langle \sin\frac{2\pi nx}{L}, \sin\frac{2\pi nx}{L} \rangle} = \frac{2}{L}\int_0^L f(x) \cdot \sin\left(\frac{2\pi nx}{L}\right) dx \end{aligned}}$$

- The fraction $\frac{2\pi}{L}$ is often written as ω_0, the fundamental angular frequency.

Complex Fourier series

✓ easier to evaluate Fourier integrals (esp. $f(x) = \exp(\,)$).
× Fourier integrals may be complex

Complex Fourier series

- the basis functions for the complex Fourier series are

$$e^{int} \quad \text{where } n \in \mathbb{Z}$$

- These basis functions form an orthogonal set under the interval $[-\pi, \pi]$.
- We can represent $f(t)$ between $-\pi$ to $+\pi$ using a linear combination of e^{int}.

$$\boxed{f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}}$$

where $\boxed{c_n = \frac{\langle f, e^{int} \rangle}{\langle e^{int}, e^{int} \rangle} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \cdot e^{-int} dt}$

* c_n can be complex.

Complex general range Fourier series

- If we want to model a periodic signal w/ period other than 2π, or even a section of a non-periodic signal of length other than 2π, we need a more general formula.
- To model a function $f(x)$ over the range 0 to L, substitute $t = \frac{2\pi x}{L}$ $\left[\frac{t}{2\pi} = \frac{x}{L}\right]$

Now, our basis functions are

$$e^{\frac{i2\pi n x}{L}} \quad \text{where } n \in \mathbb{Z}$$

- These basis functions form an orthogonal set under the interval $[0, L]$
- We can represent $f(x)$ between 0 and L using a linear combination of $e^{\frac{i2\pi n x}{L}}$

$$\boxed{f(x) = \sum_{n=-\infty}^{\infty} C_n e^{\frac{i2\pi n x}{L}}}$$

where $\boxed{C_n = \frac{\langle f, e^{\frac{i2\pi n x}{L}} \rangle}{\langle e^{\frac{i2\pi n x}{L}}, e^{\frac{i2\pi n x}{L}} \rangle} = \frac{1}{L}\int_0^L f(x)\, e^{-\frac{i2\pi n x}{L}} dx}$

* C_n can be complex.

- The fraction $\frac{2\pi}{L}$ is often written as ω_0, the fundamental angular frequency.

Conversion between real and complex Fourier series.

- Consider the real Fourier series w/ period 2π. — similar logic for general range Fourier series

$$\begin{aligned} f(t) &= d + \sum_{n=1}^{\infty}\left[a_n \cos(nt) + b_n \sin(nt)\right] \\ &= d + \sum_{n=1}^{\infty}\left[a_n\left(\frac{e^{int}+e^{-int}}{2}\right) + b_n\left(\frac{e^{int}-e^{-int}}{2i}\right)\right] \\ &= d + \sum_{n=1}^{\infty}\left(\frac{a_n - ib_n}{2}\right)e^{int} + \sum_{n=1}^{\infty}\left(\frac{a_n+ib_n}{2}\right)e^{-int} = \sum_{n=-\infty}^{\infty} C_n e^{int} \end{aligned}$$

- Comparing the coefficients, we find that

$$\boxed{C_n = \begin{cases} d, & n = 0 \\ \frac{a_n - ib_n}{2}, & n = \mathbb{Z}^+ \\ \frac{a_{-n} + ib_{-n}}{2}, & n = \mathbb{Z}^- \end{cases}} \longrightarrow \boxed{\begin{aligned} d &= C_0 \\ a_n &= 2\,\mathrm{Re}\{C_n\} \\ b_n &= -2\,\mathrm{Im}\{C_n\} \end{aligned}}$$

* a_{-n} and b_{-n} are only defined when n is -ve.

- We can show that $C_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt$ from the real Fourier series.

For $n = 0$, $\quad C_0 = d$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-0}dt$$

For +ve n, $\quad C_n = \frac{a_n - ib_n}{2}$

$$= \frac{1}{2}\left[\frac{1}{\pi}\int_{-\pi}^{\pi}[\cos(nt) - i\sin(nt)]f(t)\,dt\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt$$

For -ve n, $\quad C_n = \frac{a_{-n} + ib_{-n}}{2}$

$$= \frac{1}{2}\left[\frac{1}{\pi}\int_{-\pi}^{\pi}[\cos(-nt) + i\sin(-nt)]f(t)\,dt\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt.$$

Techniques for finding the Fourier series

Avoiding integration : symmetric signals

- odd function : $f(-t) = -f(t)$; Even function : $f(-t) = f(t)$
- The a_n terms (cos nt terms) model the even component in the function ; The b_n terms (sin nt terms) model the odd component in the function.
- The d term models the mean value of the function.
- If we can spot a symmetry in the function to be represented, we can avoid evaluating one or more of the Fourier integrals.
  - ↳ No even component → all $a_n = 0$ (no cos nt terms) / odd function
  - ↳ No odd component → all $b_n = 0$ (no sin nt terms) — even function / shifted odd function
  - ↳ zero mean → $d = 0$

  * If $a_n = 0$ or $b_n = 0$, the function must have the same symmetries as the basis function.

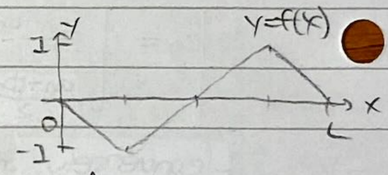
(odd function → $a_n = 0$, $d = 0$; Even function → $b_n = 0$).

Avoiding (difficult) integration : Converting to delta functions and back.

- Integrating delta functions is trivial (sifting theorem) so if we can differentiate our signal into delta functions, we can easily find the Fourier series and integrate back to obtain the Fourier series of the original function.
- Triangular / saw-tooth $\underset{\text{integrate}}{\overset{\text{differentiate}}{\rightleftarrows}}$ Square / Pulse $\underset{\text{integrate}}{\overset{\text{differentiate}}{\rightleftarrows}}$ Delta

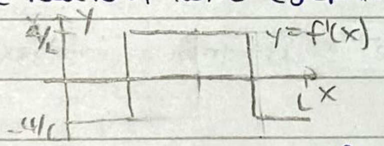
* The constant of integration is simply the constant term of the Fourier series, i.e. d, the mean value of the function.

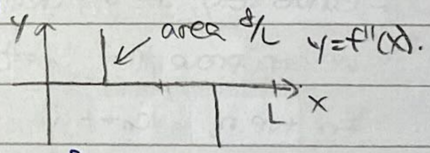
- e.g. Find the Fourier series for $f(x)$ between 0 and L



Differentiate the waveform twice (graphically)

careful w/ the magnitudes





Graphically, we can see that $f''(x) = \frac{8}{L}\delta(x - \frac{L}{4}) - \frac{8}{L}\delta(x - \frac{3L}{4})$

$f''(x)$ is odd and has zero mean → $a_n = 0$, $d = 0$.

$$
\begin{aligned}
b_n &= \frac{2}{L}\int_0^L f''(x)\sin\left(\frac{2\pi n x}{L}\right)dx \\
&= \frac{16}{L^2}\int_0^L \left[\delta(x - \tfrac{L}{4})\sin\left(\tfrac{2\pi n x}{L}\right) - \delta(x - \tfrac{3L}{4})\sin\left(\tfrac{2\pi n x}{L}\right)\right]dx \\
&= \frac{16}{L^2}\left[\sin\left(\frac{2\pi n \cdot L/4}{L}\right) - \sin\left(\frac{2\pi n \cdot 3L/4}{L}\right)\right] \\
&= \frac{16}{L^2}\left[\sin\left(\frac{\pi n}{2}\right) - \sin\left(\frac{3\pi n}{2}\right)\right] \\
&= \begin{cases} 0 & , n \text{ even} \\ \frac{32}{L^2}(-1)^{\left(\frac{n+3}{2}\right)} & , n \text{ odd} \end{cases}
\end{aligned}
$$

∴ Fourier series for $f''(x)$: $f''(x) = \frac{32}{L^2}\sum_{\substack{n=1 \\ n\,\text{odd}}}^{\infty}(-1)^{\left(\frac{n+3}{2}\right)}\sin\left(\frac{2\pi n x}{L}\right)$

Using $2m-1 = n$, $f''(x) = \frac{32}{L^2}\sum_{m=1}^{\infty}(-1)^{m+1}\sin\left(\frac{2\pi(2m-1)x}{L}\right)$.

Now, we integrate twice and set both integration constants to 0 (since mean of $f(x)$ and $f'(x)$ are both 0)

$$f''(x) = \frac{32}{L^2}\sum_{m=1}^{\infty}(-1)^{m+1}\sin\left(\frac{2\pi x(2m-1)}{L}\right)$$
$$f'(x) = \frac{16}{\pi L}\sum_{m=1}^{\infty}\frac{(-1)^m}{2m-1}\cos\left(\frac{2\pi x(2m-1)}{L}\right)$$
$$f(x) = \frac{8}{\pi^2}\sum_{m=1}^{\infty}\frac{(-1)^m}{(2m-1)^2}\sin\left(\frac{2\pi x(2m-1)}{L}\right)$$

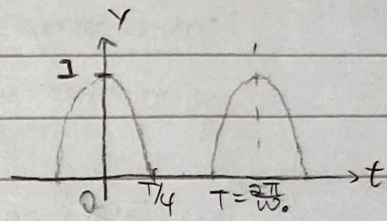
Avoiding integration: Look up the Fourier series in the databook.

- Sometimes, we can simply look up the Fourier series of a similar waveform in the databook and use a substitution of variables to find the req. Fourier series.
- Here are the Fourier series in the databook:

① Half-wave rectified cosine wave.

$$f(t) = \frac{1}{\pi} + \frac{1}{2}\cos\omega_0 t + \frac{2}{\pi}\sum_{m=1}^{\infty}(-1)^{m+1}\frac{\cos 2m\omega_0 t}{4m^2-1}$$
$$f(t) = \frac{1}{\pi} + \frac{1}{4}e^{i\omega_0 t} + \frac{1}{4}e^{-i\omega_0 t} + \frac{1}{\pi}\sum_{\substack{n=-\infty\\ n\ \text{even}\\ n\neq 0}}^{\infty}(\pm 1)\frac{e^{in\omega_0 t}}{n^2-1}$$

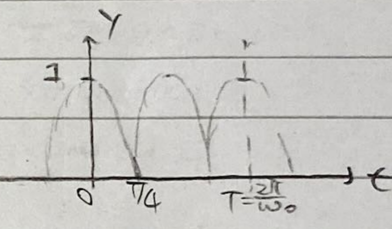
signs alternate, + for $n=2$



② Full-wave rectified cosine wave

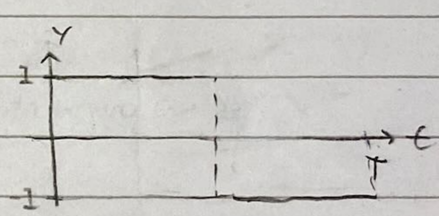
$$f(t) = \frac{2}{\pi}\left[1 + 2\sum_{m=1}^{\infty}(-1)^{m+1}\frac{\cos 2m\omega_0 t}{4m^2-1}\right]$$
$$f(t) = \frac{2}{\pi}\left[1 + \sum_{\substack{n=-\infty\\ n\ \text{even}\\ n\neq 0}}^{\infty}(\pm 1)\frac{e^{in\omega_0 t}}{n^2-1}\right]$$

signs alternate, + for $n=2$



③ Square wave

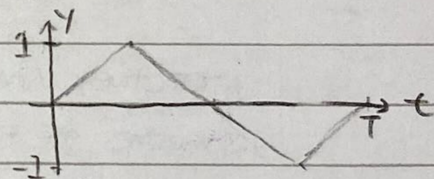
$$f(t) = \frac{4}{\pi}\sum_{m=1}^{\infty}\frac{\sin(2m-1)\omega_0 t}{2m-1}$$
$$f(t) = \sum_{\substack{n=-\infty\\ n\ \text{odd}}}^{\infty}\frac{2}{i\pi n}e^{in\omega_0 t}$$



④ Triangular wave

$$f(t) = \frac{8}{\pi^2}\sum_{m=1}^{\infty}(-1)^{m+1}\frac{\sin(2m-1)\omega_0 t}{(2m-1)^2}$$
$$f(t) = \frac{4}{i\pi^2}\sum_{\substack{n=-\infty\\ n\ \text{odd}}}^{\infty}(\pm 1)\frac{e^{in\omega_0 t}}{n^2}$$

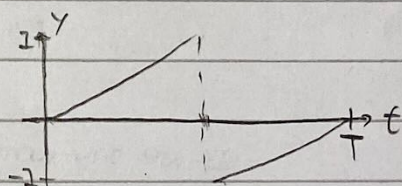
signs alternate, + for $n=1$



⑤ Saw-tooth wave

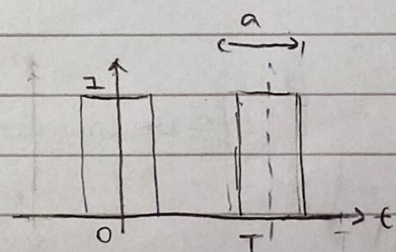
$$f(t) = \frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\sin n\omega_0 t}{n}$$
$$f(t) = \frac{1}{i\pi}\sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}(\pm 1)\frac{e^{in\omega_0 t}}{n}$$

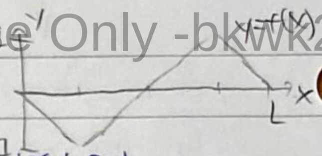
signs alternate, + for $n=1$



⑥ Pulse wave

$$f(t) = \frac{a}{T}\left[1 + 2\sum_{n=1}^{\infty}\frac{\sin\frac{n\pi a}{T}}{\frac{n\pi a}{T}}\cos n\omega_0 t\right]$$
$$f(t) = \frac{a}{T}\left[1 + \sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\frac{\sin\frac{n\pi a}{T}}{\frac{n\pi a}{T}}e^{in\omega_0 t}\right]$$





- e.g: Find the Fourier series for f(x) between 0 and L

This waveform is similar to the triangular wave in the databook.

↳ we just need to scale by -1 and transform the coords.

From the databook, $f(t) = \frac{8}{\pi^2} \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\sin(2m-1)\omega_0 t}{(2m-1)^2}$

Applying the transformation, $f(x) = -\frac{8}{\pi^2} \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\sin(2m-1)\frac{2\pi}{L} \cdot x}{(2m-1)^2}$

$= \frac{8}{\pi^2} \sum_{m=1}^{\infty} (-1)^{m} \frac{\sin(2m-1)\frac{2\pi}{L} \cdot x}{(2m-1)^2}$

Pick the start of period / interval carefully

- For functions that have delta functions, choose an interval s.t. there are no delta functions at the ends of the interval
- This is because it is not clear what to do about delta functions that coincide w/ the upper and lower limits of the integral

Odd / Even n only. and alternating sign.

- sometimes, the Fourier integral evaluates to a term that is zero for odd (even) n and non-zero for even (odd) n
- If we only want the even terms, we use the substitution $n = 2m$.

n	1	2	3	4	5
$2m$	2	4	6	8	10

- If we only want the odd terms, we use the substitution $n = 2m - 1$

n	1	2	3	4	5
$2m-1$	1	3	5	7	9

- Alternatively, simply write "n odd" or "n even" under the summation symbol.
- Sometimes, the Fourier integral evaluates to a term that alternates between +1 and -1.
- If we want all n, we use $(-1)^n$ or $(-1)^{n+1}$

n	1	2	3	4	5
$(-1)^n$	-1	1	-1	1	-1

most useful.
both cases below can be converted to this using $n=2m$ or $n=2m-1$

- If we only want even n, we use $(-1)^{\frac{n}{2}}$ or $(-1)^{(\frac{n}{2}+1)}$

even: $n = 2m$
$m = \frac{n}{2}$

n	2	4	6	8	10
$(-1)^{\frac{n}{2}}$	-1	1	-1	1	-1

- If we only want odd n, we use $(-1)^{(\frac{n+1}{2})}$ or $(-1)^{(\frac{n+1}{2}+1)}$

odd: $n = 2m-1$
$m = \frac{n+1}{2}$

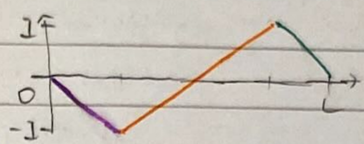
n	1	3	5	7	9
$(-1)^{\frac{n+1}{2}}$	-1	1	-1	1	-1

- Alternatively, simply write "±1" and specify true for n=1 or n=2.

Avoiding (difficult) integration: Using symmetries to combine integrals.

- sometimes, there are symmetries between piecewise function f(x) and $\sin(\frac{2\pi n x}{L})/\cos(\frac{2\pi n x}{L})$
  → we can combine integrals to simplify the steps.

(sometimes, these symmetries lead to odd/even terms evaluating to zero).



- e.g.: Find the Fourier series for f(x) between 0 and L

  The waveform can be represented as a piecewise function

piecewise function

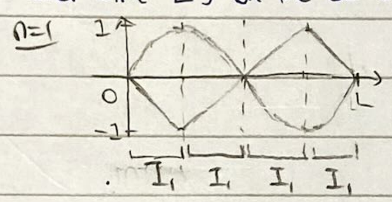
$$f(x) = \begin{cases} -\frac{4x}{L} & x \in (0, L/4) \\ \frac{4x}{L} - 2 & x \in (L/4, 3L/4) \\ 4 - \frac{4x}{L} & x \in (3L/4, L) \end{cases}$$

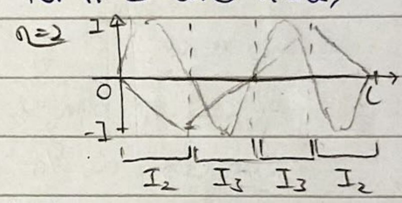
f(x) is odd and has zero mean → $a_n = 0$, $a = 0$.

$$b_n = \frac{2}{L}\int_0^L f(x)\sin(\tfrac{2\pi n x}{L})\,dx$$

$$= \frac{2}{L}\int_0^{L/4}\left(-\frac{4x}{L}\right)\sin(\tfrac{2\pi n x}{L})\,dx + \frac{2}{L}\int_{L/4}^{3L/4}\left(\frac{4x}{L}-2\right)\sin(\tfrac{2\pi n x}{L})\,dx + \frac{2}{L}\int_{3L/4}^{L}\left(4-\frac{4x}{L}\right)\sin(\tfrac{2\pi n x}{L})\,dx$$

plotting f(x) and $\sin(\frac{2\pi n x}{L})$ on the same axis for n = 1 and n = 2,





$I_2 = -I_3$

There is some symmetry between f(x) and $\sin(\frac{2\pi n x}{L})$.

For odd n, $b_n = 4I_1 = 4\cdot\frac{2}{L}\int_0^{L/4}\left(-\frac{4x}{L}\right)\sin(\tfrac{2\pi n x}{L})\,dx$

$$= \frac{8}{n^2\pi^2}\left(\frac{n\pi}{2}\cancel{\cos(\tfrac{n\pi}{2})}^{\,0} - \sin(\tfrac{n\pi}{2})\right)$$

For even n, $b_n = 0$.

consider $-\sin(\frac{n\pi}{2})$ for odd n

n	1	3	5	7
$-\sin(\frac{n\pi}{2})$	-1	1	-1	1

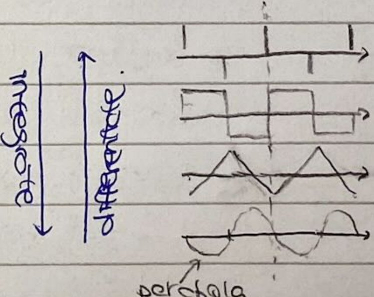
$$-\sin(\tfrac{n\pi}{2}) = (-1)^{\frac{n+1}{2}}.$$

$$\therefore f(x) = \sum_{\substack{n=1\\ n\text{ odd}}}^{\infty} \frac{8}{n^2\pi^2}(-1)^{\frac{n+1}{2}}\cdot\sin(\tfrac{2\pi n x}{L})$$

$$= \sum_{m=1}^{\infty} \frac{8}{(2m-1)^2\pi^2}(-1)^m \sin\left(\frac{2\pi(2m-1)x}{L}\right)$$

convergence of Fourier series

- series for delta functions → does not converge
- Discontinuous value → converges as $1/n$
- Discontinuous slope → converges as $1/n^2$
- Discontinuous 2nd derivative → converges as $1/n^3$

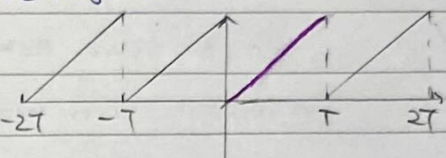


Half range series

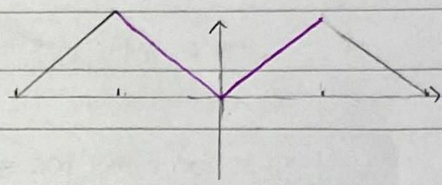
- If we are modelling a limited section of a function, we can pick a Fourier series period so as to get good convergence and a series that is easy to compute (symmetries)

- e.g: We want to model a signal $f(x) = x$ in the range 0 to T.

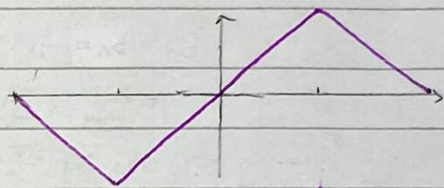
① Full range series
- Period T, $a_n = 0$.
- Converges as $1/n$



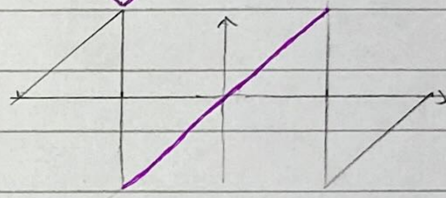
② Cosine series
- Period $2T$, $b_n = 0$
- Converges as $1/n^2$



③ Sine series
- Period $4T$, $a_n = 0$, $d = 0$
- Converges as $1/n^2$



④ Sine series
- Period $2T$, $a_n = 0$, $d = 0$
- Converges as $1/n$



① Full range series. (Period T)

Note it is a shifted odd function $\rightarrow a_n = 0$

$b_n = \frac{2}{T}\int_0^T x\sin\left(\frac{2\pi n x}{T}\right)dx = -\frac{T}{n\pi}$ $\qquad d = \frac{1}{T}\int_0^T x\,dx = \frac{T}{2}$

$\therefore f(x) = \frac{T}{2} - \frac{T}{\pi}\sum_{n=1}^{\infty}\boxed{\frac{1}{n}}\sin\left(\frac{2\pi n x}{T}\right)$ converges as $1/n$

② cosine series. (period 2T)

Note it is an even function $\rightarrow b_n = 0$.

$a_n = \frac{1}{T}\int_{-T}^{T} f(x)\cos\left(\frac{\pi n x}{T}\right)dx$ $\qquad d = \frac{1}{2T}\int_{-T}^{T}|x|\,dx = \frac{T}{2}$

$= \frac{1}{T}\left[\int_{-T}^{0} -x\cos\left(\frac{\pi n x}{T}\right)dx + \int_0^T x\cos\left(\frac{\pi n x}{T}\right)dx\right]$

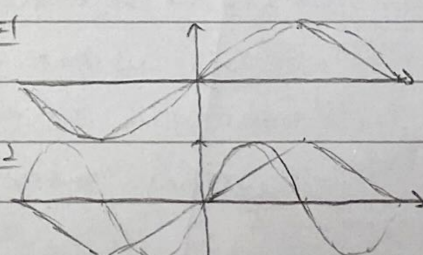
$= \frac{2T}{n^2\pi^2}\left(n\pi\cancel{\sin(n\pi)}^{0} + \cos(n\pi) - 1\right)$

$= -\frac{4T}{n^2\pi^2}$, only odd n.

$\therefore f(x) = \frac{T}{2} - \frac{4T}{\pi^2}\sum_{\substack{n=1\\ n\text{ odd}}}^{\infty}\boxed{\frac{1}{n^2}}\cos\left(\frac{\pi n x}{T}\right) = \frac{T}{2} - \frac{4T}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^2}\cos\left(\frac{\pi x(2m-1)}{T}\right)$ converges as $1/n^2$

③ sine series (period 4T)

Note it is an odd function $\rightarrow a_n = 0,\ d = 0$.



$b_n = \frac{1}{2T}\int_{-2T}^{2T} f(x)\sin\left(\frac{\pi n x}{2T}\right)dx$

$= \frac{4}{2T}\int_0^T x\sin\left(\frac{\pi n x}{2T}\right)dx$

$= \frac{4T}{n^2\pi^2}\left(2\sin\left(\frac{n\pi}{2}\right) - n\pi\cancel{\cos\left(\frac{n\pi}{2}\right)}^{0}\right)$

$= \frac{8T}{n^2\pi^2}(-1)^{\left(\frac{n+1}{2}+1\right)}$, only odd n

$\therefore f(x) = \frac{8T}{\pi^2}\sum_{\substack{n=1\\ n\text{ odd}}}^{\infty}\boxed{\frac{1}{n^2}}(-1)^{\left(\frac{n+3}{2}\right)}\sin\left(\frac{n\pi x}{2T}\right) = \frac{8T}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^2}(-1)^{m+1}\sin\left(\frac{\pi x(2m-1)}{2T}\right)$ converges as $1/n^2$

④ Sine series (Period $2T$)

Note it is an odd function $\rightarrow a_n = 0, d = 0$

$$b_n = \frac{1}{T}\int_{-T}^{T} x\sin\left(\frac{\pi n x}{T}\right)dx$$
$$= \frac{2T}{n^2\pi^2}\left(\cancel{\sin(n\pi)}^{0} - n\pi\cos(n\pi)\right)$$
$$= \frac{2T}{n\pi}(-1)^{n+1}$$
$$\therefore f(x) = \frac{2T}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}(-1)^{n+1}\sin\left(\frac{\pi n x}{T}\right)$$ converges as $1/n$

General strategy for solving Fourier series questions.

- For trigonometric functions, use identities to simplify to $\sin(nt)$ / $\cos(nt)$.
  - ↳ Harmonic form : $k\sin(nt+\phi) \rightarrow A\sin(nt) + B\cos(nt)$
  - ↳ Powers : $a\sin^k(nt) \rightarrow$ De Moivre + Binomial
- For exponentials, use complex Fourier series.
- For Half/full-wave rectified sinusoid / square / triangular / sawtooth / pulse, check the databook. If no translation in x-axis $\rightarrow$ use databook.
- For square / triangular / sawtooth / pulse, differentiate to get delta functions, find the Fourier series and integrate back.
- For delta functions, choose an interval s.t. there are no delta functions at the limits.
- Avoid piecewise functions if possible $\rightarrow$ use half/quarter range series.
- Never directly integrate piecewise functions, use other methods! If no choice, look for symmetries to combine integrals (sometimes it turns out to be 0 for odd/even n)

* Always check if a_n, b_n, d are zero before integrating (look out for shifted odd function)

* $\sin(n\pi) = 0$, $\cos\left(\frac{n\pi}{2}\right) = 0$ for $n \in \mathbb{Z}$. For alternating sign, use $(-1)^n, (-1)^{\frac{n}{2}}, (-1)^{\frac{n+1}{2}}$ etc.

Difficult Fourier series questions.

- e.g.: Find the complex Fourier series to model $f(t) = \sin(t)$.

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sin(t)e^{-int}\,dt = \frac{1}{2\pi}\left[\frac{e^{in\pi} - e^{-in\pi}}{n^2 - 1}\right].$$

$c_n = 0$ for $n \neq \pm 1$; c_n is undefined if we substitute $n = \pm 1$ directly.

$\therefore$ we set $n = \pm 1 + \varepsilon$ and calculate the limit of c_n as ε tends to 0.

$$c_1 = \lim_{\varepsilon\to 0}\frac{1}{2\pi}\left[\frac{e^{i\pi(1+\varepsilon)} - e^{-i\pi(1+\varepsilon)}}{(1+\varepsilon)^2 - 1}\right]$$
$$= \frac{1}{2\pi}\lim_{\varepsilon\to 0}\left[\frac{e^{i\pi}e^{i\pi\varepsilon} - e^{-i\pi}e^{-i\pi\varepsilon}}{\cancel{1} + 2\varepsilon + \varepsilon^2 \cancel{-1}}\right]$$
$$= \frac{1}{2\pi}\lim_{\varepsilon\to 0}\left[\frac{-e^{i\pi\varepsilon} + e^{-i\pi\varepsilon}}{2\varepsilon + O(\varepsilon^2)}\right]$$
$$= \frac{1}{2\pi}\lim_{\varepsilon\to 0}\left[\frac{\cancel{-1} - i\pi\varepsilon \cancel{+1} - i\pi\varepsilon}{2\varepsilon + O(\varepsilon^2)}\right]$$
$$= \frac{1}{2i}$$

$$c_{-1} = \lim_{\varepsilon\to 0}\frac{1}{2\pi}\left[\frac{e^{i\pi(-1+\varepsilon)} - e^{-i\pi(-1+\varepsilon)}}{(-1+\varepsilon)^2 - 1}\right]$$
$$= \frac{1}{2\pi}\lim_{\varepsilon\to 0}\left[\frac{e^{-i\pi}e^{i\pi\varepsilon} - e^{i\pi}e^{-i\pi\varepsilon}}{\cancel{1} - 2\varepsilon + \varepsilon^2 \cancel{-1}}\right]$$
$$= \frac{1}{2\pi}\lim_{\varepsilon\to 0}\left[\frac{-e^{i\pi\varepsilon} + e^{-i\pi\varepsilon}}{-2\varepsilon + O(\varepsilon^2)}\right]$$
$$= \frac{1}{2\pi}\lim_{\varepsilon\to 0}\left[\frac{\cancel{-1} - i\pi\varepsilon \cancel{+1} - i\pi\varepsilon}{-2\varepsilon + O(\varepsilon^2)}\right]$$
$$= -\frac{1}{2i}$$

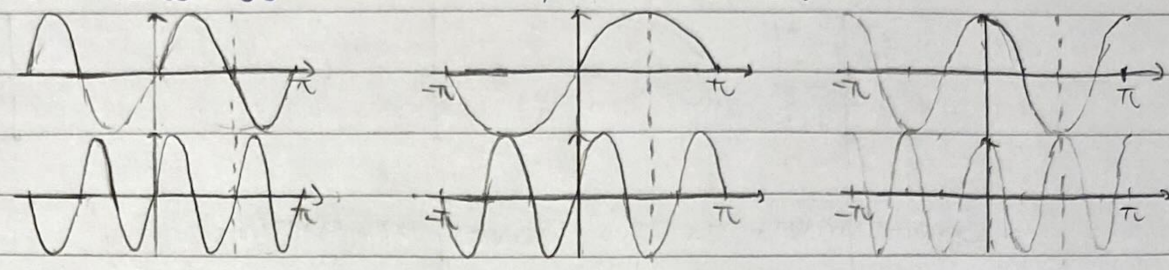
$$\therefore f(x) = \frac{1}{2i}e^{it} - \frac{1}{2i}e^{-it}$$

- e.g.: The 3 Fourier series.

(i) $\sum_{n=1}^{\infty} a_n \sin(2nx)$ (ii) $\sum_{n=1}^{\infty} b_n \sin((2n-1)x)$ (iii) $\sum_{n=1}^{\infty} c_n \cos(2nx)$

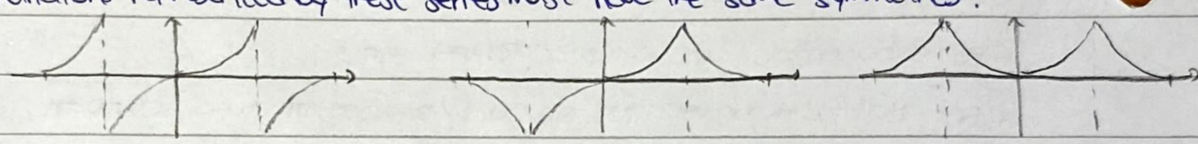
all represent the function $f(x) = x^2$ in the range $0 \le x \le \frac{\pi}{2}$.

Which series converges much more slowly than the others?



Abt $x=0$	antisymmetric	antisymmetric	symmetric
Abt $x=\frac{\pi}{2}$	antisymmetric	symmetric	symmetric.

Functions represented by these series must have the same symmetries.



Discontinuous $y \to \frac{1}{n}$ Discontinuous $\frac{dy}{dx} \to \frac{1}{n^2}$ Discontinuous $\frac{dy}{dx} \to \frac{1}{n^2}$

$\therefore$ (i) converges much slower than (ii), (iii).

- eg: Find the real Fourier series for a function $y(t) = \begin{cases} e^{-\alpha t} & t \in [0, T/2] \\ 0 & t \in (T/2, T] \end{cases}$

Exponentials → use complex Fourier series then convert back to real Fourier series

$$c_n = \frac{1}{T}\int_0^T y(t) e^{-\frac{i2\pi n t}{T}} dt$$

$$= \frac{1}{T}\int_0^{T/2} e^{-\alpha t} \cdot e^{-\frac{i2\pi n t}{T}} dt + \frac{1}{T}\int_{T/2}^{T} \cancel{0 \cdot e^{-\frac{i2\pi n t}{T}}}^{\,0} dt$$

$$= -\frac{1}{T(\alpha + i\frac{2\pi n}{T})}\left[e^{-(\alpha + \frac{i2\pi n}{T})t}\right]_0^{T/2}$$

$$= \frac{1}{\alpha T + 2\pi n i}\left(1 - e^{-\frac{\alpha T}{2} - \pi n i}\right)$$

$$d = c_0 = \frac{1}{\alpha T}\left(1 - e^{-\frac{\alpha T}{2}}\right)$$

$$a_n = 2\mathrm{Re}\{c_n\} = 2\mathrm{Re}\left\{\frac{\alpha T - 2\pi n i}{\alpha^2 T^2 + 4\pi^2 n^2}\left(1 - e^{-\frac{\alpha T}{2}} \cdot e^{-\pi n i}\right)\right\}$$

$$= 2\left[\alpha T\left(1 - e^{-\frac{\alpha T}{2}}\cos(-n\pi)\right) + 2\pi n e^{-\frac{\alpha T}{2}} \cancel{\sin(-n\pi)}^{\,0}\right] \cdot \frac{1}{\alpha^2 T^2 + 4\pi^2 n^2}$$

$$= \frac{2\alpha T\left(1 - e^{-\frac{\alpha T}{2}}\cos(n\pi)\right)}{\alpha^2 T^2 + 4\pi^2 n^2}$$

$$b_n = 2\mathrm{Im}\{c_n\} = 2\mathrm{Im}\left\{\frac{\alpha T - 2\pi n i}{\alpha^2 T^2 + 4\pi^2 n^2}\left(1 - e^{-\frac{\alpha T}{2}} e^{-\pi n i}\right)\right\}$$

$$= 2\left[-2\pi n\left(1 - e^{-\frac{\alpha T}{2}}\cos(-n\pi)\right) + \alpha T \cancel{\sin(-n\pi)}^{\,0}\right] \cdot \frac{1}{\alpha^2 T^2 + 4\pi^2 n^2}$$

$$= \frac{4n\pi\left(1 - e^{-\frac{\alpha T}{2}}\cos(n\pi)\right)}{\alpha^2 T^2 + 4\pi^2 n^2}$$

$$\therefore f(t) = \frac{1}{\alpha T}\left(1 - e^{-\frac{\alpha T}{2}}\right) + \sum_{n=1}^{\infty} \frac{2\alpha T\left(1 - e^{-\frac{\alpha T}{2}}\cos(n\pi)\right)}{\alpha^2 T^2 + 4\pi^2 n^2} \cos\left(\frac{2\pi n}{T}t\right) + \sum_{n=1}^{\infty} \frac{4n\pi\left(1 - e^{-\frac{\alpha T}{2}}\cos(n\pi)\right)}{\alpha^2 T^2 + 4\pi^2 n^2} \sin\left(\frac{2\pi n}{T}t\right)$$

Probability and statistics

refer to A level notes for more details

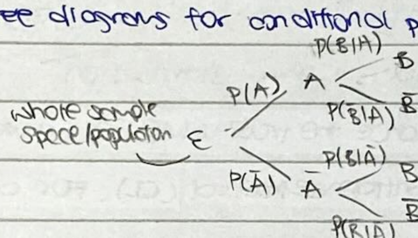
Probability

Probability

- For equiprobable events, $P(A) = \frac{n(A)}{n(\mathcal{E})}$ (event A; sample space $\mathcal{E}$).
- $P(A) = 1 \rightarrow$ certain event ; $P(A) = 0 \rightarrow$ impossible event. $0 \leq P(A) \leq 1$
- The complement of A is A' or $\bar{A}$. By defn $P(A) + P(A') = 1$.
- Addition of probability : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - ↳ If A and B are mutually exclusive, i.e. $P(A \cap B) = 0$, $\rightarrow P(A \cup B) = P(A) + P(B)$
- Multiplication of probability : $P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$
  - ↳ If A and B are independent, i.e. $P(A|B) = P(A|B')$ $\rightarrow P(A \cap B) = P(A) \cdot P(B)$.
- Law of total probability : $P(A) = \sum_{\forall n} P(A \cap B_n) = \sum_{\forall n} P(A|B_n) \cdot P(B_n)$

Conditional probability

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$ probability of A given B.
- Baye's theorem : $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$ or $P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$
- It is useful to use tree diagrams for conditional probability.



Combination and permutation

- $n!$ is the no. of ways of ordering n objects
- C^n_r is the no. of ways of choosing r objects out of n, order does not matter. $C^n_r = \frac{n!}{r!(n-r)!}$
- P^n_r is the no. of ways of choosing r objects out of n, order matters. $P^n_r = \frac{n!}{(n-r)!}$

* $P^n_r = C^n_r \times r!$; $C^n_r = C^n_{n-r}$; $C^n_r + C^n_{r-1} = C^{n+1}_r$

Statistics

Mean, variance and standard deviation

- For a population,
  - ↳ Arithmetic mean $M = \frac{\Sigma x}{n}$
  - ↳ variance $\sigma^2 = \frac{\Sigma(x-M)^2}{n} = \frac{\Sigma x^2}{n} - M^2$
  - ↳ standard deviation $\sigma = \sqrt{\frac{\Sigma(x-M)^2}{n}} = \sqrt{\frac{\Sigma x^2}{n} - M^2}$
- For a sample,
  - ↳ Arithmetic mean $m = \frac{\Sigma x}{n}$
  - ↳ variance $s^2 = \frac{\Sigma(x-m)^2}{n-1} = \frac{n}{n-1}\left(\frac{\Sigma x^2}{n} - m^2\right)$
  - ↳ standard deviation $s = \sqrt{\frac{\Sigma(x-m)^2}{n-1}} = \sqrt{\frac{n}{n-1}\left(\frac{\Sigma x^2}{n} - m^2\right)}$

Discrete probability distribution

- For a given probability mass function (pmf) $P(x)$,
  - ↳ Arithmetic mean $\mu = \sum xP(x)$
  - ↳ Variance $\sigma^2 = \sum (x-\mu)^2 P(x) = \sum x^2 P(x) - \mu^2$
  - ↳ standard deviation $\sigma = \sqrt{\sum (x-\mu)^2 P(x)} = \sqrt{\sum x^2 P(x) - \mu^2}$

Continuous probability distribution

- For a given probability density function (pdf) $f(x)$,
  - ↳ Probability that x is between a and b $P(a<x<b) = \int_a^b f(x)\,dx$.
  - ↳ Arithmetic mean $\mu = \int_{-\infty}^{\infty} xf(x)\,dx$
  - ↳ Variance $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$
  - ↳ standard deviation $\sigma = \sqrt{\int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx} = \sqrt{\int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2}$

Normal distribution (continuous probability distribution)

- The pdf of the Normal distribution is $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ $[N(\mu,\sigma^2)]$.
- The Normal distribution is symmetric about its mean.
- Central limit theorem (CLT) states if we add together a sufficient no. (say 30) iid. RVs, the sum follows a Normal distribution.

Sample mean.

- If we have n samples of x $(x_1, x_2 \ldots x_n)$, consider the sample mean $\bar{x} = \frac{\sum x}{n}$
  - ↳ $E(\bar{x}) = E(x) = \mu$
  - ↳ $Var(\bar{x}) = \frac{Var(x)}{n} = \frac{\sigma^2}{n}$.
  - ↳ $\sigma(\bar{x}) = \frac{\sigma(x)}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}}$
- The p% CI for the mean is an interval constructed from sample data in such a way that p% of such intervals will include the true population mean μ.
- The CI for the mean is given by: $[\bar{x} - k\frac{\sigma}{\sqrt{n}}, \bar{x} + k\frac{\sigma}{\sqrt{n}}]$
  - ↳ 50% CI, $k = 0.67$
  - ↳ 68% CI, $k = 1$
  - ↳ 95% CI, $k = 2$
  - ↳ 99.73% CI $k = 3$.

Functions of several variables

Representing functions of several variables

Graphical representations

- consider a function of 2 variables

$$z = f(x,y)$$

where x, y are the independent variables and z is the dependent variable

- we can represent such a function using
  - ↳ A surface in 3D
  - ↳ contour map
  - ↳ Heat map
  - ↳ Plot families of curves w/ x or y fixed.
- For functions of 3 variables $w=f(x,y,z)$, it is possible to use a sequence of pictures of one of the above types.

Parametric representations

- It is sometimes useful to represent a surface parametrically:

$$\underline{r}(u,v) = x(u,v)\underline{i} + y(u,v)\underline{j} + z(u,v)\underline{k}$$

where u, v are parameters and $\underline{i}, \underline{j}, \underline{k}$ are the unit coordinate vectors.

- Setting $u = u_1, u_2 \ldots$ constant and varying v (and vice versa) gives a set of curves in the surface

Partial derivatives

Partial derivatives

- consider a function of 2 variables $z=f(x,y)$.
- If we fix y, the partial derivative of f wrt x at a general pt (x,y) is defined as.

$$\boxed{\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(x,y) = \lim_{\delta x \to 0} \frac{f(x+\delta x, y) - f(x,y)}{\delta x}}$$

- If we fix x, the partial derivative of f wrt y at a general pt (x,y) is defined as

$$\boxed{\frac{\partial f}{\partial y} = \frac{\partial f}{\partial y}(x,y) = \lim_{\delta y \to 0} \frac{f(x, y+\delta y) - f(x,y)}{\delta y}}$$

- The idea of partial derivative extends to any no. of variables.

$f(x_1, x_2 \ldots x_n)$ $\quad$ $x_1, x_2 \ldots x_n$ are independent variables.

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial x_i}(a_1, a_2 \ldots a_n) = \lim_{\delta x_i \to 0} \frac{f(a_1, a_2 \ldots a_i + \delta x_i \ldots a_n) - f(a_1, a_2 \ldots a_n)}{\delta x_i}$$

* sometimes the notation $\left(\frac{\partial f}{\partial x}\right)_y$ is used to emphasise the partial derivative wrt x keeping y fixed.

Higher order partial derivatives

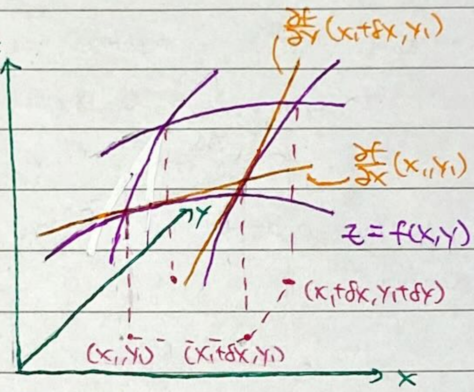
- For $z=f(x,y)$, the functions $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are also functions of x and y. If we differentiate again, we get
  - ↳ $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$ ↳ $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$
  - ↳ $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$ ↳ $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$

* Note the order for mixed partials $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$

- If $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ are continuous, they are equal. As this is a very mild condition, we always have equality in practice.
- $\frac{\partial^2 f}{\partial x \partial y}$ is the rate of change in the x-direction of the slope in the y-direction. It corresponds to the local "twist" or "torsion" in the surface

Total differential and chain rule

- consider $z=f(x,y)$



- To first order accuracy, (linearization)

Perturbation in x : $z(x_1+\delta x, y_1) = z(x_1,y_1) + \frac{\partial z}{\partial x}(x_1,y_1)\,\delta x$

Perturbation in y : $z(x_1+\delta x, y_1+\delta y) = z(x_1+\delta x, y_1) + \frac{\partial z}{\partial y}(x_1+\delta x, y_1)\,\delta y$

$$\simeq z(x_1+\delta x, y_1) + \frac{\partial z}{\partial y}(x_1,y_1)\,\delta y$$

$$\therefore \delta z = z(x_1+\delta x, y_1+\delta y) - z(x_1,y_1)$$

$$\boxed{\delta z = \frac{\partial z}{\partial x}(x_1,y_1)\,\delta x + \frac{\partial z}{\partial y}(x_1,y_1)\,\delta y} \quad [*]$$

- Taking limits, we have the defn of the total differential.

$$\boxed{dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy}$$

- Suppose x and y are both functions of t, i.e. $x=x(t)$ and $y=y(t)$. $\frac{dz}{dt}$ is given by dividing [*] by δt and taking limits. [Here $z=f(x(t),y(t))$]

$$\boxed{\frac{dz}{dt} = \left(\frac{\partial z}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial z}{\partial y}\right)_x \frac{dy}{dt}} \quad \text{[chain rule]}$$

- The chain rule extends to functions of more than 2 variables. Suppose $w=f(x(t),y(t),z(t))$.

$$\frac{dw}{dt} = \left(\frac{\partial w}{\partial x}\right)_{y,z}\frac{dx}{dt} + \left(\frac{\partial w}{\partial y}\right)_{x,z}\frac{dy}{dt} + \left(\frac{\partial w}{\partial z}\right)_{x,y}\frac{dz}{dt}$$

- The chain rule also applies if x, y, z are functions of more than 1 variable, i.e. $x(\theta,\phi)$ etc.

$$\left(\frac{\partial w}{\partial \theta}\right)_\phi = \left(\frac{\partial w}{\partial x}\right)_{y,z}\left(\frac{\partial x}{\partial \theta}\right)_\phi + \left(\frac{\partial w}{\partial y}\right)_{x,z}\left(\frac{\partial y}{\partial \theta}\right)_\phi + \left(\frac{\partial w}{\partial z}\right)_{x,y}\left(\frac{\partial z}{\partial \theta}\right)_\phi \;;\; \left(\frac{\partial w}{\partial \phi}\right)_\theta = \left(\frac{\partial w}{\partial x}\right)_{y,z}\left(\frac{\partial x}{\partial \phi}\right)_\theta + \left(\frac{\partial w}{\partial y}\right)_{x,z}\left(\frac{\partial y}{\partial \phi}\right)_\theta + \left(\frac{\partial w}{\partial z}\right)_{x,y}\left(\frac{\partial z}{\partial \phi}\right)_\theta$$

Implicit differentiation

- The eqn. $f(x,y)=k$ defines an implicit function of x. If we can solve to get $y=\phi(x)$ then we have an explicit function, and we can differentiate to find $\frac{dy}{dx}$. If not,

$$\frac{d}{dx}[f(x,y)] = \frac{d}{dx}[k]$$

$$\left(\frac{\partial f}{\partial x}\right)_y \cancel{\frac{dx}{dx}}^{1} + \left(\frac{\partial f}{\partial y}\right)_x \frac{dy}{dx} = 0$$

$$\boxed{\frac{dy}{dx} = \left[\left(\frac{\partial f}{\partial y}\right)_x\right]^{-1}\left(\frac{\partial f}{\partial x}\right)_y}$$

Perfect differentials (Exact differentials)

- Given a function $z=f(x,y)$, we have seen that

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

- If there exists a function f s.t. df can be expressed in the form

$$\boxed{df = P(x,y)dx + Q(x,y)dy} \quad [\#]$$

then $P(x,y)dx + Q(x,y)dy$ is a perfect differential.

- Comparing the 2 eqns above, we see that for a perfect differential,

$$P(x,y) = \left(\frac{\partial f}{\partial x}\right)_y \qquad Q(x,y) = \left(\frac{\partial f}{\partial y}\right)_x$$

$$\left(\frac{\partial P}{\partial y}\right)_x = \frac{\partial^2 f}{\partial y \partial x} \qquad \left(\frac{\partial Q}{\partial x}\right)_y = \frac{\partial^2 f}{\partial x \partial y}$$

Equality of the mixed second partial derivatives implies

$$\boxed{\left(\frac{\partial P}{\partial y}\right)_x = \left(\frac{\partial Q}{\partial x}\right)_y}$$

which is a sufficient condition for [#]. i.e. $\left(\frac{\partial P}{\partial y}\right)_x = \left(\frac{\partial Q}{\partial x}\right)_y \Leftrightarrow df = P(x,y)dx + Q(x,y)dy$

- The integral of a perfect differential along a path in the (x,y) plane only depends on the end pts.

- Given [#], we can find $f(x,y)$.

$$f = \int P(x,y)dx = g_1(x,y) + h_1(y) \qquad f = \int Q(x,y)dy = g_2(x,y) + h_2(x)$$

Equating both sides, $g_1(x,y) + h_1(y) = g_2(x,y) + h_2(x)$

Cancelling terms, $h_1(y) + \phi_1(y) = h_2(x) + \Phi_2(x)$

LHS independent of x; RHS independent of y → both equal a constant, c.

$$h_1(y) + \phi_1(y) = c \quad \rightarrow \quad h_1(y) = c - \phi_1(y).$$

- e.g: Given $df = (2x-2y)dx + (4-2x-6y)dy$, find $f(x,y)$.

$f = \int (2x-2y)dx = x^2 - 2xy + h_1(y)$; $f = \int (4-2x-6y)dy = 4y - 2xy - 3y^2 + h_2(x)$

$\therefore x^2 - 2xy + h_1(y) = 4y - 2xy - 3y^2 + h_2(x) \rightarrow h_1(y) - 4y + 3y^2 = h_2(x) - x^2$

$h_1(y) - 4y + 3y^2 = c \rightarrow h_1(y) = c + 4y - 3y^2 \quad \therefore f = x^2 - 2xy - 3y^2 + 4y + c.$

Directional derivatives

- The directional derivative $D_{\underline{b}}f$ is the rate of change of f when we move in the direction $\underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, where $\underline{b}$ is a unit vector. Moving in a small step $\underline{b}\,\delta s = \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$.

$$\boxed{D_{\underline{b}}f = D_{\underline{b}}f(x,y) = \lim_{\delta s \to 0} \frac{f(x+\delta x, y+\delta y) - f(x,y)}{\delta s}}$$

$\underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \frac{\delta x}{\delta s} \\ \frac{\delta y}{\delta s} \end{bmatrix}$

Recall that for $z = f(x,y)$, $\delta z = f(x+\delta x, y+\delta y) - f(x,y) = \left(\frac{\partial f}{\partial x}\right)_y \delta x + \left(\frac{\partial f}{\partial y}\right)_x \delta y$

$$D_{\underline{b}}f = \lim_{\delta s \to 0} \frac{1}{\delta s}\left[\left(\frac{\partial f}{\partial x}\right)_y \delta x + \left(\frac{\partial f}{\partial y}\right)_x \delta y\right]$$

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

$$\boxed{D_{\underline{b}}f = \left(\frac{\partial f}{\partial x}\right)_y b_1 + \left(\frac{\partial f}{\partial y}\right)_x b_2}$$

- Consider the tangent plane at (x_1, y_1) containing tangent lines w/ slopes $\left(\frac{\partial f}{\partial x}\right)_y(x_1,y_1)$, $\left(\frac{\partial f}{\partial y}\right)_x(x_1,y_1)$. $D_{\underline{b}}f$ is the slope of the line in the tangent plane lying over $\underline{b}$, i.e. the local slope of the surface in the $\underline{b}$ direction.

* Note $D_{\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]}f = \left(\frac{\partial f}{\partial x}\right)_y$ and $D_{\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]}f = \left(\frac{\partial f}{\partial y}\right)_x$.

The gradient of f(x,y)

- For a function of 2 variables $z = f(x,y)$, the gradient is defined as.

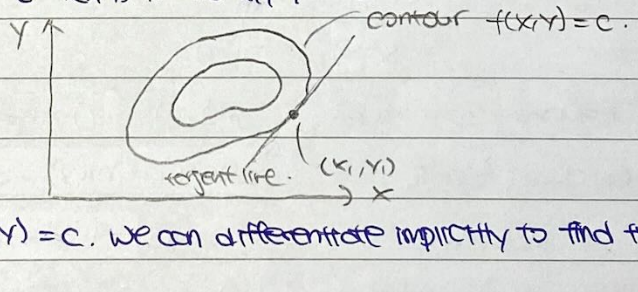
$$\boxed{\nabla f = \frac{\partial f}{\partial x}\underline{i} + \frac{\partial f}{\partial y}\underline{j} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}}$$

- We can rewrite the expression for the directional derivative as.

$$\boxed{D_{\underline{b}}f = \nabla f \cdot \underline{b} = |\nabla f|\cancel{|\underline{b}|}^{1}\cos\theta}$$

where θ is the angle between $\underline{b}$ and ∇f

- $D_{\underline{b}}f$ is max. when $\cos\theta = 1$, i.e. $\theta = 0$, when $\underline{b}$ is in the direction of ∇f. $\left(\underline{b} = \frac{\nabla f}{|\nabla f|}\right)$
- In other words, ∇f gives the direction in the xy-plane which corresponds to the largest directional derivative on the surface $z = f(x,y)$.
- Consider the contours of $z = f(x,y)$ in the x,y-plane.



Consider the contour $f(x,y) = c$. We can differentiate implicitly to find the slope of the tangent line.

$$f(x,y) = c \quad \rightarrow \quad \frac{dy}{dx} = -\left[\left(\frac{\partial f}{\partial y}\right)_x\right]^{-1}\left(\frac{\partial f}{\partial x}\right)_y$$

A vector // to the tangent line is $\begin{bmatrix} 1 \\ \frac{dy}{dx} \end{bmatrix}$

If we dot ∇f w/ this vector $\begin{bmatrix} 1 \\ \frac{dy}{dx} \end{bmatrix}$, we find that

$$\nabla f \cdot \begin{bmatrix} 1 \\ \frac{dy}{dx} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -\left[\frac{\partial f}{\partial y}\right]^{-1}\frac{\partial f}{\partial x} \end{bmatrix} = \frac{\partial f}{\partial x} - \frac{\partial f}{\partial x} = 0.$$

$\therefore \nabla f$ is $\perp$ to the contours.

The gradient of $f(x,y,z)$.

- For a function of 3 variables $w=f(x,y,z)$, the gradient is defined as

$$\boxed{\underline{\nabla} f = \frac{\partial f}{\partial x}\underline{i} + \frac{\partial f}{\partial y}\underline{j} + \frac{\partial f}{\partial z}\underline{k} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix}}$$

- A function $w=f(x,y,z)$ cannot be graphed directly as we would req. 4 dimensions. However, we could use contours – a family of surfaces w/ constant w.

$$f(x,y,z)=c \quad , \text{ where } c \text{ is a constant}$$

such surfaces are level surfaces.

- Within a level surface, $f(x,y,z)$ is constant, so

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = 0$$

$$\underline{\nabla} f \cdot \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = 0$$

where $\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}$ is a general vector in the level surface.

$\therefore \underline{\nabla} f$ is normal to the level surface.

Taylor series in 2 variables

Taylor series

- A sufficiently smooth (can be differentiated repeatedly) function $f(x,y)$ can be expanded in a Taylor series about a general pt (x,y).

$$\boxed{f(x+\Delta x, y+\Delta y) = f(x,y) + \Delta x \frac{\partial f}{\partial x} + \Delta y \frac{\partial f}{\partial y} + \frac{1}{2!}\left[\Delta x^2 \frac{\partial^2 f}{\partial x^2} + 2\Delta x \Delta y \frac{\partial^2 f}{\partial x \partial y} + \Delta y^2 \frac{\partial^2 f}{\partial y^2}\right] + \dots}$$

- Alternatively, we can write the Taylor series using more compact notation.

$$\boxed{f(x+\Delta x, y+\Delta y) = f(x,y) + \underline{r}^T \underline{\nabla} f + \frac{1}{2!}\underline{r}^T \underline{\underline{S}}\, \underline{r} + \dots}$$

where $\underline{r} = \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$ and $\underline{\underline{S}} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$

- $\underline{\underline{S}}$ is the Hessian matrix and is symmetric due to the equality of mixed partials.

- Note $\underline{r}^T \underline{\nabla} f = [\Delta x \ \Delta y] \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \Delta x \frac{\partial f}{\partial x} + \Delta y \frac{\partial f}{\partial y}$

$$\underline{r}^T \underline{\underline{S}}\, \underline{r} = [\Delta x \ \Delta y] \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \Delta x^2 \frac{\partial^2 f}{\partial x^2} + 2\Delta x \Delta y \frac{\partial^2 f}{\partial x \partial y} + \Delta y^2 \frac{\partial^2 f}{\partial y^2}$$

Maxima, minima and stationary points.

- The function $f(x,y)$ is said to have a maximum (minimum) at $x=x_1, y=y_1$ if we can draw a small circle around (x_1,y_1) s.t.

$$\overset{\text{maximum}}{f(x,y) < f(x_1,y_1)} \qquad \overset{\text{minimum}}{(f(x,y) > f(x_1,y_1))}$$

for all pts $(x,y) \neq (x_1,y_1)$ within the circle.

- In order for (x_1,y_1) to be a maximum/minimum, we need the directional derivative to be zero in all directions, i.e. the tangent plane is flat. Thus we need

$$\boxed{\underline{\nabla} f = \underline{0}} \quad \text{or} \quad \boxed{\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0} \quad \text{[necessary condition]}$$

- In general, any pt. satisfying the above eqn. is a stationary pt.

classification of stationary points

- Given that (x_1,y_1) is a stationary pt, i.e. $\underline{\nabla} f(x_1,y_1)=\underline{0}$, we can find the sufficient conditions for maxima, minima and saddle pts from the 2nd order terms in the Taylor series.
- Near a stationary pt, ($\underline{\nabla} f = \underline{0}$)

$$\Delta f = f(x_1+\Delta x, y_1+\Delta y) - f(x_1,y_1) = \underline{r}^T \cancel{\underline{\nabla} f}^{\,0} + \frac{1}{2!}\underline{r}^T \underline{\underline{S}}\,\underline{r} + \dots$$

$$\therefore \boxed{\Delta f \approx \frac{1}{2!}\underline{r}^T\underline{\underline{S}}\,\underline{r}}$$

Thus

$\underline{r}^T\underline{\underline{S}}\,\underline{r} < 0$ for all $\underline{r} \neq 0$	$\rightarrow$	(x_1,y_1) is a maximum
$\underline{r}^T\underline{\underline{S}}\,\underline{r} > 0$ for all $\underline{r} \neq 0$	$\rightarrow$	(x_1,y_1) is a minimum

- Since $\underline{\underline{S}}$ is symmetric, it has real eigenvalues λ_1, λ_2 and real orthogonal eigenvectors $\underline{u}_1, \underline{u}_2$.

If $\underline{\underline{\Lambda}} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, and $\underline{\underline{U}} = \begin{bmatrix} \uparrow & \uparrow \\ \underline{u}_1 & \underline{u}_2 \\ \downarrow & \downarrow \end{bmatrix}$, ($\underline{\underline{U}}$ is an orthogonal matrix), we can write

$$\underline{\underline{S}} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^T$$

$$\therefore \underline{r}^T\underline{\underline{S}}\,\underline{r} = \underline{r}^T\underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^T\underline{r}$$

We can define $\underline{v}$ s.t. $\underline{v} = \underline{\underline{U}}^T\underline{r}$ (so $\underline{v}^T = (\underline{\underline{U}}^T\underline{r})^T = \underline{r}^T\underline{\underline{U}}$), where $\underline{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

$$\therefore \underline{r}^T\underline{\underline{S}}\,\underline{r} = \underline{v}^T\underline{\underline{\Lambda}}\,\underline{v} = v_1^2\lambda_1 + v_2^2\lambda_2$$

* $\underline{v} = \underline{\underline{U}}^T\underline{r}$ represents the perturbation from x_1, y_1 in a new coordinate frame

- The choice between maximum, minimum and saddle is determined by the eigenvalues of $\underline{\underline{S}}$.

$\lambda_1 < 0,\ \lambda_2 < 0$	$\rightarrow$	$\underline{r}^T\underline{\underline{S}}\,\underline{r} < 0$	so maximum
$\lambda_1 > 0,\ \lambda_2 > 0$	$\rightarrow$	$\underline{r}^T\underline{\underline{S}}\,\underline{r} > 0$	so minimum
$\lambda_1\lambda_2 < 0$	$\rightarrow$	saddle pt.	

- If one or both eigenvalues are zero, the test is inconclusive $\rightarrow$ we need to look at higher order derivatives or use a sketch. usually the easiest method
- An alternative and equivalent test, directly on the second partial derivatives is as follows:

suppose $\underline{\nabla} f = \underline{0}$, define $\boxed{\Delta = |\underline{\underline{S}}| = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2}$

$\frac{\partial^2 f}{\partial x^2} < 0,\ \frac{\partial^2 f}{\partial y^2} < 0$ $[\Delta > 0]$	$\rightarrow$	maximum
$\frac{\partial^2 f}{\partial x^2} > 0,\ \frac{\partial^2 f}{\partial y^2} > 0$ $[\Delta > 0]$	$\rightarrow$	minimum
$\Delta < 0$	$\rightarrow$	saddle pt.

works as
$\lambda_1\lambda_2 = |\underline{\underline{S}}| = \Delta$
$\lambda_1 + \lambda_2 = \mathrm{Tr}(\underline{\underline{S}}) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$

- $\Delta = 0$ can be a maximum, minimum, saddle pt. or none of these.

Laplace Transforms

Definition of the Laplace Transform

- Let $g(t)$ be defined for all $t>0$ (the function may be zero or undefined for $t<0$). Usually, $g(t)=0$ for $t<0$.
- The Laplace Transform is defined as

$$\boxed{G(s) = L[g(t)] = \int_0^\infty g(t)e^{-st}dt}$$

- $G(s)$ is a function of the general complex variable s.
- The lower limit of 0 is important in problems w/ boundary conditions.
- * The Laplace Transform is an integral transform of the time signal $g(t)$

Existence of Laplace Transforms

- $g(t)$ is of exponential order if $|g(t)| \leq Me^{ct}$ for large t, w/ some M, C. i.e.

$$\lim_{t\to\infty} \frac{f(t)}{e^{ct}} \text{ exists}$$

- If $g(t)$ is continuous and of exponential order w/ constant c, then

$$G(s) = L[g(t)]$$

is defined for all $s>c$ (i.e. sufficiently far to the right in the complex plane)

Linearity of Laplace Transforms

- consider the Laplace Transform of the linear combination of $f(t)$ and $g(t)$.

$$L[\alpha f(t) + \beta g(t)] = \int_0^\infty [\alpha f(t) + \beta g(t)]e^{-st}dt$$

$$= \alpha \int_0^\infty f(t)e^{-st}dt + \beta \int_0^\infty g(t)e^{-st}dt$$

$$\boxed{L[\alpha f(t) + \beta g(t)] = \alpha L[f(t)] + \beta L[g(t)]}$$

i.e. The Laplace Transform is a linear operation.

Inverse of Laplace Transforms

- The function $g(t)$ is uniquely defined by its Laplace Transform $G(s)$ → we can think of Laplace Transform pairs, therefore, we can define the inverse Laplace Transform

$$\boxed{g(t) = L^{-1}[G(s)]}$$

- This is useful for solving DEs — we take the Laplace Transforms, find $G(s)$ and then convert it back to $g(t)$.

Common Laplace Transforms.

- consider the Laplace Transform of $g(t) = 1$ (or $g(t) = H(t)$)

$$L[g(t)] = \int_0^\infty 1 \cdot e^{-st} dt$$
$$= -\frac{1}{s}[e^{-st}]_0^\infty$$
$$\boxed{L[1] = \frac{1}{s}}$$

- consider the Laplace Transform of $g(t) = e^{at}$.

$$L[g(t)] = \int_0^\infty e^{at} \cdot e^{-st} dt$$
$$= -\frac{1}{s-a}[e^{-(s-a)t}]_0^\infty$$
$$\boxed{L[e^{at}] = \frac{1}{s-a}}$$

- consider the Laplace Transforms of $g_1(t) = \sin(\omega_0 t)$ and $g_2(t) = \cos(\omega_0 t)$

$$L[g_1(t)] = L\left[\frac{1}{2i}(e^{i\omega_0 t} - e^{-i\omega_0 t})\right]$$
$$= \frac{1}{2i}\left(\frac{1}{s-i\omega_0} - \frac{1}{s+i\omega_0}\right)$$
$$= \frac{1}{2i} \frac{2i\omega_0}{s^2+\omega_0^2}$$
$$\boxed{L[\sin(\omega_0 t)] = \frac{\omega_0}{s^2+\omega_0^2}}$$

$$L[g_2(t)] = L\left[\frac{1}{2}(e^{i\omega_0 t} + e^{-i\omega_0 t})\right]$$
$$= \frac{1}{2}\left(\frac{1}{s-i\omega_0} - \frac{1}{s+i\omega_0}\right)$$
$$= \frac{1}{2} \frac{2s}{s^2+\omega_0^2}$$
$$\boxed{L[\cos(\omega_0 t)] = \frac{s}{s^2+\omega_0^2}}$$

Derivatives

- Consider the Laplace Transform of $\frac{dg}{dt}$

$$L\left[\frac{dg}{dt}\right] = \int_0^\infty \frac{dg}{dt} e^{-st} dt$$
$$= [e^{-st} g(t)]_0^\infty + s\int_0^\infty g(t) e^{-st} dt$$
$$= -g(0) + sL[g(t)]$$
$$\boxed{L\left[\frac{dg}{dt}\right] = sG(s) - g(0)}$$

- Consider the Laplace Transform of $\frac{d^2g}{dt^2}$

$$L\left[\frac{d^2g}{dt^2}\right] = sL\left[\frac{dg}{dt}\right] - \dot{g}(0)$$
$$= s(sG(s) - g(0)) - \dot{g}(0)$$
$$\boxed{L\left[\frac{d^2g}{dt^2}\right] = s^2G(s) - sg(0) - \dot{g}(0)}$$

- In general, the Laplace Transform of $\frac{d^ng}{dt^n}$ is given by

$$\boxed{L\left[\frac{d^ng}{dt^n}\right] = s^nG(s) - s^{n-1}g(0) - s^{n-2}\dot{g}(0) - \dots - \frac{d^{n-1}g}{dt^{n-1}}(0)}$$

Multiplication by t:

- consider the derivative of $G(s)$ wrt. s.

$$\frac{d}{ds}G(s) = \frac{d}{ds}\int_0^\infty g(t) e^{-st} dt$$
$$= \int_0^\infty g(t) \frac{d}{ds}(e^{-st}) dt$$
$$= -\int_0^\infty t g(t) dt = -L[tg(t)]$$
$$\therefore \boxed{L[tg(t)] = -\frac{d}{ds}G(s)}$$

Steps and impulses

- Usually, we regard the integral $\int_0^{\varepsilon}\delta(t)dt$ as not well defined. To define the Laplace Transform of $\delta(t)$, we need a modified defn.

$$\boxed{G(s) = L[g(t)] = \int_{0^-}^{\infty} g(t)e^{-st}dt}$$

- Using this modified definition, we can find the Laplace Transform of $\delta(t)$.

$$\boxed{L[\delta(t)] = \int_{0^-}^{\infty}\delta(t)e^{-st}dt = e^{-s(0)} = 1}$$

- There is a slight change to the derivative rule using this modified definition.

$$\boxed{L[\tfrac{dg}{dt}] = sG(s) - g(0^-)}$$

- From this, we can find the Laplace Transform of $\delta(t)$ by

$$L[\delta(t)] = L[\tfrac{d}{dt}H(t)] = s\,L[H(t)] - H(0^-) = s\cdot 1/s - 0 = 1.$$

* To be safe, we should always stick to the 0^- convention.

(A 0^+ convention can also be used w/ initial conditions given at 0^+).

Translation property (shifting)

- Consider the Laplace Transform of $e^{at}g(t)$.

$$L[e^{at}g(t)] = \int_0^{\infty} g(t)e^{-(s-a)t}dt$$

$$\boxed{L[e^{at}g(t)] = G(s-a)}.$$

Time shift and the pulse

- Consider a function $g(t)$ which is zero for $t<0$. To emphasise that it is zero for $t<0$, we write $g(t)H(t)$. Shifting it to the right by T, we have $g(t-T)H(t-T)$.
- The Laplace Transform of this shifted function $g(t-T)H(t-T)$ is given by

$$L[g(t-T)H(t-T)] = \int_0^{\infty} g(t-T)H(t-T)e^{-st}dt$$

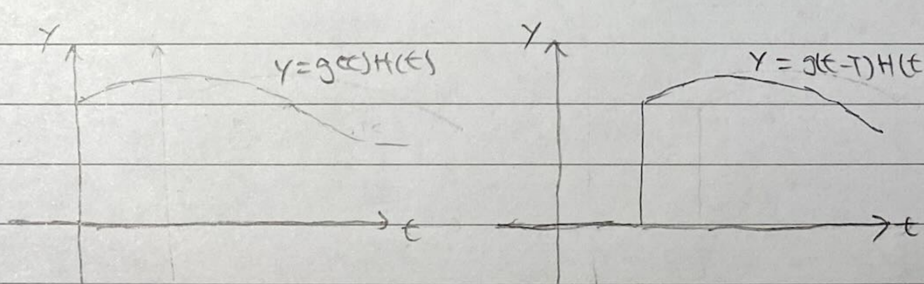
$$= \int_T^{\infty} g(t-T)e^{-st}dt$$

Using the substitution $u=t-T$, we have $du=dt$

$$L[g(t-T)H(t-T)] = \int_0^{\infty} g(u)e^{-s(u+T)}du$$

$$= e^{-sT}\int_0^{\infty} g(u)e^{-su}du$$

$$\boxed{L[g(t-T)H(t-T)] = e^{-sT}G(s)}$$



Convolution

- The Laplace Transform of a convolution equals the product of the Laplace Transforms.

$$\boxed{L\left[\int_0^t f(\tau)g(t-\tau)d\tau\right] = L[f(t)]\,L[g(t)]}$$

- It is slightly easier to prove the above w/ double-infinite limits in the integrals. As usual, we assume that $f(t)=0$ and $g(t)=0$ for $t<0$, so $g(t-\tau)=0$ for $\tau>t$. Thus,

$$\int_0^t \int_0^t f(\tau)g(t-\tau)d\tau\, e^{-st}dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau\, e^{-st}dt$$
$$= \int_{-\infty}^{\infty} e^{-s\tau}f(\tau)\int_{-\infty}^{\infty} e^{-s(t-\tau)}g(t-\tau)\,dt\,d\tau$$

Using the substitution $u=t-\tau$, we have $du=dt$

$$L\left[\int_0^t f(\tau)g(t-\tau)d\tau\right] = \int_{-\infty}^{\infty} e^{-s\tau}f(\tau)d\tau\int_{-\infty}^{\infty} e^{-su}g(u)du$$
$$= L[f(t)]\,L[g(t)]$$

Solving differential equations

- eg: solve $\ddot{y}+2\dot{y}+y=\cos 2t$, given $y(0)=0$, $\dot{y}(0)=1$.

Taking the Laplace Transform on $\ddot{y}+2\dot{y}+y=\cos 2t$,

$$s^2Y(s) - s\cancel{y(0)}^{0} - \dot{y}(0) + 2sY(s) - 2\cancel{y(0)}^{0} + Y(s) = \frac{s}{s^2+4}$$
$$Y(s)(s^2+2s+1) = \frac{s}{s^2+4} + 1$$
$$Y(s)(s+1)^2 = \frac{s^2+s+4}{s^2+4}$$
$$Y(s) = \frac{s^2+s+4}{(s^2+4)(s+1)^2}$$

* $\frac{s^2+s+4}{(s^2+4)(s+1)^2} = \frac{As+B}{s^2+4} + \frac{C}{s+1} + \frac{D}{(s+1)^2} \rightarrow s^2+s+4 = (As+B)(s+1)^2 + C(s^2+4)(s+1) + D(s^2+4)$

Equating coefficients, we find that $A=-3/25$, $B=8/25$, $C=3/25$, $D=4/5$.

$$\therefore\ Y(s) = -\frac{3}{25}\cdot\frac{s}{s^2+4} + \frac{8}{25}\cdot\frac{1}{2}\cdot\frac{2}{s^2+4} + \frac{3}{25}\frac{1}{s+1} + \frac{4}{5}\frac{1}{(s+1)^2}$$
$$y(t) = -\frac{3}{25}\cos 2t + \frac{4}{25}\sin 2t + \frac{3}{25}e^{-t} + \frac{4}{5}te^{-t}$$

Simultaneous differential equations

- Laplace Transforms are particularly useful w/ simultaneous DEs since it is often easier to take Laplace Transforms directly, than to manipulate algebraically to solve for the req. variable.