Algebra and Geometry

Prof. P. Haynes and Prof. J Saxl

Michaelmas 2002

Contents

1 Complex Numbers

1.1 Introduction

Real Numbers (denoted by $\mathbb{R}$) consist of:

integers	(denoted by $\mathbb{Z}$) $\ldots, -3, -2, -1, 0, 1, 2, \ldots$
rationals	(denoted by $\mathbb{Q}$) $\frac{p}{q}$ where p, q are integers
irrationals	$\sqrt{2}, \pi, e, \pi^2$ etc

It is often useful to visualise real numbers as lying on a line

Complex Numbers (denoted by $\mathbb{C}$):
If $a, b \in \mathbb{R}$ then $z = a + ib \in \mathbb{C}$ ('$\in$' means "belongs to") where i is such that $i^2 = -1$.
If $z = a + ib$ then write

$a = \Re(z)$ or $\mathrm{Re}(z)$	(real part of z)
$b = \Im(z)$ or $\mathrm{Im}(z)$	(imaginary part of z)

Extending the number system from real ($\mathbb{R}$) to complex ($\mathbb{C}$) allows certain important generalisations. For example, in complex numbers the quadratic equation

$$\alpha x^2 + \beta x + \gamma = 0 \quad : \quad \alpha, \beta, \gamma \in \mathbb{R}, \alpha \neq 0$$

always has two roots

$$x_1 = -\frac{\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha} \qquad x_2 = -\frac{\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

where

$x_1, x_2 \in \mathbb{R}$	if $\beta^2 > 4\alpha\gamma$
$x_1, x_2 \in \mathbb{C}$	if $\beta^2 < 4\alpha\gamma$

in which latter case

$$x_1 = -\frac{\beta}{2\alpha} + i\frac{\sqrt{4\alpha\gamma - \beta^2}}{2\alpha}, \qquad x_2 = -\frac{\beta}{2\alpha} - i\frac{\sqrt{4\alpha\gamma - \beta^2}}{2\alpha}$$

Note: $\mathbb{C}$ contains all real numbers i.e. if $a \in \mathbb{R}$ then $a + i.0 \in \mathbb{C}$
A complex number $0 + ib$ is said to be "pure imaginary"

Algebraic manipulation for complex number: Simply follow the rules for reals, adding the rule $i^2 = -1$
Hence:

addition/subtraction:	$(a + ib) \pm (c + id)$ $= (a \pm c) + i(b \pm d)$
multiplication:	$(a + ib)(c + id)$ $= ac + ibc + ida + (ib)(id)$ $= (ac - bd) + i(bc + ad)$
inverse:	$(a + ib)^{-1} = \frac{a}{a^2+b^2} - \frac{ib}{a^2+b^2}$

[Check from the above that $zz^{-1} = 1 + i.0$]
All of these operations on elements of $\mathbb{C}$ result in new elements of $\mathbb{C}$ (This is described as "closure": $\mathbb{C}$ is 'closed under addition' etc)
We may extend the idea of functions to complex numbers. The complex-valued function f takes any complex number (z) as 'input' and defines a new complex number $f(z)$ as 'output'.

Complex Conjugate of $z = a + ib$ is defined as $a - ib$, and written as $\overline{z}$ (sometimes z^*)
The complex conjugate has the properties $\overline{z_1 \pm z_2} = \overline{z_1} \pm \overline{z_2}$, $\overline{z_1 z_2} = \overline{z_1}\ \overline{z_2}$, $\overline{z^{-1}} = (\overline{z})^{-1}$

Modulus of $z = a + ib$ is defined as $\sqrt{a^2 + b^2}$ and written as $|z|$
Note that $|z|^2 = z\overline{z}$ and $z^{-1} = \frac{\overline{z}}{|z|^2}$

Theorem 1. The representation of a complex number z in terms of real and imaginary parts is unique

Proof. Assume $\exists a, b, c, d \in \mathbb{R}$ such that

$$z = a + ib = c + id$$

Then by arranging $a - c = i(d - b)$, so $\overbrace{(a-c)^2}^{\geq 0} = \overbrace{-(d-b)^2}^{\leq 0}$, so $(a-c) = 0 \Rightarrow a = c$ and $(d - b) = 0 \Rightarrow b = d$
If follows that if $z_1 = z_2$ with $z_1, z_2 \in \mathbb{C}$ then $\Re(z_1) = \Re(z_2)$ and $\Im(z_1) = \Im(z_2)$ □

Complex Conjugate Function - Given a complex-valued function f, the complex conjugate function $\overline{f}$ is defined by:

$$\overline{f}(\overline{z}) = \overline{f(z)}$$

For example, if $f(z) = pz^2 + qz + r$ with $p, q, r \in \mathbb{C}$ then $\overline{f}(\overline{z}) = \overline{f(z)} = \overline{p}\overline{z}^2 + \overline{q}\overline{z} + \overline{r}$. Hence $\overline{f}(z) = \overline{p}z^2 + \overline{q}z + \overline{r}$
This example generalises to any function defined by addition, subtraction, multiplication and inverse

1.2 The Argand Diagram

Consider the set of points in 2D referred to Cartesian axes

We can represent each $z = x + iy \in \mathbb{C}$ by the point (x, y).

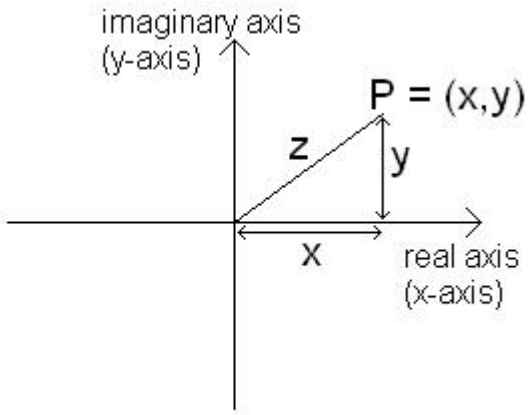

Figure 1: Representation on the Argand Diagram

Label the 2D vector $\vec{OP}$ by the complex number z. This defines the Argand diagram (or the 'complex plan') (see figure 1). [Invented by Caspar Wessel (1797) and re-invented by Jean Robert Argand (1806)]
Call the x-axis the 'real' axis and the y-axis the 'imaginary' axis.
Modulus: The modulus of z corresponds to the magniture of the vector $\vec{OP}$, $|z| = \sqrt{x^2 + y^2}$
Complex conjugate: If $\vec{OP}$ represents z, then $\vec{OP'}$ represents $\overline{z}$, where P' is the point $(x, -y)$ (i.e. P reflected in the x-axis)
Addition: If $z_1 = x_1 + iy_1$ associated with P_1, $z_2 = x_2 + iy_2$ associated with P_2, then $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$

$z_1 + z_2 = z_3$ is associated with the point P_3, obtained by completing the parallelogram $P_1OP_2P_3$ i.e. as vector addition $\vec{OP_3} = \vec{OP_1} + \vec{P_1P_3}$ (sometimes called the 'triangle law').

Theorem 2. If $z_1, z_2 \in \mathbb{C}$ then

i) $|z_1 + z_2| \leq |z_1| + |z_2|$
ii) $|z_1 - z_2| \geq |(|z_1| - |z_2|)|$

Proof. i) is the triangle inequality:

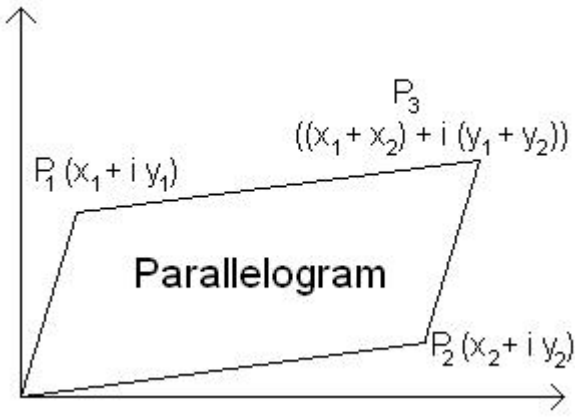

Figure 2: Addition on the Argand Diagram

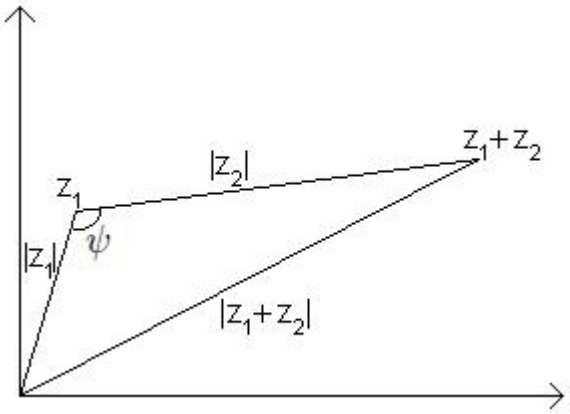

Figure 3: The triangle inequality

By the cosine rule

$$\begin{aligned}|z_1+z_2|^2 &= |z_1|^2+|z_2|^2-2|z_1||z_2|cos\psi \\ &\leq |z_1|^2+|z_2|^2+2|z_1||z_2| \\ &= (|z_1|+|z_2|)^2\end{aligned}$$

ii) follows from i), putting $z_1+z_2=z_1', z_2=z_2'$, so $z_1=z_1'-z_2'$. Hence by i) $|z_1'|\leq|z_1'-z_2'|+|z_2'|$ and $|z_1'-z_2'|\geq|z_i'|-|z_2'|$. Now, interchanging z_1' and z_2', we have $|z_2'|-|z_1'|\leq|z_2'-z_1'|=|z_1'-z_2'|$, hence the result □

Polar (modulus/argument) representation

Use plane polar co-ordinate to represent position in Argand diagram. $x=r\cos\theta$ and $y=r\sin\theta$, hence:

$$z=x+iy=r\cos\theta+ri\sin\theta=r(\cos\theta+i\sin\theta)\quad r\in\mathbb{R}, r\geq 0$$

Note that $|z|=\sqrt{x^2+y^2}=r$, so r is the modulus of z ('mod z)' for short). θ is called the 'argument' of z ('arg z) for short). The expression for z in terms of r and θ is called the 'modulus/argument form'.
The pair (r,θ) specifies z uniquely, but z does not specify (r,θ) uniquely, since adding $2n\pi$ to θ (n integer) does not change z. For each z there is a unique value of the argument θ such that $-\pi<\theta\leq\pi$, sometimes called the principal value of the argument.

Geometric interpretation of multiplication Consider z_1, z_2, written in modulus argument form

$$\begin{aligned}z_1 &= r_1(\cos\theta_1+i\sin\theta_1) \\ z_2 &= r_2(\cos\theta_2+i\sin\theta_2)\end{aligned}$$

$$\begin{aligned}z_1z_2 &= r_1r_2(\cos\theta_1.\cos\theta_2-\sin\theta_1.\sin\theta_2 \\ &\quad +i(\sin\theta_1.\cos\theta_2+\sin\theta_2.\cos\theta_1)) \\ &= r_1r_2(\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2))\end{aligned}$$

Multiplication of z_2 by z_1 rotates z_2 by θ_1 and scales z_2 by $|z_1|$

$|z_1z_2|=|z_1||z_2|$
$\arg(z_1z_2)=\arg(z_1)+\arg(z_2)$ ($+2k\pi$, with k an arbitrary integer)

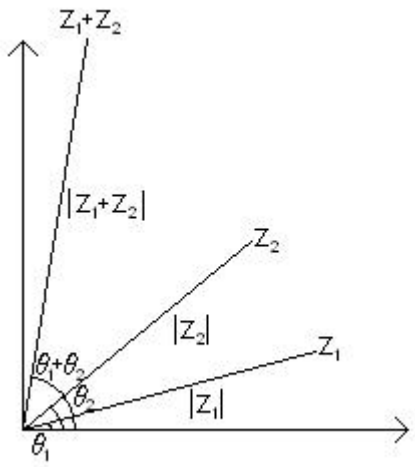

Figure 4: Multiplication on the Argand Diagram

1.3 De Moivre's Theorem: complex exponentials

Theorem 3. $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ where $\theta \in \mathbb{R}$ and $n \in \mathbb{Z}$

Proof. For $n > 0$ prove by induction.

Assume true for $n = p$: $(\cos\theta + i\sin\theta)^p = \cos p\theta + i\sin p\theta$
Then

$$\begin{aligned}(\cos\theta + i\sin\theta)^{p+1} &= (\cos\theta + i\sin\theta)(\cos\theta + i\sin\theta)^p \\ &= (\cos\theta + i\sin\theta)(\cos p\theta + i\sin p\theta) \\ &= \cos\theta.\cos p\theta - \sin\theta.\sin p\theta + i(\sin\theta.\cos p\theta + \cos\theta.\sin p\theta) \\ &= \cos(p+1)\theta + i\sin(p+1)\theta\end{aligned}$$

Hence true for $n = p + 1$
Trivially true for $n = 0, 1$, hence true $\forall n \in \mathbb{N} \cup \{0\}$ by induction

Now consider $n < 0$, say $n = -p$

$$\begin{aligned}(\cos\theta + i\sin\theta)^{-p} &= ((\cos\theta + i\sin\theta)^p)^{-1} \\ &= (\cos p\theta + i\sin p\theta)^{-1} \\ &= \frac{1}{\cos p\theta + i\sin p\theta} \\ &= \frac{\cos p\theta + i\sin p\theta}{(\cos p\theta + i\sin p\theta)^2} \\ &= \frac{\cos p\theta + i\sin p\theta}{1} \\ &= \cos n\theta + i\sin n\theta\end{aligned}$$

Hence true $\forall n \in \mathbb{Z}$ □

Exponential Function: $\exp x = e^x$
Defined by the power series

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

(This series converges $\forall x \in \mathbb{R}$ - see Analysis course)

It follows from the series that $(\exp x)(\exp y) = \exp(x + y)$ for $x, y \in \mathbb{R}$. This, plus $\exp(1) = 1 + 1 + \frac{1}{2} + \cdots$ may be used to justify the equivalence $\exp(x) = e^x$

The Complex Exponential defined by $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad z \in \mathbb{C}$ converges for all finite $|z|$
For short, we write $\exp(z) = e^z$ as above

Theorem 4.

$$\exp(iw) = e^{iw} = \cos w + i\sin w \quad w \in \mathbb{C}$$

Proof. First consider w real:

$$\begin{aligned}\exp(iw) &= \sum_{n=0}^{\infty} \frac{(iw)^n}{n!} = 1 + iw - \frac{w^2}{2} - \frac{iw^3}{3!} + \cdots \\ &= (1 - \frac{w^2}{2!} + \frac{w^4}{4!} - \cdots) + i(w - \frac{w^3}{3!} + \frac{w^5}{5!} - \cdots) \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{w^{2n}}{(2n)!} + i\sum_{n=0}^{\infty} (-1)^n \frac{w^{2n+1}}{(2n+1)!} \\ &= \cos w + i\sin w\end{aligned}$$

Now define the complex functions

$$\cos w = \sum_{n=0}^{\infty} (-1)^n \frac{w^{2n}}{(2n)!} \quad \sin w = \sum_{n=0}^{\infty} (-1)^n \frac{w^{2n+1}}{(2n+1)!}$$

For $w \in \mathbb{C}$
Then $\exp iw = e^{iw} = \cos w + i \sin w, w \in \mathbb{C}$
Similarly $\exp(-iw) = e^{iw} = \cos w - i \sin w$ □

It follows that $\cos w = \frac{1}{2}(e^{iw} + e{-iw})$ and $\sin w = \frac{1}{2i}(e^{iw} - e^{-iw})$

Relation to modulus/argument form
Put $w = \theta, \theta \in \mathbb{R}$, then $e^{i\theta} = \cos\theta + i\sin\theta$
Hence $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$, with (again) $r = |z|$, θ =arg(z).

Note that de Moivre's theorem

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

may be argued to follow from $e^{in\theta} = (e^{i\theta})^n$

Mulitiplication of two complex numbers:

$$z_1 z_2 = (r_1 e^{i\theta_1})(r_2 e^{i\theta_2}) = \overbrace{r_1 r_2}^{\text{multiply magnitudes}} \quad e^{i(\overbrace{\theta_1 + \theta_2}^{\text{add arguments}})}$$

Modulus/argument expression for 1
Consider solutions of $e^{i\theta} = 1$, hence $\cos\theta + i\sin\theta = 1, \cos\theta = 1, \sin\theta = 0$, hence $\theta = 2k\pi$, with $k \in \mathbb{Z}$ i.e.

$$e^{2k\pi i} = 1$$

Roots of Unity
A root of unity is a solution of $z^n = 1$ with $z \in \mathbb{C}$ and n a positive integer

Theorem 5. There are n solutions of $z^n = 1$ (i.e. n n^{th} roots of unity)

Proof. One solution is $z = 1$
Seek more general solutions of the form $re^{i\theta}, (re^{i\theta})^n = r^n e^{ni\theta} = 1$, hence $r = 1$, $e^{i\theta} = 1$, hence $n\theta = 2k\pi, k \in \mathbb{Z}$ with $0 \le \theta < 2\pi$
$\theta = \frac{2k\pi}{n}$ gives n distinct roots for $k = 0, 1, \ldots, n-1$ with $0 \le \theta < 2\pi$
Write $w = e^{\frac{2\pi i}{n}}$, then the roots of $z^n = 1$ are $1, w, w^2, \ldots, w^{n-1}$
□

Note: $w^n = 1$, also $\sum_{k=0}^{n-1} w^k = 1 + w + \cdots + w^{n-1} = 0$ because $\sum_{k=0}^{n-1} w^k = \frac{(w^n-1)}{(w-1)} = \frac{0}{w-1} = 0$

Example: $z^5 = 1$
Put $z = e^{i\theta}$, hence $e^{5i\theta} = e^{2\pi ki}$, hence $\theta = \frac{2\pi k}{5}, k = 0, 1, 2, 3, 4$ and $w = e^{\frac{2\pi i}{5}}$
Roots are $1, w, w^2, w^3, w^4$, with $1 + w + w^2 + w^3 + w^4 = 0$ (each root corresponds to a point of a pentagon)

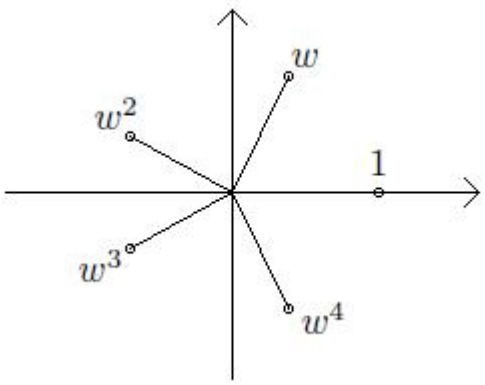

Figure 5: 5th roots of unity

1.4 Logarithms and complex powers

If $v \in \mathbb{R}, v > 0$ the complex equation $e^u = v$ has a unique real solution $u = \log v = \ln v$

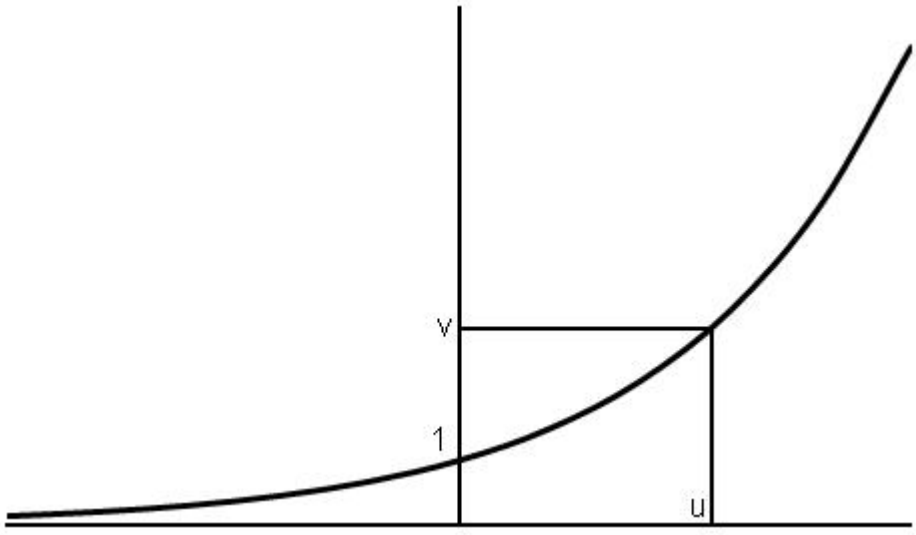

Figure 6: Graph of exp (x)

$\log(z)$ for $z \in \mathbb{C}$ is the solution w of $e^w = z$

Set $w = u + iv, u, v \in \mathbb{R}$, then $e^{u+iv} = z = re^{i\theta}$, hence

$$\begin{aligned} e^u &= |z| = r \\ v &= \arg(z) = \theta + 2k\pi \text{ for any } k \in \mathbb{R} \end{aligned}$$

Thus $w = \log z = \log \underbrace{|z|}_{\text{real}} + i \underbrace{\arg z}_{\text{imaginary}}$ with $\arg(z)$, and hence $\log(z)$ a multivalued function.

The principal value of $\log z$ is such that

$$-\pi < \arg z = \Im(\log z) \leq \pi$$

Example: If $z = -x, x \in \mathbb{R}, x > 0$ then $\log z = \log|-x| + i\arg(-x) = log|x| + i\pi + 2ki\pi$ for $k \in \mathbb{Z}$. The principal value of $\log(-x)$ is $\log|x| + i\pi$

Powers
Recall the definition of x^a for $x, a \in \mathbb{R}, x > 0$

$$x^a = e^{a \log x} = \exp(a \log x)$$

For complex numbers we use a similiar definition, given $z \in \mathbb{C}, z \neq 0$

$$z^w = e^w \log z$$

Note that since $\log z$ is multivalued so is z^w (arbitrary multiples of $e^{2\pi ikw}$ for $k \in \mathbb{Z}$)

Example: $i^i = e^{i \log i} = e^{i(\log|i| + i\arg(i))} = e^{i(\log 1 + 2ki\pi + \frac{i\pi}{2})} = e^{\frac{-\pi}{2}} \times e^{-2k\pi}$ for $k \in \mathbb{Z}$

1.5 Lines and circles in the complex plan

Line: For fixed z_0 and $c \in \mathbb{C}, z = z_0 + \lambda c, \lambda \in \mathbb{R}$ represents points on a straight line through z_0 and parallel to c

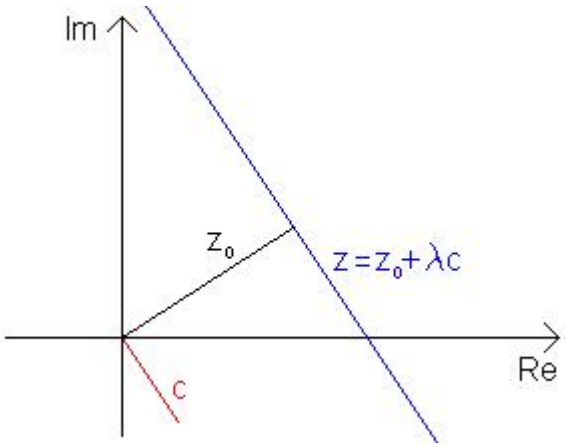

Figure 7: Lines on the complex plane

Note that $\lambda = \frac{(z-z_0)}{c} \in \mathbb{R}$, hence $\lambda = \overline{\lambda}$, so

$$\frac{(z-z_0)}{c} - \frac{(\overline{z}-\overline{z_0})}{\overline{c}}$$

Hence

$$z\overline{c} - \overline{z}c = z_0\overline{c} - \overline{z_0}c$$

is an alternative representation of the line.

Circle: circle radius r, centre $a \quad r \in \mathbb{R}, a \in \mathbb{C}$ is given by

$$S = \{z \in \mathbb{C} : |z-a| = r\}$$

the set of complex numbers z such that $|z-a| = r$

The $a = p + iq, z = x + iy$ then $|z-a|^2 = (x-p)^2 + (y-q)^2 = r^2$ i.e. the expression for a circle with centre (p, q), radius r in Cartesian coordinates.

An alternative description of the circle comes from $|z-a|^2 = (\overline{z}-\overline{a})(z-a)$, so

$$z\overline{z} - \overline{a}z - a\overline{z} + |a|^2 - r^2 = 0$$

1.6 Möbius Transformations

Consider a 'map' of $\mathbb{C} \to \mathbb{C}$ ('$\mathbb{C}$ into $\mathbb{C}$')

$$z \mapsto z' = f(z) = \frac{az+b}{cz+d}$$

where $a, b, c, d \in \mathbb{C}$ (all constant) and

i) c, d not both zero
ii) a, b not both zero
iii) a, c not both zero

i) ensures that $f(z)$ finite for some z, ii) and iii) ensure that different z map into different points. We can combine all these conditions into $ad - bc \neq 0$
$f(z)$ maps every point of the complex plane, except $z = -\frac{d}{c}$, into another.

Inverse: $z = \frac{(-dz'+b)}{(cz'-a)}$, which represents another Möbius transformation.
For every z' except $\frac{a}{c}$ there is a corresponding z, thus f maps $\mathbb{C}\backslash\{-\frac{d}{c}\}$ to $\mathbb{C}\backslash\{\frac{a}{c}\}$

Composition: Consider a second Möbius transformation

$$z' \mapsto z'' = g(z') = \frac{\alpha z' + \beta}{\gamma z' + \delta} \quad \alpha, \beta, \gamma, \delta \in \mathbb{C}, \alpha\delta - \beta\gamma \neq 0$$

Then the combined map $z \mapsto z''$ is also a Möbius transformation

$$\begin{aligned} z'' = g(z') &= g(f(z)) \\ &= \frac{\alpha z' + \beta}{\gamma z' + \delta} \\ &= \frac{\alpha(az+b) + \beta(cz+d)}{\gamma(az+b) + \delta(cz+d)} \\ &= \frac{(\alpha a + \beta c)z + \alpha b + \beta d}{(\gamma a + \delta c)z + \gamma b + \delta d} \end{aligned}$$

The set of all Möbius maps is there closed under composition.

Examples:
i) $(a = 1, c = 0, d = 1), z' = z + b$ is translation. Lines map to parallel lines. Circles maps to similar circles (same radius, new centre)

ii) $(b = 0, c = 0, d = 1), z' = az$, scales z by $|a|$ and rotates by $\arg a$ about O.
Line: $z = z_0 + \lambda p \quad (\lambda \in \mathbb{R})$ becomes $z' = az_0 + \lambda ap = z'_0 + \lambda c'$, another line.
Circle: $|z - q| = r$ becomes $|\frac{z'}{a} - q| = r$, hence $|z' - aq| = |a|r$, equivalently $|z' - q'| = r'$, another circle.

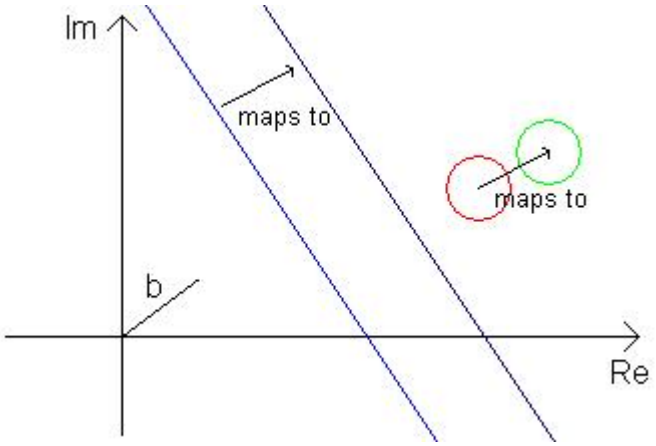

Figure 8: Translations of lines and circles

iii) $(a = 0, b = 1, c = 1, d = 0), z' = \frac{1}{z}$, described as 'inversion' with respect to O.
Line $z = z_0 + \lambda p$ or $z\overline{p} - \overline{z}p = z_0\overline{p} - \overline{z_0}p$ becomes

$$\frac{\overline{p}}{z'} - \frac{p}{\overline{z'}} = z_0\overline{p} - \overline{z_0}p$$

Hence

$$\overline{z'}\overline{p} - z'p = (z_0\overline{p} - \overline{z_0}pz'\overline{z'}$$

$$z'\overline{z'} - \frac{\overline{z'}\overline{p}}{z_0\overline{p} - \overline{z_0}p} - \frac{z'p}{\overline{z_0}p - z_0\overline{p}} = 0$$

$$|z' - \frac{\overline{p}}{z_0\overline{p} - \overline{z_0}p}|^2 = |\frac{p}{\overline{z_0}p - z_0\overline{p}}|^2$$

This is a circle through the origin, except when $\overline{z_0}p - z_0\overline{p} = 0$ (which is the condition that the straight line passes through the origin). Then $\overline{z'}\overline{p} - z'p = 0$ i.e. a straight line through the origin.
Circle: $|z - q| = r$ becomes $|\frac{1}{z'} - q| = r$ i.e. $|1 - qz'| = r|z'|$, hence $(1 - qz')(1 - \overline{q}\overline{z'}) = r^2\overline{z'}z'$, hence $z'\overline{z'}(|q|^2 - r^2) - qz'\overline{q}\overline{z'} + 1 = 0$, hence

$$|z' - \frac{\overline{q}}{|q|^2 - r^2}|^2 = \frac{|q|^2}{(|q|^2 - r^2)^2} - \frac{1}{|q|^2 - r^2}$$
$$= \frac{r^2}{(|q|^2 - r^2)^2}$$

2 Vector Algebra

2.1 Scalars and Vectors in 3-D

Scalars are represented by a single real number $\lambda \in \mathbb{R}$ e.g. temperature, density

Vectors have magnitude and direction e.g. velocity, force
Vectors are written as follows: $\underline{v}$ or $\boldsymbol{v}$

Magnitude is written as $|\underline{\boldsymbol{v}}|$
Two vectors $\underline{\boldsymbol{u}}$ are equal $\underline{\boldsymbol{v}}$ are equal if $|\underline{\boldsymbol{u}}| = |\underline{\boldsymbol{v}}|, \underline{\boldsymbol{u}} \parallel \underline{\boldsymbol{v}}$ and $\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}}$ in the same sense (or direction)

Geometric representation:
Represent $\underline{\boldsymbol{v}}$ as a line segment $\vec{\boldsymbol{AB}}$ in 3-D, with length $|\underline{\boldsymbol{v}}|$ and same direction as $\underline{\boldsymbol{v}}$
e.g. every point $\boldsymbol{P}$ in 3-D space has position vector $\underline{\boldsymbol{x}} = \vec{\mathrm{OP}}$ with $|\underline{\boldsymbol{x}}| = |\vec{\boldsymbol{OP}}| = |\boldsymbol{r}|$ (the distance from the origin)
Every complex number is associated with a unique point in the 2-D plane, and hence with the position vector of that point.

Notation:	set of all position vectors in 3-D	$\mathbb{R}^{\mathbf{3}}$
	set of all position vectors in 2-D	$\mathbb{R}^{\mathbf{2}}$

Properties of Vectors
1) Addition: use the parallelogram law

$$\vec{\mathrm{OC}} = \vec{\mathrm{OA}} + \vec{\mathrm{OB}} = \vec{\mathrm{OA}} + \vec{\mathrm{AC}}$$

$$\therefore \vec{\mathrm{OB}} = \vec{\mathrm{AC}}$$

$$\underline{\boldsymbol{c}} = \underline{\boldsymbol{a}} + \underline{\boldsymbol{b}}$$

Properties of addition:

i) $\underline{\boldsymbol{a}} + \underline{\boldsymbol{b}} = \underline{\boldsymbol{b}} + \underline{\boldsymbol{a}}$ (commutative)
ii) $|\underline{\boldsymbol{a}} + \underline{\boldsymbol{b}}| \leq |\underline{\boldsymbol{a}}| + |\underline{\boldsymbol{b}}|$ (triangle law)
iii) if $|\underline{\boldsymbol{a}}| = \mathbf{0}$ then $\underline{\boldsymbol{a}} = \underline{\mathbf{0}}$, the zero or null vector
$\underline{\boldsymbol{a}} = \underline{\boldsymbol{a}} + \underline{\mathbf{0}}\ \forall \underline{\boldsymbol{a}}$
iv) Given $\underline{\boldsymbol{a}}$, then $\exists(-\underline{\boldsymbol{a}})$ such that $(-\underline{\boldsymbol{a}}) + \underline{\boldsymbol{a}} = \underline{\mathbf{0}} = \underline{\boldsymbol{a}} + (-\underline{\boldsymbol{a}})$
$(-\underline{\boldsymbol{a}})$ is parallet to $\underline{\boldsymbol{a}}$, same magnitude, opposite sense

2) Multiplication by a scalar:
If $\boldsymbol{\lambda} \in \mathbb{R}$, then $\boldsymbol{\lambda}\underline{\boldsymbol{a}}$ has magnitude $|\boldsymbol{\lambda}||\underline{\boldsymbol{a}}|$, and is parallel to $\underline{\boldsymbol{a}}$, with the same sense if $\boldsymbol{\lambda} \geq \mathbf{0}$ and the opposite if $\boldsymbol{\lambda} < \mathbf{0}$
If $\underline{\boldsymbol{a}} \neq \mathbf{0}$, then $\hat{\underline{\boldsymbol{a}}} = \frac{\underline{\boldsymbol{a}}}{|\underline{\boldsymbol{a}}|}$ is a unit vector since $|\hat{\underline{\boldsymbol{a}}}| = \mathbf{1}$

Important properties:

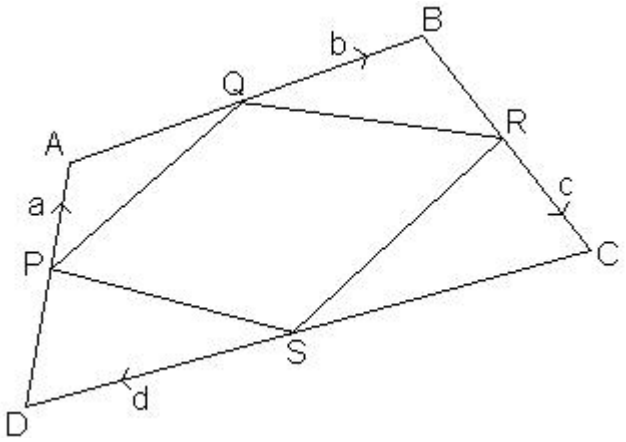

Figure 9: Forming a parallelogram from any quadrilateral

$$\begin{aligned}
(\boldsymbol{\lambda}+\boldsymbol{\mu})\underline{\boldsymbol{a}} &= \boldsymbol{\lambda}\underline{\boldsymbol{a}}+\boldsymbol{\mu}\underline{\boldsymbol{a}} \\
\boldsymbol{\lambda}(\underline{\boldsymbol{a}}+\underline{\boldsymbol{b}}) &= \boldsymbol{\lambda}\underline{\boldsymbol{a}}+\boldsymbol{\lambda}\underline{\boldsymbol{b}} \\
\boldsymbol{\lambda}(\boldsymbol{\mu}\underline{\boldsymbol{a}}) &= (\boldsymbol{\lambda}\boldsymbol{\mu})\underline{\boldsymbol{a}} \\
\underline{\boldsymbol{a}}+(\underline{\boldsymbol{b}}+\underline{\boldsymbol{c}}) &= (\underline{\boldsymbol{a}}+\underline{\boldsymbol{b}})+\underline{\boldsymbol{c}} \\
\underline{\mathbf{0}}.\underline{\boldsymbol{a}} &= \underline{\mathbf{0}} \\
\mathbf{1}.\underline{\boldsymbol{a}} &= \underline{\boldsymbol{a}} \\
(-\mathbf{1}).\underline{\boldsymbol{a}} &= (-\underline{\boldsymbol{a}})
\end{aligned}$$

Where the first 2 are distributive laws, the next 2 are associative laws.
If $\underline{\boldsymbol{c}} = \boldsymbol{\lambda}\underline{\boldsymbol{a}} + \boldsymbol{\mu}\underline{\boldsymbol{b}} \quad \boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}$ we say that $\underline{\boldsymbol{c}}$ is a linear combination of $\underline{\boldsymbol{a}}$ and $\underline{\boldsymbol{b}}$

Example: Using vector algebra to prove a geometric result
Proposition: The midpoints of the sides of any quadrilateral form a parallogram:

Let $\vec{\mathrm{DA}} = \underline{\boldsymbol{a}}, \vec{\mathrm{AB}} = \underline{\boldsymbol{b}}, \vec{\mathrm{BC}} = \underline{\boldsymbol{c}}, \vec{\mathrm{CD}} = \underline{\boldsymbol{d}}$
Since ABCD is closed, $\underline{\boldsymbol{a}} + \underline{\boldsymbol{b}} + \underline{\boldsymbol{c}} + \underline{\boldsymbol{d}} = \underline{\mathbf{0}}$

$$\begin{aligned}
\vec{\mathrm{PQ}} &= \vec{\mathrm{PA}} + \vec{\mathrm{AQ}} \\
&= \frac{\mathbf{1}}{\mathbf{2}}\vec{\mathrm{DA}} + \frac{\mathbf{1}}{\mathbf{2}}\vec{\mathrm{AB}} = \frac{\mathbf{1}}{\mathbf{2}}\underline{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{2}}\underline{\boldsymbol{b}} = \frac{\mathbf{1}}{\mathbf{2}}(\underline{\boldsymbol{a}}+\underline{\boldsymbol{b}}) \\
\vec{\mathrm{RS}} &= \frac{\mathbf{1}}{\mathbf{2}}\underline{\boldsymbol{c}} + \frac{\mathbf{1}}{\mathbf{2}}\underline{\boldsymbol{d}} = -\frac{\mathbf{1}}{\mathbf{2}}(\underline{\boldsymbol{a}}+\underline{\boldsymbol{b}}) = -\vec{\mathrm{PQ}} = \vec{\mathrm{QP}}
\end{aligned}$$

hence PQRS is a parallelogram

2.2 Scalar Product

Scalar Product - The scalar product of $\underline{a}, \underline{b}$ is $|\underline{a}||\underline{b}| \cos\theta = \underline{a}.\underline{b}$ with θ the angle between $\underline{a}$ and $\underline{b}$, $0 \leq \theta \leq \pi$
This is also called the dot product

Properties:

i) $\underline{a}.\underline{a} < 0$ if $\frac{\pi}{2} < \theta \leq \pi$
ii) $\underline{a}.\underline{a} = |\underline{a}|^2$
iii) $\underline{a}, \underline{b} \neq \underline{0}$ and $\underline{a}.\underline{b} = 0 \Rightarrow \underline{a}, \underline{b} \perp$ (or orthogonal)

Projection: Given $\underline{a}$, define $\underline{a'}$ as the projection of $\underline{a}$ onto the direction of $\underline{b}$ (that is, the part of $\underline{a}$ that is parallel to $\underline{b}$

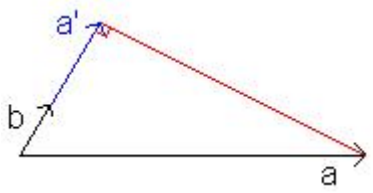

Figure 10: Projection of vector a onto vector b

By geometry

$$
\begin{aligned}
|\underline{a'}| &= |\underline{a}| \cos\theta \\
\Rightarrow \underline{a'} &= |\underline{a}| \cos\theta . \hat{\underline{b}} \\
\therefore \underline{a'} &= |\underline{a}||\underline{b}| \cos\theta \frac{\underline{b}}{|\underline{b}|^2} \\
\therefore \underline{a'} &= \frac{(\underline{a}.\underline{b})\underline{b}}{|\underline{b}|^2} = (\underline{a}.\hat{\underline{b}})\hat{\underline{b}}
\end{aligned}
$$

Properties of scalar product:

i) $\underline{a}.\underline{b} = \underline{b}.\underline{a}$ (commutative)
ii) $\underline{a}.\lambda\underline{b} = \lambda(\underline{a}.\underline{b}) = (\lambda\underline{a}).\underline{b}$

Proof. If $\lambda > 0$ then $\underline{a}.(\lambda\underline{b}) = |\underline{a}||\lambda\underline{b}| \cos\theta = \lambda|\underline{a}||\underline{b}| \cos\theta = \lambda(\underline{a}.\underline{b})$
If $\lambda < 0$ then $\underline{a}.(\lambda\underline{b}) = |\underline{a}||\lambda\underline{b}| \cos(\pi - \theta) = |\underline{a}||\underline{b}|(-\lambda) \cos(\pi - \theta) = \lambda|\underline{a}||\underline{b}| \cos\theta = \lambda(\underline{a}.\underline{b})$ □

iii) $\underline{a}.(\underline{b} + \underline{c}) = \underline{a}.\underline{b} + \underline{a}.\underline{c}$
trivial if $\underline{a} = \underline{0}$

If $\underline{\boldsymbol{a}} \neq \mathbf{0}$ define $\hat{\underline{\boldsymbol{a}}} = \frac{\underline{\boldsymbol{a}}}{|\underline{\boldsymbol{a}}|}$ as before

$(\hat{\underline{\boldsymbol{a}}}.(\underline{\boldsymbol{b}} + \underline{\boldsymbol{c}}))\hat{\underline{\boldsymbol{a}}}$ = projection of $\underline{\boldsymbol{b}} = \underline{\boldsymbol{c}}$ onto $\hat{\underline{\boldsymbol{a}}}$
$= (\hat{\underline{\boldsymbol{a}}}.\underline{\boldsymbol{a}})\hat{\underline{\boldsymbol{a}}} + (\hat{\underline{\boldsymbol{a}}}.\underline{\boldsymbol{c}})\hat{\underline{\boldsymbol{a}}}$
= projection of $\underline{\boldsymbol{b}}$ + projection of $\underline{\boldsymbol{c}}$ onto $\underline{\boldsymbol{a}}$

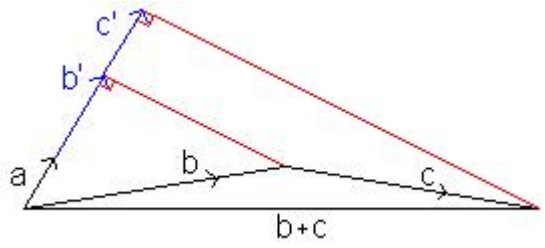

Figure 11: Projection of two vectors

$\Rightarrow \underline{\hat{a}}.(\underline{b}+\underline{c}) = \underline{\hat{a}}.\underline{b} + \underline{\hat{a}}.\underline{c}$
multiply through by $|\underline{a}|$
$\Rightarrow \underline{a}.(\underline{b}.\underline{c}) = \underline{a}.\underline{b} + \underline{a}.\underline{c}$

Example: Cosine rule

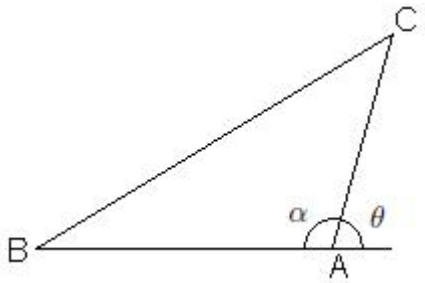

Figure 12: Cosine rule

$$\begin{aligned}
\mathrm{BC}^{\mathbf{2}} &= |\vec{\mathrm{BC}}|^{\mathbf{2}} \\
&= |\vec{\mathrm{BA}} + \vec{\mathrm{AC}}|^{\mathbf{2}} \\
&= (\vec{\mathrm{BA}} + \vec{\mathrm{AC}}).(\vec{\mathrm{BA}} + \vec{\mathrm{AC}}) \\
&= \vec{\mathrm{BA}}.\vec{\mathrm{BA}} + \mathbf{2}\vec{\mathrm{AC}}.\vec{\mathrm{BA}} + \vec{\mathrm{AC}}.\vec{\mathrm{AC}} \\
&= \mathrm{BA}^{\mathbf{2}} + \mathbf{2}\mathrm{AC}.\mathrm{BA}\,\mathbf{cos}\,\boldsymbol{\theta} + \mathrm{AC}^{\mathbf{2}} \\
&= \mathrm{BA}^{\mathbf{2}} + \mathrm{AC}^{\mathbf{2}} - \mathbf{2}\mathrm{AC}.\mathrm{BA}\,\mathbf{cos}\,\boldsymbol{\alpha}
\end{aligned}$$

2.3 Vector Product

Vector Product - The vector product is defined as a vector with the following properties:

i) $|\underline{a} \times \underline{b}| = |\underline{a}||\underline{b}| \sin\theta$ where θ is the angle between $\underline{a}$ and $\underline{b}$
ii) $\underline{a} \times \underline{b}$ is $\perp$ to $\underline{a}$ and $\underline{b}$
iii) It is in the same sense as a right-handed screw rotating from $\underline{a}$ to $\underline{b}$ (if $\underline{a} \longrightarrow \underline{b}$ looks anti-clockwise it is towards the viewer) - see figure 12

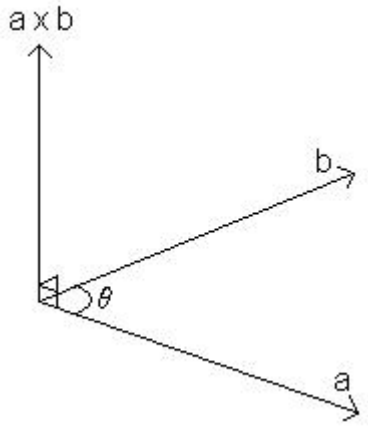

Figure 13: Right-handed rule for vector sense

Uses i) $\underline{a} \times \underline{a} = \underline{0}$
ii) $\underline{a} \times \underline{b} = -\underline{b} \times \underline{a}$
iii) $\underline{a} \times \underline{b} = \underline{0} \quad \underline{a}, \underline{b} \neq \underline{0} \Rightarrow \underline{a} \parallel \underline{b}$ i.e. $\theta = \mathbf{0}$ or π ($\exists \lambda \in \mathbb{R}$ such that $\underline{a} = \lambda \underline{b}$
iv) $\underline{a} \times (\lambda \underline{b}) = (\lambda \underline{a}) \times \underline{b} = \lambda(\underline{a} \times \underline{b})$
v) $\underline{a} \times (\underline{b} + \underline{c}) = (\underline{a} \times \underline{b}) + (\underline{a} \times \underline{c})$
The first 4 are obvious, the $\mathbf{5}^{\text{th}}$ needs proof

Proof. Consider $\frac{\underline{a}}{|\underline{a}|} \times \underline{b} = \underline{b}''$. This vector is the projection of $\underline{b}$ onto the plane perpendicular to $\underline{a}$, rotated by $\frac{\pi}{2}$ clockwise about $\underline{a}$. Consider this as two steps, first projection of $\underline{b}$ to give $\underline{b}'$, then rotation of $\underline{b}'$ to give $\underline{b}''$

Now note that if $\underline{x}'$ is the projection of the vector $\underline{x}$ onto the plane perpendicular to $\underline{a}$ then $\underline{b}' + \underline{c}' = (\underline{b} + \underline{c})'$

Rotating $\underline{b}', \underline{c}'$ and $(\underline{b} + \underline{c})'$ by $\frac{\pi}{2}$ gives the required result. □

Area Interpretation:

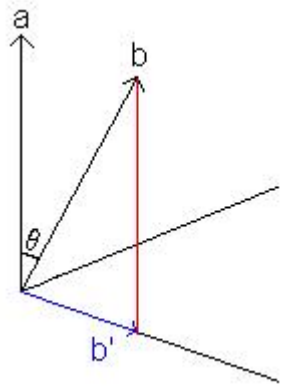

Figure 14: Projection of b to give b'

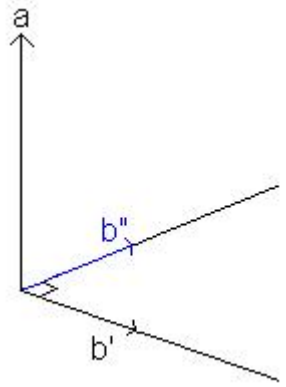

Figure 15: Rotation of b' to give b"

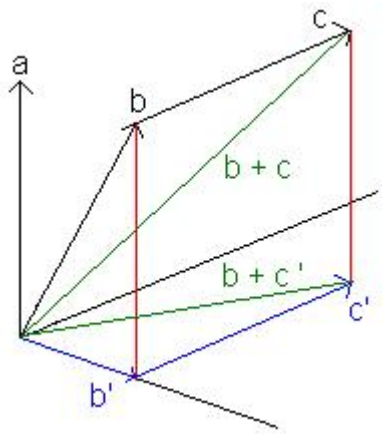

Figure 16: Projection of b and c onto a plane

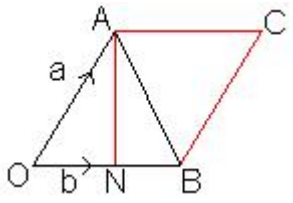

Figure 17: Parallelogram OABC

$$\begin{aligned}\text{Area OAB} &= \frac{\mathbf{1}}{\mathbf{2}}\text{OA.BN} \\ &= \frac{\mathbf{1}}{\mathbf{2}}\text{OA.OB}\,\mathbf{sin}\,\boldsymbol{\theta} \\ &= \frac{\mathbf{1}}{\mathbf{2}}|\underline{\boldsymbol{a}} \times \underline{\boldsymbol{b}}| \\ \text{Area OACB} &= |\underline{\boldsymbol{a}} \times \underline{\boldsymbol{b}}|\end{aligned}$$

$\underline{\boldsymbol{a}} \times \underline{\boldsymbol{b}}$ is in the direction normal to the plane containing the triangle, so can consider $\frac{\mathbf{1}}{\mathbf{2}}\underline{\boldsymbol{a}} \times \underline{\boldsymbol{b}}$ = vector area of triangle, as it contains information about the size and orientation

2.4 Scalar Triple Product

Given

$\cdot$ (scalar product)
$\times$ (vector product)

can form

i) $(\underline{a} \times \underline{b}).\underline{c}$ (scalar triple product)
ii) $(\underline{a} \times \underline{b}) \times \underline{c}$ (vector triple product)

Properties of i)

① $(\underline{a} \times \underline{b}).\underline{c} = \underline{c}.(\underline{a} \times \underline{b})$
② $(\underline{a} \times \underline{b}).\underline{c} = -(\underline{b} \times \underline{a}).\underline{c}$
③ $|(\underline{a} \times \underline{b}).\underline{c}|$ is the volume of a parellelipiped with edges $\underline{a}, \underline{b}$ and $\underline{c}$

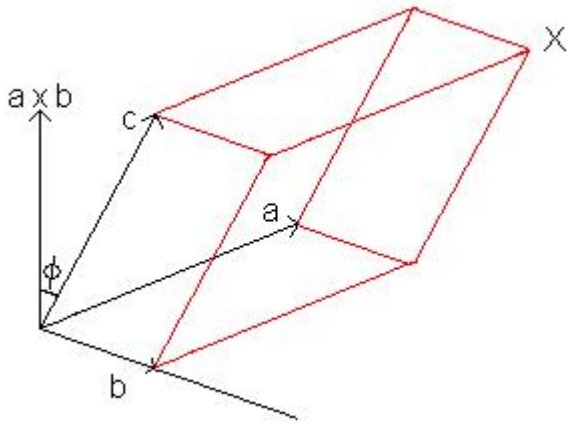

Figure 18: Parallelipiped formed by a, b and c

$$\begin{aligned} \text{Volume} &= \text{base area} \times \text{height} \\ &= |\underline{a} \times \underline{b}||\underline{c}| \cos\phi \\ &= |(\underline{a} \times \underline{b}).\underline{c}| \end{aligned}$$

Now $(\underline{a} \times \underline{b}).\underline{c} \geq \mathbf{0}$ if, when viewed from X (the furthest point from the origin) $\underline{a}, \underline{b}, \underline{c}$ are anti-clockwise

④ $(\underline{a} \times \underline{b}).\underline{c} = (\underline{b} \times \underline{c}).\underline{a}$
$\because$ if $\underline{a}, \underline{b}, \underline{c}$ is anticlockwise then so is $\underline{b}, \underline{c}, \underline{a}$ so the volume is unchanged
⑤ $(\underline{a} \times \underline{b}).\underline{c} = \underline{a}.(\underline{b} \times \underline{c})$ (from ② and ④)
⑥ If $\underline{a}, \underline{b}, \underline{c}$ are coplanar then $\underline{a}.(\underline{b} \times \underline{c}) = \mathbf{0}$

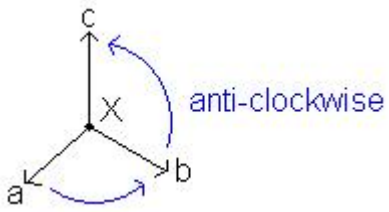

Figure 19: a, b, c viewed from X

Notation: $[\underline{\boldsymbol{a}}, \underline{\boldsymbol{b}}, \underline{\boldsymbol{c}}] = (\underline{\boldsymbol{a}} \times \underline{\boldsymbol{b}}).\underline{\boldsymbol{c}}$

2.5 Bases and Components

Consider a 2-D plane.
Given O, $\underline{a} \neq \underline{O}$ and $\underline{b} \neq \underline{O}$ with $\underline{a}, \underline{b}$ not $\parallel$ (i.e. $\underline{a} \times \underline{b} \neq \underline{0}$)
For all points P in the plane:

$$\vec{OP} = \underline{x} = \lambda\underline{a} + \mu\underline{b} \qquad (\lambda, \mu \in \mathbb{R})$$

for suitable λ, μ, which are unique given $\underline{x}$

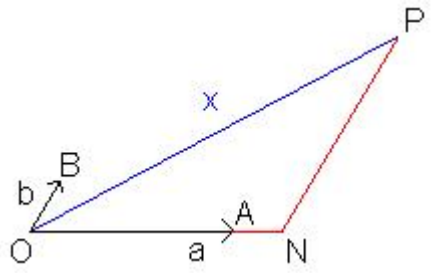

Figure 20: Construction of points in 2-D

Draw NP $\parallel$ OB. Let N be the intersection with the extended line OA.

$$\begin{aligned} \therefore \vec{OP} &= \vec{ON} + \vec{NP} \\ &= \lambda\underline{a} + \mu\underline{b} \end{aligned}$$

We say that $\{\underline{a}, \underline{b}\}$ spans the plane (or the set of vectors lying in the plane).

Lemma 6. λ, μ are unique

Proof. Assume $\exists\ \lambda, \lambda', \mu, \mu'$ such that $\underline{x} = \lambda\underline{a} + \mu\underline{b} = \lambda'\underline{a} + \mu'\underline{b}$
$\Rightarrow (\lambda - \lambda')\underline{a} = (\mu - \mu')\underline{b}$
$\Rightarrow \lambda - \lambda' = 0 = \mu - \mu' \qquad \because \underline{a}, \underline{b}$ not $\parallel$
$\Rightarrow$ uniqueness of λ, μ □

Linearly Independent - If $\alpha\underline{a} + \beta\underline{b} = \underline{0} \Rightarrow \alpha = 0, \beta = 0$ then $\{\underline{a}, \underline{b}\}$ is linearly independent

Basis - We say $\{\underline{a}, \underline{b}\}$ is a basis for the set of vectors in a plane if $\{\underline{a}, \underline{b}\}$ spans the plane and $\{\underline{a}, \underline{b}\}$ is linearly independent

Now consider 3-D
Given O, $\underline{a}, \underline{b}, \underline{c}$ with $\underline{a}, \underline{b}, \underline{c}$ not coplanar and $(\underline{a} \times \underline{b}).\underline{c} \neq 0$
For any point P in 3-D

$$\vec{OP} = \underline{x} = \lambda\underline{a} + \mu\underline{b} + \nu\underline{c} \qquad \lambda, \mu, \nu \in \mathbb{R}$$

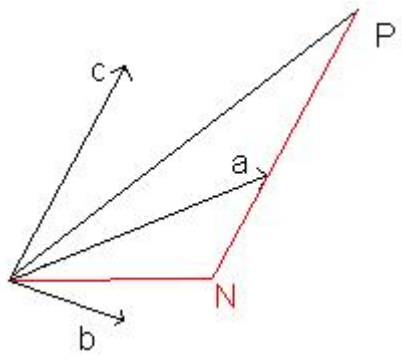

Figure 21: Construction of points in 3-D

Consider the line thru' P $\parallel$ to $\underline{c}$
This intersects $\mathbf{\Pi}_{\boldsymbol{a}}\boldsymbol{b}$ at N ($\because \underline{c}$ is not coplanar with $\underline{a}, \underline{b}$)

$$\begin{aligned}\therefore \vec{OP} &= \vec{ON} + \vec{NP} \\ &= \underbrace{\lambda\underline{a} + \mu\underline{b}}_{\because \text{ N is in } \mathbf{\Pi}_{ab}} + \underbrace{\nu\underline{c}}_{\vec{NP} \parallel \underline{c}}\end{aligned}$$

$\therefore \underline{a}, \underline{b}, \underline{c}$ span the 3-D plane

Lemma 7. $\boldsymbol{\lambda, \mu, \nu}$ are unique

Proof. Take

$$\underline{x}.(\underline{b} \times \underline{c}) = \lambda\underline{a}.(\underline{b} \times \underline{c}) + \underbrace{\mu\underline{b}.(\underline{b} \times \underline{c}) + \nu\underline{c}.(\underline{b} \times \underline{c})}_{\mathbf{0}}$$

$$\therefore \boldsymbol{\lambda} = \frac{\underline{x}.(\underline{b} \times \underline{c})}{\underline{a}.(\underline{b} \times \underline{c})}$$

Note that $\underline{a}.(\underline{b} \times \underline{c}) \neq \mathbf{0}$ as $\underline{a}, \underline{b}, \underline{c}$ are not coplanar

$$\therefore \boldsymbol{\lambda} = \frac{[\underline{x}, \underline{b}, \underline{c}]}{[\underline{a}, \underline{b}, \underline{c}]}$$

Which defines $\boldsymbol{\lambda}$ uniquely, and a similar argument will define $\boldsymbol{\mu, \nu}$ □

$\boldsymbol{\lambda, \mu, \nu}$ unique $\Rightarrow \{\underline{a}, \underline{b}, \underline{c}\}$ is a linearly independent set
hence $\{\underline{a}, \underline{b}, \underline{c}\}$ is a basis for the set of 3-D vectors

Dimension - The dimension of a space is the number of vectors needed for a basis

Components - If $\{\underline{a}, \underline{b}, \underline{c}\}$ is a basis for a 3-D set of vectors, then $\exists$ unique $\boldsymbol{\lambda, \mu, \nu}$ such that $\underline{x} = \lambda\underline{a} + \mu\underline{b} + \nu\underline{c}$
We call $\boldsymbol{\lambda, \mu, \nu}$ the components of $\underline{x}$ with respect to $\{\underline{a}, \underline{b}, \underline{c}\}$
Note that it would be different for a different basis

2.6 Standard Cartesian basis in 3-D

OX, OY and OZ are defined as the right-handed Cartesian axes:

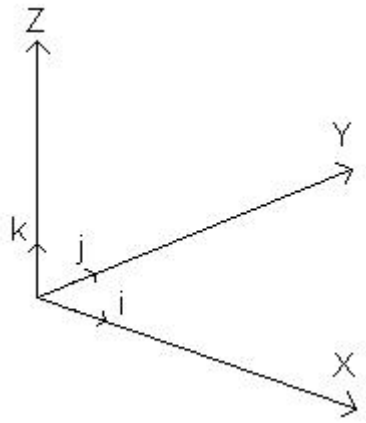

Figure 22: Standard Cartesian Axes

Viewed from the +ve quadrant, the axes go anti-clockwise $\Rightarrow$ right-handed. Let

$\underline{\boldsymbol{i}}$ be the unit vector along OX
$\underline{\boldsymbol{j}}$ be the unit vector along OY
$\underline{\boldsymbol{k}}$ be the unit vector along OZ

$\{\underline{\boldsymbol{i}}, \underline{\boldsymbol{j}}, \underline{\boldsymbol{k}}\}$ is a basis for 3-D spaces

Properties:

i) $\underline{\boldsymbol{i}}^2 = \underline{\boldsymbol{j}}^2 = \underline{\boldsymbol{k}}^2 = \mathbf{1}$
ii) $\underline{\boldsymbol{i}}.\underline{\boldsymbol{j}} = \underline{\boldsymbol{j}}.\underline{\boldsymbol{k}} = \underline{\boldsymbol{k}}.\underline{\boldsymbol{i}} = \mathbf{0}$
iii) $\underline{\boldsymbol{i}} \times \underline{\boldsymbol{j}} = \underline{\boldsymbol{k}} \quad \underline{\boldsymbol{j}} \times \underline{\boldsymbol{k}} = \underline{\boldsymbol{i}} \quad \underline{\boldsymbol{k}} \times \underline{\boldsymbol{i}} = \underline{\boldsymbol{j}}$
iv) $[\underline{\boldsymbol{i}}, \underline{\boldsymbol{j}}, \underline{\boldsymbol{k}}] = \mathbf{1}$

Orthonomal basis - A set of basis vectors satisfying properties i) and ii) above

We define the Cartesian Components of $\underline{\boldsymbol{v}}$ as $(\boldsymbol{v_x}, \boldsymbol{v_y}, \boldsymbol{v_z}) \quad \boldsymbol{v_i} \in \mathbb{R} \boldsymbol{i} = \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$, with

$$\underline{\boldsymbol{v}} = \boldsymbol{v_x}\underline{\boldsymbol{i}} + \boldsymbol{v_y}\underline{\boldsymbol{j}} + \boldsymbol{v_z}\underline{\boldsymbol{k}}$$

$\underline{\boldsymbol{v}}.\underline{\boldsymbol{i}} = \boldsymbol{v_x}\underline{\boldsymbol{i}}.\underline{\boldsymbol{i}} = \boldsymbol{v_x} \Rightarrow \underline{\boldsymbol{v}} = (\underline{\boldsymbol{v}}.\underline{\boldsymbol{i}})\underline{\boldsymbol{i}} + (\underline{\boldsymbol{v}}.\underline{\boldsymbol{j}})\underline{\boldsymbol{j}} + (\underline{\boldsymbol{v}}.\underline{\boldsymbol{k}}).\underline{\boldsymbol{k}}$
Often we write

$\underline{\boldsymbol{v}} = (\boldsymbol{v_x}, \boldsymbol{v_y}, \boldsymbol{v_z})$ (assuming all w.r.t. standard basis)
then $\underline{\boldsymbol{i}} = (\mathbf{1}, \mathbf{0}, \mathbf{0})$
$\underline{\boldsymbol{j}} = (\mathbf{0}, \mathbf{1}, \mathbf{0})$
$\underline{\boldsymbol{k}} = (\mathbf{0}, \mathbf{0}, \mathbf{1})$

Consider P with Cartesian coordinates $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$
Then $\vec{\mathrm{OP}} = \underline{\boldsymbol{x}} = \boldsymbol{x}\underline{\boldsymbol{i}} + \boldsymbol{y}\underline{\boldsymbol{j}} + \boldsymbol{z}\underline{\boldsymbol{k}}$
Or, for shorthand, $\underline{\boldsymbol{x}} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$
Direction Cosines

If $\underline{\boldsymbol{t}}$ is a unit vector $(\boldsymbol{t_x}, \boldsymbol{t_y}, \boldsymbol{t_z})$ with respect to the standard basis

$$\boldsymbol{t_x} = \underline{\boldsymbol{t}}.\underline{\boldsymbol{i}} = |\underline{\boldsymbol{t}}||\underline{\boldsymbol{i}}| \cos \boldsymbol{\alpha} = \cos \boldsymbol{\alpha}$$
$$\boldsymbol{t_y} = \ldots = \cos \boldsymbol{\beta}$$
$$\boldsymbol{t_z} = \ldots = \cos \boldsymbol{\gamma}$$
$$\therefore \underline{\boldsymbol{t}} = \underbrace{(\cos \boldsymbol{\alpha}, \cos \boldsymbol{\beta}, \cos \boldsymbol{\gamma})}_{\text{direction cosines}}$$

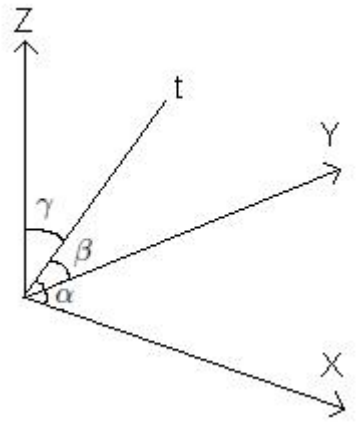

Figure 23: Direction Cosines

2.7 Vector identities (in terms of components)

Let

$$\underline{a} = a_x\underline{i} + a_y\underline{j} + a_z\underline{k}$$
$$\underline{b} = b_x\underline{i} + b_y\underline{j} + b_z\underline{k}$$
$$\underline{c} = c_x\underline{i} + c_y\underline{j} + c_z\underline{k}$$

i) Addition + Scalar Multiplication:

$$\lambda\underline{a} + \mu\underline{b} = (\lambda a_x + \mu b_x)\underline{i} + (\lambda a_y + \mu a_y)\underline{j} + (\lambda a_z + \mu b_z)\underline{k}$$

ii) Scalar Product

$$\underline{a}.\underline{b} = a_x b_x + a_y b_y + a_z b_z$$

iii) Vector Product

$$\underline{a} \times \underline{b} = (a_y b_z - a_z b_y)\underline{i} + (a_z b_x - a_x b_z)\underline{j} + (a_x b_y - a_y b_x)\underline{k}$$

iv) Scalar Triple Product

$$\underline{a}.(\underline{b} \times \underline{c}) = a_x(b_y c_z - b_z c_y) + a_y(b_z c_x - b_x c_z) + a_z(b_x c_y - b_y c_x)$$

v) Vector Triple Product

Theorem 8. $\underline{a} \times (\underline{b} \times \underline{c}) = (\underline{a}.\underline{c})\underline{b} - (\underline{a}.\underline{b})\underline{c}$

Proof. Consider the x component:

$$\begin{aligned} a_y(\underline{b} \times \underline{c})_z - a_z(\underline{b} \times \underline{c})_y &= a_y(b_x c_y - b_y c_x) - a_z(b_z c_x - b_x c_z) \\ &= (a_y c_y + a_z c_z)b_x - (a_y b_y + a_z b_z)c_x + a_x b_x c_x - a_x b_x c_x \\ &= (a_x c_x + a_y c_y + a_z c_z)b_x - (a_x b_x + a_y b_y + a_z b_z)c_x \\ &= (\underline{a}.\underline{c})b_x - (\underline{a}.\underline{b})c_x \\ &= [(\underline{a}.\underline{c})\underline{b} - (\underline{a}.\underline{b})\underline{c}]_x \end{aligned}$$

□

2.8 Polar Coordinates

These are alternative choices for our basis vectors

Plane Polar Coordinates

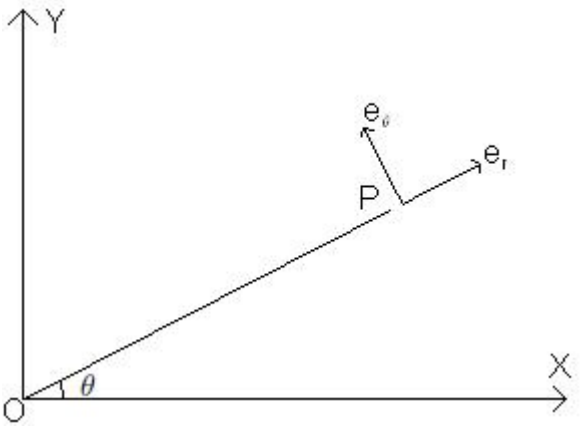

Figure 24: Plane Polar Coordinates

In 2-D, $\boldsymbol{x = r\cos\theta, y = r\sin\theta}$ with $\boldsymbol{0 \leq r < \infty, 0 \leq \theta < 2\pi}$. Note that curves of constant $\boldsymbol{\theta}$ and $\boldsymbol{r}$ are mutually orthogonal.
We take $\underline{\boldsymbol{e_r}}$ as the unit vector $\perp$ to constant $\boldsymbol{r}$, in the direction of increasing $\boldsymbol{r}$, and $\underline{\boldsymbol{e_\theta}}$ as the unit vector $\perp$ to constant $\boldsymbol{\theta}$ in direction of increasing $\boldsymbol{\theta}$

$$\begin{aligned} \underline{\boldsymbol{e_r}} &= \underline{\boldsymbol{i}}\cos\theta + \underline{\boldsymbol{j}}\sin\theta \\ \underline{\boldsymbol{e_\theta}} &= -\underline{\boldsymbol{i}}\sin\theta + \underline{\boldsymbol{j}}\cos\theta \\ \underline{\boldsymbol{e_r}}.\underline{\boldsymbol{e_\theta}} &= 0 \end{aligned}$$

Note that $\underline{\boldsymbol{e_r}}, \underline{\boldsymbol{e_\theta}}$ are position dependent. We can relate polar coordinates to Cartesian coordinates as follows:

$$\underline{\boldsymbol{x}} = \vec{\mathrm{OP}} = x\underline{\boldsymbol{i}} + y\underline{\boldsymbol{j}} = r\cos\theta\underline{\boldsymbol{i}} + r\sin\theta\underline{\boldsymbol{j}} = r\underline{\boldsymbol{e_r}}$$

Cylindrical Polar Coordinates

In 3-D, $\boldsymbol{x = \rho\cos\phi, y = \rho\sin\phi, z = z}$ with $\boldsymbol{0 \leq \rho < \infty, 0 \leq \phi < 2\pi, -\infty < z < \infty}$. Take:

$\underline{\boldsymbol{e_\rho}}$ is the unit vector $\perp$ to surfaces of constant $\boldsymbol{\rho}$, in the direction of $\boldsymbol{\rho}$ increasing
$\underline{\boldsymbol{e_\phi}}$ is the unit vector $\perp$ to surfaces of constant $\boldsymbol{\phi}$, in the direction of $\boldsymbol{\phi}$ increasing
$\underline{\boldsymbol{e_z}} = \underline{\boldsymbol{k}}$

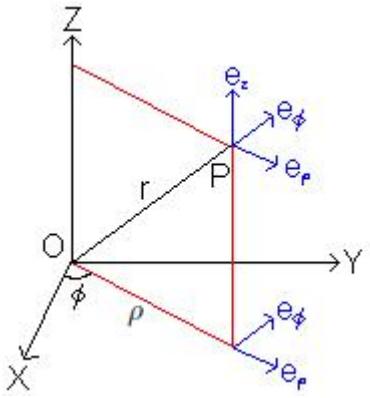

Figure 25: Cylindrical Polar Coordinates

Note that these basis vectors are mutually orthogonal as they are normals to the surfaces $\boldsymbol{\rho}$ = constant, $\boldsymbol{\phi}$ = constant, $\boldsymbol{z}$ = constant. They form a right-handed triad of mutually orthogonal unit vectors.

$$\underline{\boldsymbol{e_\rho}}.\underline{\boldsymbol{e_\phi}} = \underline{\boldsymbol{e_z}}.\underline{\boldsymbol{e_\rho}} = \underline{\boldsymbol{e_\phi}}.\underline{\boldsymbol{e_z}} = \mathbf{0}$$

$$\underline{\boldsymbol{e_\rho}} \times \underline{\boldsymbol{e_\phi}} = \underline{\boldsymbol{e_z}} \qquad \underline{\boldsymbol{e_z}} \times \underline{\boldsymbol{e_\rho}} = \underline{\boldsymbol{e_\phi}} \qquad \underline{\boldsymbol{e_\phi}} \times \underline{\boldsymbol{e_z}} = \underline{\boldsymbol{e_\rho}}$$

$$\underline{\boldsymbol{e_\rho}}.(\underline{\boldsymbol{e_\phi}} \times \underline{\boldsymbol{e_z}}) = \mathbf{1}$$

$$\underline{\boldsymbol{x}} = \vec{\mathrm{ON}} + \vec{\mathrm{NP}} = \boldsymbol{\rho}\underline{\boldsymbol{e_\rho}} + \boldsymbol{z}\underline{\boldsymbol{e_z}}$$

Spherical Polar Coordinates

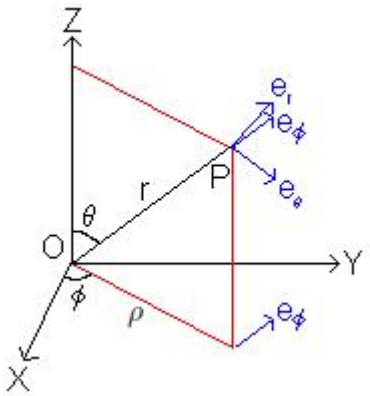

Figure 26: Spherical Polar Coordinates

In 3-D, $\boldsymbol{x} = \boldsymbol{r}\sin\boldsymbol{\theta}\cos\boldsymbol{\phi}, \boldsymbol{y} = \boldsymbol{r}\sin\boldsymbol{\theta}\sin\boldsymbol{\phi}, \boldsymbol{z} = \boldsymbol{r}\cos\boldsymbol{\theta}$ with $\mathbf{0} \leq \boldsymbol{r} < \boldsymbol{\infty}, \mathbf{0} \leq \boldsymbol{\theta} < \boldsymbol{\pi}, \mathbf{0} \leq \boldsymbol{\phi} < \mathbf{2}\boldsymbol{\pi}$. Take:

$\underline{\boldsymbol{e_r}}$ is the unit vector $\perp$ to surfaces of constant $\boldsymbol{r}$, in the direction of $\boldsymbol{r}$ increasing
$\underline{\boldsymbol{e_\theta}}$ is the unit vector $\perp$ to surfaces of constant $\boldsymbol{\theta}$, in the direction of $\boldsymbol{\theta}$ increasing
$\underline{\boldsymbol{e_\phi}}$ is the unit vector $\perp$ to surfaces of constant $\boldsymbol{\phi}$, in the direction of $\boldsymbol{\phi}$ increasing
Note that these basis vectors are mutually orthogonal as they are normals to the surfaces $\boldsymbol{r}$ = constant, $\boldsymbol{\theta}$ = constant, $\boldsymbol{\phi}$ = constant. They form a right-handed triad of mutually orthogonal unit vectors.

$$\underline{\boldsymbol{e_r}}.\underline{\boldsymbol{e_\theta}} = \underline{\boldsymbol{e_\theta}}.\underline{\boldsymbol{e_\phi}} = \underline{\boldsymbol{e_\phi}}.\underline{\boldsymbol{e_r}} = \boldsymbol{0}$$

$$\underline{\boldsymbol{e_r}} \times \underline{\boldsymbol{e_\theta}} = \underline{\boldsymbol{e_\phi}} \qquad \underline{\boldsymbol{e_\theta}} \times \underline{\boldsymbol{e_\phi}} = \underline{\boldsymbol{e_r}} \qquad \underline{\boldsymbol{e_\phi}} \times \underline{\boldsymbol{e_r}} = \underline{\boldsymbol{e_\theta}}$$

$$\underline{\boldsymbol{e_r}}.(\underline{\boldsymbol{e_\theta}} \times \underline{\boldsymbol{e_\phi}}) = \boldsymbol{1}$$

$$\underline{\boldsymbol{x}} = \vec{\mathrm{OP}} = \boldsymbol{r}\underline{\boldsymbol{e_r}}$$

2.9 Suffix Notation

Note that was can describe a vector $\underline{v}$ as

$$\begin{aligned}\underline{v} &= v_x\underline{e_x} + v_y\underline{e_y} + v_z\underline{e_z} \\ &= v_x\underline{i} + v_y\underline{j} + v_z\underline{k} \\ &= (v_x, v_y, v_z) \\ &= v_1\underline{i} + v_2\underline{j} + v_3\underline{k} \\ &= \{v_i : i = 1, 2, 3\}\end{aligned}$$

Using suffix notation we just write $\underline{v} = \{v_i\}$ with the $i = 1, 2, 3$ understood

Similarly $\underline{x} = (x, y, z) = (x_1, x_2, x_3) = \{x_i\}$

This is just a way of writing the components of vectors which allows us to consider n-dimensional ones (otherwise, what would come after $\{v_x, v_y, v_z\}$ for a 4-dimensional vector). In general we also don't need to include the brackets

$$\underline{a} = \underline{b} \Rightarrow a_i = b_i$$

A single suffix like above indicates that the suffix should range through all the possible values (or even, that it is valid for all the possible values that they could take). It is called a free suffix. Similarly

$$\begin{aligned}\underline{c} = \lambda\underline{a} + \mu\underline{b} \Leftrightarrow c_i &= \lambda a_i + \mu b_i \\ \text{or } c_j &= \lambda a_j + \mu b_j \\ \text{or } c_f &= \lambda a_f + \mu b_f\end{aligned}$$

We can use any letter, but it must be consistent.

Scalar Product

$$\underline{a}.\underline{b} = a_1b_1 + a_2b_2 + a_3b_3 = \sum_{i=1}^{3} a_ib_i$$

Here, we have no free suffices. The i is a dummy suffix - as it is repeated within a sum. You can combine both:

$$(\underline{a}.\underline{b})\underline{c} = \underline{d} \Leftrightarrow (\sum_{k=1}^{3} a_kb_k)c_i = d_i$$

In the above equation, $\boldsymbol{k}$ is a dummy suffix and $\boldsymbol{i}$ is a free suffix. Below is one with two dummy suffices

$$(\underline{a}.\underline{b})(\underline{c}.\underline{d}) = (\sum_{i=1}^{3} a_i b_i)(\sum_{j=1}^{3} c_j d_j)$$
$$= \sum_{i=1}^{3}\sum_{j=1}^{3} a_i b_i c_j d_j$$

To avoid ambiguities we only allow suffices to appear twice in each expression (so not $\sum_{i=1}^{3} a_i b_i c_i d_i$)

Einstein Summation Convention

We omit the explicit $\sum$
If a suffix appears twice - it's a dummy suffix, so sum
If a suffix appears once - it's a free suffix, and ranges over the possible values
Examples:

$$\underline{a} = \underline{b} + \underline{c} \quad a_i = b_i + c_i$$
$$(\underline{a}.\underline{b})\underline{c} = \underline{d} \quad a_i b_i c_k = d_k$$

Not allowed are

$$a_k = b_j \text{ - the free suffices are different}$$
$$a_j b_j c_j = d_j \text{ - The } j \text{ appears 3 times}$$

Kronecker Delta - Define δ_{ij} where i, j can be any of $1, 2, 3$ by:

$$\delta_{11} = \delta_{22} = \delta_{33} = 1 \qquad \delta_{ij} = 0 \text{ for } i \neq j$$

$$\begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Properties:

i) $a_i\delta_{ij} = \sum_{i=1}^{3} a_i\delta_{ij} = a_1\delta_{1j} + a_2\delta_{2j} + a_3\delta_{3j} = a_j$
ii) $\delta_{ij}\delta_{jk} = \sum_{j=1}^{3} \delta_{ij}\delta_{jk} = \delta_{i1}\delta_{1k} + \delta_{i2}\delta_{2k} + \delta_{i3}\delta_{3k} = \delta_{ik}$
iii) $\delta_{ii} = \sum_{i=1}^{3} \delta_{ii} = 3$
iv) $a_i\delta_{ij}b_j = a_j b_j = \underline{a}.\underline{b}$

Recall that we have basis vectors $\underline{e_i}$ where $\underline{e_1} = \underline{i}, \underline{e_2} = \underline{j}, \underline{e_3} = \underline{k}$

$$\underline{e_i}.\underline{e_j} = \delta_{ij} \qquad \underline{a}.\underline{e_i} = a_i$$
$$\therefore \underline{e_j}.\underline{e_i} = \delta_{ij} = (\underline{e_j})_i \qquad \text{the } i^{\text{th}} \text{ component of } \underline{e_j}$$

The ϵ symbol - We define ϵ_{ijk} as:

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } \{i,j,k\} \text{ is an even permutation of } \{1,2,3\} \\ -1 & \text{if } \{i,j,k\} \text{ is an odd permutation of } \{1,2,3\} \\ 0 & \text{if any of the } \{i,j,k\} \text{are equal} \end{cases}$$

(Permutations can be thought of as a series of swaps. If it takes an even number of swaps it is an even permutation, otherwise it is odd.)

$$\epsilon_{123} = \epsilon_{312} = \epsilon_{231} = 1 \qquad \epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$$

$$\overbrace{\epsilon_{112} = \epsilon_{122} = \dots}^{\text{21 combinations}} = 0$$

Hence:

$$\epsilon_{kij} = \epsilon_{jki} = \epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{kji} = -\epsilon_{jik}$$

Note: ϵ_{ijk} is also referred to as the alternating tensor - see IB

Vector Product

$$\begin{aligned} (\underline{a} \times \underline{b})_i &= \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{ijk} a_j b_k \\ &= \epsilon_{ijk} a_j b_k \end{aligned}$$

We can make this explicit to check:

$(\underline{a} \times \underline{b})_1 = \epsilon_{123} a_2 b_3 + \epsilon_{132} a_3 b_2$ (all others have a repeated suffix on ϵ and so are 0)
$= a_2 b_3 - a_3 b_2$

Example:

$$\begin{aligned} (\underline{e_j} \times \underline{e_k})_i &= \epsilon_{ilm} \underbrace{\underline{e_j}_l}_{\delta_{jl}} \underbrace{\underline{e_k}_m}_{\delta_{km}} \\ &= \epsilon_{ijm} \delta_{km} \\ &= \epsilon_{ijk} \end{aligned}$$

Theorem 9. $\epsilon_{ijk}\epsilon_{ipq} = \delta_{jp}\delta_{kq} - \delta jq\delta_{kp}$

Proof. Consider $j = k = 1$
LHS: $\epsilon_{i11}\epsilon_{ipq} = 0$
RHS: $\delta_{1p}\delta_{1q} - \delta_{1q}\delta_{1p} = 0$
This also covers the case when $j = k = 2, 3$ and, by symmetry, $p = q = 1, 2, 3$

Consider $j = 1, k - 2$
LHS: $\underbrace{\epsilon_{i12}}_{\text{non-zero only when } i=3} \qquad \epsilon_{ipq} = \epsilon_{312}\epsilon_{3pq} = \epsilon_{3pq}$

Which is 1 if $\boldsymbol{p=1, q=2}$, -1 if $\boldsymbol{p=2, q=1}$ and 0 otherwise
RHS: $\boldsymbol{\delta_{1p}\delta_{2q} - \delta_{1q}\delta_{2p}}$ which is non-zero only if $\boldsymbol{p=1,2}$ and $\boldsymbol{q=1,2}$
$\boldsymbol{(p,q)=(1,2)} \quad \boldsymbol{\delta_{1p}\delta_{2q} - \delta_{1q}\delta_{2p} = \delta_{11}\delta_{22} - \delta_{12}\delta_{21} = 1}$
$\boldsymbol{(p,q)=(2,1)} \quad \boldsymbol{\delta_{1p}\delta_{2q} - \delta_{1q}\delta_{2p} = \delta_{12}\delta_{21} - \delta_{11}\delta_{22} = -1}$
$\boldsymbol{(p,q)=(1,1)} \quad \boldsymbol{\delta_{1p}\delta_{2q} - \delta_{1q}\delta_{2p} = \delta_{11}\delta_{21} - \delta_{11}\delta_{21} = 0}$
$\boldsymbol{(p,q)=(2,2)} \quad \boldsymbol{\delta_{1p}\delta_{2q} - \delta_{1q}\delta_{2p} = \delta_{12}\delta_{22} - \delta_{12}\delta_{22} = 0}$
i.e. the same as the LHS, and similar arguments hold for the other values of $\boldsymbol{j,k}$ □

Corollary 10. Take $\boldsymbol{p=j}$

$$\begin{aligned}\boldsymbol{\epsilon_{ijk}\epsilon_{ijq}} &= \underbrace{\boldsymbol{\delta_{jj}}}_{\boldsymbol{3}}\boldsymbol{\delta_{kq}} - \underbrace{\boldsymbol{\delta jq\delta_{kj}}}_{\boldsymbol{\delta_{kq}}} \\ &= \boldsymbol{2\delta_{kq}}\end{aligned}$$

Vector Triple Product - Proof of the formula

$$\begin{aligned}(\underline{a}\times(\underline{b}\times\underline{c}))_i &= \boldsymbol{\epsilon_{ijk}a_j(\underline{b}\times\underline{c})_k} \\ &= \boldsymbol{\epsilon_{ijk}a_j\epsilon_{kpq}b_pc_q} \\ &= \underbrace{\boldsymbol{\epsilon_{(ijk)}}}_{\boldsymbol{=\epsilon_{kij}}}\boldsymbol{\epsilon_{kpq}a_jb_pc_q} \\ &= \boldsymbol{(\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})(a_jb_pc_q)} \\ &= \boldsymbol{a_jb_ic_j - a_jc_ib_j} \quad \text{using } \boldsymbol{t_{ij}\delta_{jk} = t_{ik}} \\ &= \boldsymbol{(\underline{a}.\underline{c})b_i - (\underline{a}.\underline{b})c_i} \\ &= \boldsymbol{((\underline{a}.\underline{c})\underline{b} - (\underline{a}.\underline{b})\underline{c})_i}\end{aligned}$$

2.10 Vector equations for geometric objects

Lines:

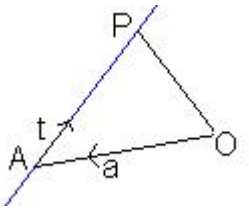

Figure 27: Vector equation of a line

The line shown is a line through A $\parallel$ to $\underline{\boldsymbol{t}}$

$$\vec{\mathrm{OP}} = \vec{\mathrm{OA}} + \vec{\mathrm{AP}}$$
$$\underline{\boldsymbol{x}} = \underline{\boldsymbol{a}} + \boldsymbol{\lambda}\underline{\boldsymbol{t}}$$

Or, $\underline{\boldsymbol{x}} - \underline{\boldsymbol{a}} = \boldsymbol{\lambda}\underline{\boldsymbol{t}} \Rightarrow (\underline{\boldsymbol{x}} - \underline{\boldsymbol{a}}) \times \underline{\boldsymbol{t}} = \boldsymbol{\lambda}\underline{\boldsymbol{t}} \times \underline{\boldsymbol{t}} = \underline{\mathbf{0}}$
$\therefore (\underline{\boldsymbol{x}} - \underline{\boldsymbol{a}}) \times \underline{\boldsymbol{t}} = \underline{\mathbf{0}}$, which is the vector equation for a line.
Since $\underline{\boldsymbol{t}} \neq \underline{\mathbf{0}} \Rightarrow (\underline{\boldsymbol{x}} - \underline{\boldsymbol{a}}) = \underline{\mathbf{0}}$ or $(\underline{\boldsymbol{x}} - \underline{\boldsymbol{a}}) \parallel \underline{\boldsymbol{t}}$ i.e. $\underline{\boldsymbol{x}} - \underline{\boldsymbol{a}} = \boldsymbol{\lambda}\underline{\boldsymbol{t}}$ for some $\boldsymbol{\lambda}$

Note that the equation has a non-unique solution, as we have 3 equations for 3 unknowns in terms of components. The multiplicity of solutions is represented by a single unknown scalar (we have a one-dimensional set of solutions)

Example: Let $\underline{\boldsymbol{x}} \times \underline{\boldsymbol{t}} = \underline{\boldsymbol{u}}$

$$\therefore \underbrace{(\underline{\boldsymbol{x}} \times \underline{\boldsymbol{t}}).\underline{\boldsymbol{t}}}_{\text{volume of parallelipiped} = 0} = \underline{\boldsymbol{u}}.\underline{\boldsymbol{t}}$$
$$\therefore \underline{\boldsymbol{u}}.\underline{\boldsymbol{t}} = \mathbf{0} \quad (\text{if } \underline{\boldsymbol{u}}.\underline{\boldsymbol{t}} \neq \mathbf{0} \text{ there are no solutions})$$
$$\therefore (\underline{\boldsymbol{x}} \times \underline{\boldsymbol{t}}) \times \underline{\boldsymbol{t}} = (\underline{\boldsymbol{x}}.\underline{\boldsymbol{t}})\underline{\boldsymbol{t}} - (\underline{\boldsymbol{t}}.\underline{\boldsymbol{t}})\underline{\boldsymbol{x}} = \underline{\boldsymbol{u}} \times \underline{\boldsymbol{t}}$$
$$\Rightarrow \underline{\boldsymbol{x}} = \frac{(\underline{\boldsymbol{x}}.\underline{\boldsymbol{t}})\underline{\boldsymbol{t}}}{|\underline{\boldsymbol{t}}|^2} - \frac{\underline{\boldsymbol{u}} \times \underline{\boldsymbol{t}}}{|\underline{\boldsymbol{t}}|^2}$$
$$\Rightarrow \underline{\boldsymbol{x}} = \frac{\boldsymbol{\lambda}\underline{\boldsymbol{t}}}{|\underline{\boldsymbol{t}}|^2} - \frac{\underline{\boldsymbol{u}} \times \underline{\boldsymbol{t}}}{|\underline{\boldsymbol{t}}|^2} \text{ with } \boldsymbol{\lambda} \in \mathbb{R}$$

So the solutions lie on a straight line $\parallel \underline{\boldsymbol{t}}$ and through $-\frac{\underline{\boldsymbol{u}} \times \underline{\boldsymbol{t}}}{|\underline{\boldsymbol{t}}|^2}$

Planes

Consider a plane through A, $\perp$ to the unit vector $\underline{\boldsymbol{n}}$ ($\underline{\boldsymbol{n}}$ is the normal to the plane)

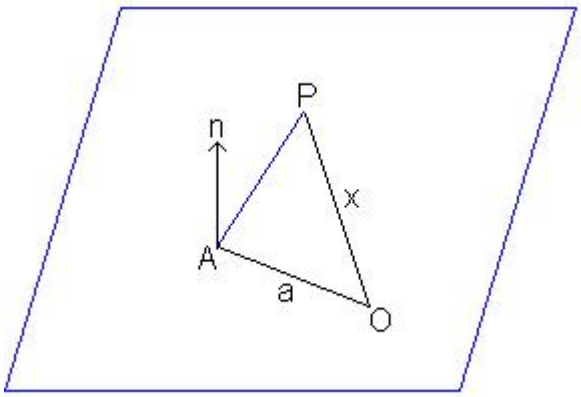

Figure 28: Vector equation of a plane

$$\vec{AP} \perp \underline{n} \Rightarrow \vec{AP}.\underline{n} = 0$$
$$(\vec{OP} - \vec{OA}).\underline{n} = 0$$
$$\therefore (\underline{x} - \underline{a}).\underline{n} = 0$$

The Z in the plane with $\vec{OZ} \perp$ to plane $\Rightarrow \vec{OZ} = d\underline{n}$ with d the $\perp$ distance from the plane to the origin.

Z lies in the plane, $\therefore (d\underline{n} - \underline{a}).\underline{n} = 0 \Rightarrow \underbrace{\underline{a}.\underline{n}}_{\text{distance from plane to origin}} = d\underline{n}^2 = d$,

giving the following equation for the plane:

$$\underline{x}.\underline{n} = \underline{a}.\underline{n} = d$$

If $\underline{l}, \underline{m}$ are linearly independent and $\perp$ to $\underline{n}$ (i.e. $\underline{l}.\underline{n} = \underline{m}.\underline{n} = 0$) then $\underline{x} = \underline{a} + \lambda\underline{l} + \mu\underline{m}$ $(\lambda, \mu \in \mathbb{R})$ is the solution of the equation for the plane.
Two independent scalars describe the multiplicity of solutions.
Example: Take two lines
$\boldsymbol{L_1 : (\underline{x} - \underline{a}) \times \underline{t} = 0}$
$\boldsymbol{L_2 : (\underline{x} - \underline{b}) \times \underline{u} = 0}$
Under which circumstances do L_1, L_2 intersect?

Assume L_1, L_2 not $\parallel$ i.e. $\underline{t}, \underline{u}$ linearly independent.
Consider the plane Π containing L_1 and $\parallel$ to L_2 (contains $L_2' \parallel L_2$)
The normal is $\perp$ to $\underline{t}, \underline{u}$ i.e. $\parallel$ to $\underline{t} \times \underline{u}$
Hence Π has equation

$$(\underline{x} - \underline{a}).(\underline{t} \times \underline{u}) = 0$$

$\boldsymbol{L_2} \parallel \boldsymbol{\Pi} \Rightarrow$ no intersection of $\boldsymbol{L_2}$ lies in $\boldsymbol{\Pi}$, which means $\underline{\boldsymbol{b}}$ lies in $\boldsymbol{\Pi}$

$$\Rightarrow (\underline{\boldsymbol{b}} - \underline{\boldsymbol{a}}).(\underline{\boldsymbol{t}} \times \underline{\boldsymbol{u}}) = \boldsymbol{0}$$

This is necessary for intersection. But is this sufficient?
This $\Rightarrow \underline{\boldsymbol{b}} - \underline{\boldsymbol{a}}$ lies on a plane through the origin $\perp \underline{\boldsymbol{t}} \times \underline{\boldsymbol{u}}$
Hence $\underline{\boldsymbol{b}} - \underline{\boldsymbol{a}} = \boldsymbol{\lambda}\underline{\boldsymbol{t}} + \boldsymbol{\mu}\underline{\boldsymbol{u}}$ for some $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}$
$\therefore \exists \underline{\boldsymbol{x}} = \underbrace{\underline{\boldsymbol{a}} + \boldsymbol{\lambda}\underline{\boldsymbol{t}}}_{\text{lies in } \boldsymbol{L_1}} = \underbrace{\underline{\boldsymbol{b}} - \boldsymbol{\mu}\underline{\boldsymbol{u}}}_{\text{lies in } \boldsymbol{L_2}}$
So this is sufficient, and so $(\underline{\boldsymbol{b}} - \underline{\boldsymbol{a}}).(\underline{\boldsymbol{t}} \times \underline{\boldsymbol{u}}) = \boldsymbol{0}$ is sufficient for intersection.

Vector Equations

One approach is to write out components:

e.g. $(\underline{\boldsymbol{x}} - \underline{\boldsymbol{a}}.\underline{\boldsymbol{n}} = \boldsymbol{0} \Rightarrow \boldsymbol{x_1 n_1 + x_2 n_2 + x_3 n_3 = a_1 n_1 + a_2 n_2 + a_3 n_3}$, a single linear equation for 3 unknowns.
For another approach, use vector algebra operations to reorganise:

Example:

$$\begin{aligned} \underline{\boldsymbol{x}} - (\underline{\boldsymbol{x}} \times \underline{\boldsymbol{a}}) \times \underline{\boldsymbol{b}} &= \underline{\boldsymbol{c}} \\ \underline{\boldsymbol{x}} - \underline{\boldsymbol{a}}(\underline{\boldsymbol{b}}.\underline{\boldsymbol{x}}) + \underline{\boldsymbol{x}}(\underline{\boldsymbol{b}}.\underline{\boldsymbol{a}}) &= \underline{\boldsymbol{c}} \\ \therefore \underline{\boldsymbol{x}}(\boldsymbol{1} + \underline{\boldsymbol{b}}.\underline{\boldsymbol{a}}) &= \underline{\boldsymbol{c}} + \underline{\boldsymbol{a}}(\underline{\boldsymbol{c}}.\underline{\boldsymbol{b}}) \\ \therefore \underline{\boldsymbol{x}} &= \frac{\underline{\boldsymbol{c}} + \underline{\boldsymbol{a}}(\underline{\boldsymbol{c}}.\underline{\boldsymbol{b}})}{\boldsymbol{1} + \underline{\boldsymbol{b}}.\underline{\boldsymbol{a}}} \end{aligned}$$

Alternatively we could go down a different route:

$$\begin{aligned} \underline{\boldsymbol{x}} - (\underline{\boldsymbol{x}} \times \underline{\boldsymbol{a}}) \times \underline{\boldsymbol{b}} &= \underline{\boldsymbol{c}} \\ \underline{\boldsymbol{b}}.(\underline{\boldsymbol{x}} - (\underline{\boldsymbol{x}} \times \underline{\boldsymbol{a}}) \times \underline{\boldsymbol{b}}) &= \underline{\boldsymbol{b}}.\underline{\boldsymbol{c}} \\ \Rightarrow \underline{\boldsymbol{b}}.\underline{\boldsymbol{x}} - ((\underline{\boldsymbol{x}} \times \underline{\boldsymbol{a}}) \times \underline{\boldsymbol{b}}).\underline{\boldsymbol{b}} &= \underline{\boldsymbol{b}}.\underline{\boldsymbol{c}} \\ \therefore \underline{\boldsymbol{b}}.\underline{\boldsymbol{x}} &= \underline{\boldsymbol{b}}.\underline{\boldsymbol{c}} \end{aligned}$$

As another alternative we could use $\underline{\boldsymbol{a}}, \underline{\boldsymbol{c}}$ as two elements of the basis, and take $\underline{\boldsymbol{c}} \times \underline{\boldsymbol{a}}$ as the $\boldsymbol{3}^{\text{rd}}$ element (if $\underline{\boldsymbol{c}} \not\parallel \underline{\boldsymbol{a}}$

2.11 Cones and conic sections

Take a right circular cone surface containing points P on which lines OP make a fixed angle $\boldsymbol{\alpha}$ with the axis (with O as the vertex).

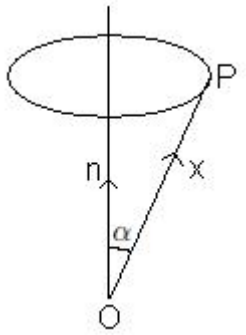

Figure 29: Right circular cone

$$\underline{x}.\underline{n} = |\underline{x}||\underline{n}| \cos \alpha$$
$$\therefore (\underline{x}.\underline{n})^2 = x^2 \cos^2 \alpha$$

Note: This also allows the 'reverse cone', reflected below point O

We can generalise to a cone with the vertex at an arbitrary point $\underline{a} \Rightarrow ((\underline{x} - \underline{a}).\underline{n})^2 = (\underline{x} - \underline{a})^2 \cos^2 \alpha$

We can also write this in terms of coordinates (x, y, a)

$\left.\begin{array}{rcl} \underline{a} & = & (a, b, c) \\ \underline{n} & = & (l, m, n) \end{array}\right\}$ with respect to standard basis

such that $l^2 + m^2 + n^2 = 1$

$$\therefore ((x-a)l+(y-b)m+(z-c)n)^2 = ((x-a)^2+(y-b)^2+(z-c)^2) \cos^2 \alpha$$

which is the algebraic equation for a cone.

Consider the intersection with $z = 0$

$$((x - a)l + (y - b)m - cn)^2 = ((x - a)^2 + (y - b)^2 + c^2) \cos^2 \alpha$$

which is a quadratic function of the form $f(x, y) =$ constant

WLOG take $\boldsymbol{l = 0}$, hence $\boldsymbol{(l, m, n) = (0, \sin\beta, \cos\beta)}$
Write $\boldsymbol{X = x - a, Y = y - b - \frac{cmn}{\sin^2\beta - \cos^2\beta}}$

$$\boldsymbol{\Rightarrow X^2\cos^2\alpha + Y^2(\cos^2\alpha - \sin^2\beta) = \frac{c^2\sin^2\alpha\cos^2\alpha}{\cos^2\alpha - \sin^2\beta}}$$

We have 3 potential cases:
<u>case 1</u>

$\boldsymbol{\alpha + \beta < \frac{\pi}{2} \Rightarrow \beta < \frac{\pi}{2} - \alpha \Rightarrow \cos\alpha > \sin\beta}$

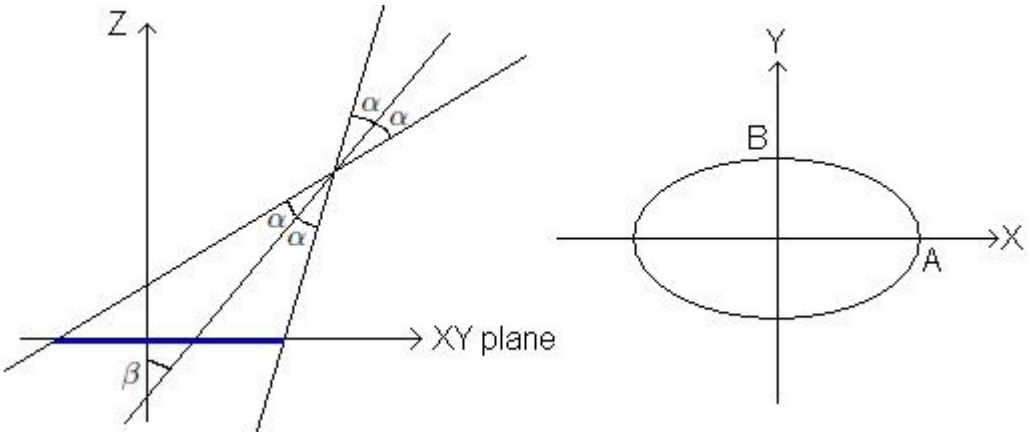

Figure 30: Ellipse as a conic section

Let $\boldsymbol{A^2 = \frac{c^2\sin^{-2}}{\cos^2\alpha - \sin^2\beta}}$ and $\boldsymbol{B^2 = \frac{c^2\sin^{-2}\cos^2\alpha}{\cos^2\alpha - \sin^{-2}\beta}}$
Then we get the equation for a ELLIPSE - a single bounded curve:

$$\boldsymbol{\frac{X^2}{A^2} + \frac{Y^2}{B^2} = 1}$$

<u>case 2</u>

$\boldsymbol{\alpha + \beta > \frac{\pi}{2} \Rightarrow \cos^2\alpha < \sin^{-2}\beta}$

Let $\boldsymbol{A^2 = \frac{c^2\sin^{-2}\alpha}{\sin^2\beta - \cos^2\alpha}}$ and $\boldsymbol{B^2 = \frac{c^2\sin^{-2}\alpha\cos^2\alpha}{(\cos^2\alpha - \sin^{-2}\beta)^2}}$
Then we get the equation for HYPERBOLAE - a pair of double unbounded curves:

$$\boldsymbol{-\frac{X^2}{A^2} + \frac{Y^2}{B^2} = 1}$$

<u>case 3</u>

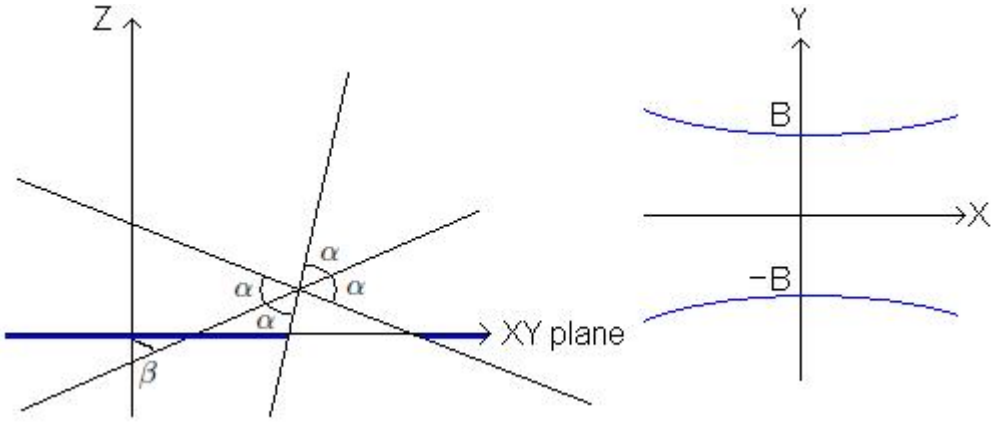

Figure 31: Hyperbolae as conic sections

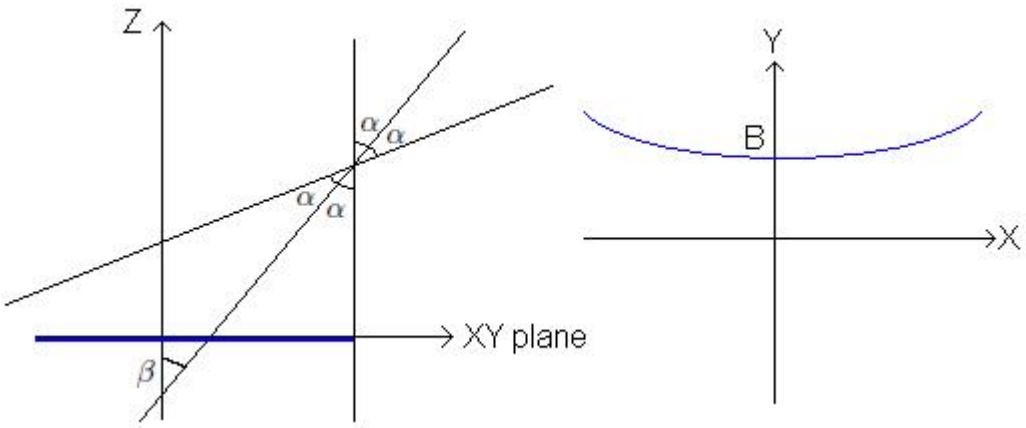

Figure 32: Parabola as a conic section

$\boldsymbol{\cos\alpha = \sin\beta \Rightarrow \alpha + \beta = \frac{\pi}{2}}$

In this case, $\boldsymbol{-\cos^2\alpha(x-a)^2 - 2c\sin\alpha\cos\alpha(y-b) + c^2(\cos^2\beta - \cos^2\alpha) = 0}$, (there are no second order terms in $\boldsymbol{y}$)

Let $\boldsymbol{X = x - a, Y = y - b - \frac{c}{2}(\cos^2\beta - \cos^2\alpha)}$ Thus we get the equation for a PARABOLA - a single unbounded curve

$$\boldsymbol{X^2 = -2c\tan\alpha Y}$$

This represents slicing the cone in a plane $\parallel$ to the tangent plane.

2.12 Maps: Isometrics and Inversions

Isometry - An Isometry is a mapping from $\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ under which $\underline{\boldsymbol{x}} \mapsto \underline{\boldsymbol{x}'}$ such that, for any pair $\underline{\boldsymbol{x_1}}, \underline{\boldsymbol{x_2}}$: $(\underline{\boldsymbol{x_1}} \mapsto \underline{\boldsymbol{x_1'}}, \underline{\boldsymbol{x_2}} \mapsto \underline{\boldsymbol{x_2'}}, \underline{\boldsymbol{x_1}}, \underline{\boldsymbol{x_1'}}, \underline{\boldsymbol{x_2}}, \underline{\boldsymbol{x_2'}} \in \mathbb{R}^3)$

$$|\underline{\boldsymbol{x_1}} - \underline{\boldsymbol{x_2}}| = |\underline{\boldsymbol{x_1'}} - \underline{\boldsymbol{x_2'}}|$$

i.e. distance is preserved

Examples:
1) Reflection in the plane $\boldsymbol{\Pi}$ with $\boldsymbol{\Pi} = \{\underline{\boldsymbol{x}} \in \mathbb{R}^3 : \underline{\boldsymbol{x}}.\underline{\boldsymbol{n}} = 0\}$ with $\underline{\boldsymbol{n}}$ a constant unit vector.

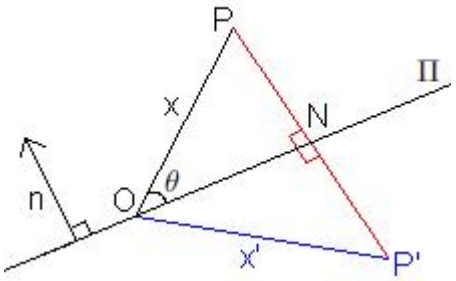

Figure 33: Reflection in a plane

$$\begin{aligned}
\therefore \underline{\boldsymbol{x}} &= \vec{\mathrm{OP}} \\
\underline{\boldsymbol{x'}} &= \vec{\mathrm{OP}}\text{'} \\
\vec{\mathrm{NP}}\text{'} &= \vec{\mathrm{PN}} = -\vec{\mathrm{NP}} \\
\underline{\boldsymbol{x'}} = \vec{\mathrm{OP}}\text{'} &= \vec{\mathrm{OP}} + \vec{\mathrm{PP}}\text{'} \\
&= \vec{\mathrm{OP}} + \mathbf{2}\vec{\mathrm{NP}}\text{'} \\
&= \vec{\mathrm{OP}} - \mathbf{2}\vec{\mathrm{NP}}
\end{aligned}$$

But $|\vec{\mathrm{NP}}| = |\underline{\boldsymbol{x}}| \sin \boldsymbol{Q}$ and $\vec{\mathrm{NP}} = |\underline{\boldsymbol{x}}|\underline{\boldsymbol{n}} \sin \boldsymbol{Q}$
$\therefore \vec{\mathrm{NP}} = (\underline{\boldsymbol{x}}.\underline{\boldsymbol{n}})\underline{\boldsymbol{n}}$

Finally $\underline{\boldsymbol{x'}} = \underline{\boldsymbol{x}} - \mathbf{2}(\underline{\boldsymbol{x}}.\underline{\boldsymbol{n}})\underline{\boldsymbol{n}}$

$\underline{\boldsymbol{x'}}$ is a linear function of $\underline{\boldsymbol{x}}$ (the components of $\underline{\boldsymbol{x'}}$ are linear functions of components of $\underline{\boldsymbol{x}}$).

Is distance preserved?
Take $\underline{x_1'}, \underline{x_2'}$ such that:
$\underline{x_1'} = \underline{x_1} - 2(\underline{x_1}.\underline{n})\underline{n}$
$\underline{x_2'} = \underline{x_2} - 2(\underline{x_2}.\underline{n})\underline{n}$

$$\begin{aligned}
|\underline{x_1'} - \underline{x_2'}|^2 &= (\underline{x_1'} - \underline{x_2'}).(\underline{x_1'} - \underline{x_2'}) \\
&= (\underline{x_1} - \underline{x_2} - 2(\underline{x_1}.\underline{n})\underline{n} + 2(\underline{x_2}.\underline{n})\underline{n})^2 \\
&= (\underline{x_{12}} - 2(\underline{x_{12}}.\underline{n})\underline{n})^2 \\
&= \underline{x_{12}}.\underline{x_{12}} - 4(\underline{x_{12}}.\underline{n})(\underline{x_{12}}.\underline{n}) + 4(\underline{x_{12}}.\underline{n})^2\underline{n}^2 \\
&= \underline{x_{12}}.\underline{x_{12}} = |\underline{x_1} - \underline{x_2}|^2
\end{aligned}$$

$\therefore$ it is an isometry.

2) Translation - $\underline{x} \mapsto \underline{x'} = \underline{x} + \underline{a}$ where $\underline{a}$ is a constant.
This is trivially an isometry

3) Inversion in a sphere

Let $\Sigma - \{\underline{x} \in \mathbb{R}^3 : |\underline{x}| = k \in \mathbb{R}(k > 0)\}$, a sphere centred on the origin with radius k

Inverse Point - Given a point P, the inverse point P' (with respect to Σ) lies on the line OP such that

$$|\vec{\text{OP}}'| = \frac{k^2}{|\vec{\text{OP}}|}$$

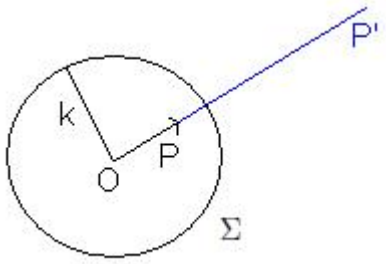

Figure 34: Inversion in a sphere

Let $\vec{\text{OP}} = \underline{x}, \vec{\text{OP}}' = \underline{x'}$, then $\underline{x'} = \frac{|\underline{x'}|}{|\underline{x}|}\underline{x}$ as they are in the same direction

$$\Rightarrow \underline{x'} = \frac{k^2}{|\underline{x}|^2}\underline{x}$$

Therefore this is NOT an isometry

Note that (P')' = P, and that inversion of the points on a sphere gives another sphere:

Add another sphere with centre $\boldsymbol{a}$ and radius $\boldsymbol{r}$

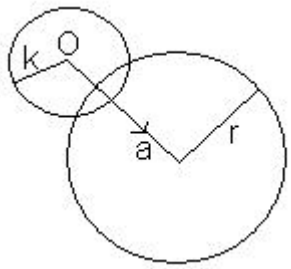

Figure 35: Inversion of one sphere in another

Our new sphere has equation $|\underline{x} - \underline{a}|^2 = r^2 \Rightarrow \underline{x}^2 - 2\underline{x}.\underline{a} + \underline{a}^2 = r^2$

Under the inverse $\underline{x'} = \frac{k^2}{|\underline{x}|^2}\underline{x}$:

$$\frac{k^4}{|\underline{x'}|^2} - \frac{2\underline{a}.\underline{x'}k^2}{|\underline{x'}|^2} + a^2 = r^2$$
$$\therefore (a^2 - r^2)|\underline{x'}|^2 - 2\underline{a}.\underline{x'}k^2 + k^2 = 0$$
$$\therefore |\underline{x'}|^2 - \frac{2\underline{a}.\underline{x'}k^2}{a^2 - r^2} + \frac{k^2}{a^2 - r^2} = 0$$
$$|\underline{x} - \frac{\underline{a}k^2}{a^2 - r^2}|^2 = \frac{r^2k^4}{(a^2 - r^2)^2}$$

This is a new sphere with centre $\frac{\underline{a}k^2}{a^2 - r^2}$ and radius $\frac{r^2k^4}{(a^2 - r^2)^2}$. We only have a problem when $\boldsymbol{r = a}$ i.e. the sphere to be inverted passes through the origin.

3 Vector Spaces

3.1 Definition of a Vector Space

Vector Space - A set $\boldsymbol{V}$ of vectors is a vector space over $\mathbb{R}$ if it satisfies the following rules:

A: Addition - for each pair $\underline{\boldsymbol{x}}, \underline{\boldsymbol{y}} \in \boldsymbol{V}$ then $\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} \in \boldsymbol{V}$ (i.e. closed under addition) and:

A1: $\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} = \underline{\boldsymbol{y}} + \underline{\boldsymbol{x}}$ (commutative)
A2: $(\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}}) + \underline{\boldsymbol{z}} = \underline{\boldsymbol{x}} + (\underline{\boldsymbol{y}} + \underline{\boldsymbol{z}})$ (associative)
A3: $\exists$ a zero vector $\underline{\mathbf{0}} \in \boldsymbol{V}$ such that $\underline{\boldsymbol{x}} + \underline{\mathbf{0}} = \underline{\boldsymbol{x}}\ \forall \underline{\boldsymbol{x}} \in \boldsymbol{V}$
A4: For each $\underline{\boldsymbol{x}} \in \boldsymbol{V} \exists$ negative vector $(-\underline{\boldsymbol{x}})$ such that
$\underline{\boldsymbol{x}} + (-\underline{\boldsymbol{x}}) = (-\underline{\boldsymbol{x}}) + \underline{\boldsymbol{x}} = \underline{\mathbf{0}}$

B: Multiplication by scalars - real numbers for our purposes
For each $\boldsymbol{\lambda} \in \mathbb{R}, \underline{\boldsymbol{x}} \in \boldsymbol{V} \exists!$ vector $\boldsymbol{\lambda}\underline{\boldsymbol{x}}$ with $\boldsymbol{\lambda}\underline{\boldsymbol{x}} \in \boldsymbol{V}$ (closed under scalar multiplication) and:

B1: $\boldsymbol{\lambda}(\boldsymbol{\mu}\underline{\boldsymbol{x}}) = (\boldsymbol{\lambda}\boldsymbol{\mu})\underline{\boldsymbol{x}}$ (associative)
B2: $\mathbf{1}.\underline{\boldsymbol{x}} = \underline{\boldsymbol{x}}\ \forall \underline{\boldsymbol{x}} \in \boldsymbol{V}$
B3: $\boldsymbol{\lambda}(\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}}) = \boldsymbol{\lambda}\underline{\boldsymbol{x}} + \boldsymbol{\lambda}\underline{\boldsymbol{y}}$ (distributive)
B4: $(\boldsymbol{\lambda} + \boldsymbol{\mu})\underline{\boldsymbol{x}} = \boldsymbol{\lambda}\underline{\boldsymbol{x}} + \boldsymbol{\mu}\underline{\boldsymbol{x}}$ (distributive)

Implications of the above
①: The zero vector is unique:
If we have two zero vectors, $\underline{\mathbf{0}}$ and $\underline{\mathbf{0}'}$ such that $\forall \underline{\boldsymbol{x}}, \underline{\mathbf{0}} + \underline{\boldsymbol{x}} = \underline{\mathbf{0}'} + \underline{\boldsymbol{x}} = \underline{\boldsymbol{x}}$, then $\underline{\mathbf{0}} + \underline{\mathbf{0}'} =$ both $\underline{\mathbf{0}}$ and $\underline{\mathbf{0}'}$ by the properties of the zero vector, so $\underline{\mathbf{0}} = \underline{\mathbf{0}'}$

②: $\mathbf{0}.\underline{\boldsymbol{x}} = \underline{\mathbf{0}}\ \forall \underline{\boldsymbol{x}}$ since
$\underline{\boldsymbol{x}} + \mathbf{0}.\underline{\boldsymbol{x}} = (\mathbf{1} + \mathbf{0})\underline{\boldsymbol{x}} = \mathbf{1}\underline{\boldsymbol{x}} = \underline{\boldsymbol{x}}$
$\therefore -\underline{\boldsymbol{x}} + (\underline{\boldsymbol{x}} + \mathbf{0}.\underline{\boldsymbol{x}}) = -\underline{\boldsymbol{x}} + \underline{\boldsymbol{x}} = \underline{\mathbf{0}}$
Rearraning we get $(-\underline{\boldsymbol{x}} + \underline{\boldsymbol{x}}) + \mathbf{0}.\underline{\boldsymbol{x}} = \mathbf{0}.\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$

③: The negative is unique.
If we have $\underline{\boldsymbol{y}}, \underline{\boldsymbol{z}}$ such that $\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} = \underline{\boldsymbol{x}} + \underline{\boldsymbol{z}} = \underline{\mathbf{0}}$
$\underline{\boldsymbol{z}} = \underline{\mathbf{0}} + \underline{\boldsymbol{z}} = \underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} + \underline{\boldsymbol{z}} = \underline{\boldsymbol{x}} + \underline{\boldsymbol{z}} + \underline{\boldsymbol{y}} = \underline{\mathbf{0}} + \underline{\boldsymbol{y}} = \underline{\boldsymbol{y}}$

④: $(-\mathbf{1})\underline{\boldsymbol{x}} = -\underline{\boldsymbol{x}}$
$\because (-\mathbf{1})\underline{\boldsymbol{x}} + \underline{\boldsymbol{x}} = (-\mathbf{1} + \mathbf{1})\underline{\boldsymbol{x}} = \mathbf{0}.\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$ so $(-\mathbf{1})\underline{\boldsymbol{x}}$ is the negative vector for $\underline{\boldsymbol{x}}$

⑤: $\boldsymbol{\lambda}\underline{\mathbf{0}} = \underline{\mathbf{0}}\ \forall \boldsymbol{\lambda} \in \mathbb{R}$
$\because \boldsymbol{\lambda}\underline{\mathbf{0}} + \boldsymbol{\lambda}\underline{\boldsymbol{x}} = \boldsymbol{\lambda}(\underline{\mathbf{0}} + \underline{\boldsymbol{x}}) = \boldsymbol{\lambda}\underline{\boldsymbol{x}}$
$\Rightarrow \boldsymbol{\lambda}\underline{\mathbf{0}} = \underline{\mathbf{0}}$

⑥: If $\boldsymbol{\lambda}\underline{\boldsymbol{x}} = \mathbf{0}$ then either $\boldsymbol{\lambda} = \mathbf{0}$ or $\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$
Because if $\boldsymbol{\lambda} \neq \mathbf{0}$ then $\exists \boldsymbol{\lambda}^{-\mathbf{1}}$ such that $\boldsymbol{\lambda}\boldsymbol{\lambda}^{-\mathbf{1}} = \mathbf{1}$

$\therefore \boldsymbol{\lambda^{-1}\lambda\underline{x} = \lambda^{-1}\underline{0} = \underline{0}}$, but also $\boldsymbol{\lambda^{-1}\lambda\underline{x} = 1.\underline{x} = \underline{x}}$, therefore $\boldsymbol{\underline{x} = \underline{0}}$

We use rules A1-A4, B1-B4 as the axiomatic definition of a vector space. This allows the generalisation of the ideas of vectors to other mathematical objects e.g. polynomials, functions etc.

Examples:
i) The set of all n-tuples: $\boldsymbol{\underline{X} = (x_1, x_2, \ldots, x_n)}$ with $\boldsymbol{x_i} \in \mathbb{R}$ (called $\mathbb{R}^{\boldsymbol{n}}$)
If $\boldsymbol{\underline{X}, \underline{Y}} \in \mathbb{R}^{\boldsymbol{n}}$ with $\boldsymbol{\underline{Y} = (y_1, y_2, \ldots, y_n)}$ we define
$\boldsymbol{\underline{X} + \underline{Y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)}$
$\boldsymbol{\lambda\underline{X} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)}$
$\boldsymbol{\underline{0} = (0, 0, \ldots, 0)}$
$\boldsymbol{-\underline{X} = (-x_1, -x_2, \ldots, -x_n)}$
All of the rules A1-4, B1-4 are satisfied, hence $\mathbb{R}^{\boldsymbol{n}}$ is a vector space over $\mathbb{R}$

ii) Functions $\boldsymbol{f(x)}$ of a real variable $\boldsymbol{x} \in [\boldsymbol{0, 1}]$. We define
$\boldsymbol{(f + g)(x) = f(x) + g(x)}$
$\boldsymbol{(\lambda f)(x) = \lambda f(x)}$
$\boldsymbol{0(x) = 0\ \forall x : 0 \leq x \leq 1}$
$\boldsymbol{(-f)(x) = -f(x)}$

3.2 Subspaces

Subspace - A subset $\boldsymbol{U}$ of a vector space $\boldsymbol{V}$ is a subspace if $\boldsymbol{U}$ itself is a vector space (under the same definition of addition and scalar multiplication used for $\boldsymbol{V}$). Strictly both $\boldsymbol{V}$ and $\{\underline{\mathbf{0}}\}$ are subspaces of $\boldsymbol{V}$, a proper subspace is any subspace which is neither $\boldsymbol{V}$ nor $\{\underline{\mathbf{0}}\}$

Theorem 11. A subset $\boldsymbol{U}$ of a vector space $\boldsymbol{V}$ is a subspace under the operations defined on $\boldsymbol{V}$ iff:
i) for each $\underline{\boldsymbol{x}}, \underline{\boldsymbol{y}} \in \boldsymbol{U} \quad \underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} \in \boldsymbol{U}$
ii) for each $\underline{\boldsymbol{x}} \in \boldsymbol{U} \quad \boldsymbol{\lambda}\underline{\boldsymbol{x}} \in \boldsymbol{U} \; \forall \boldsymbol{\lambda} \in \mathbb{R}$

Proof. only if: If $\boldsymbol{U}$ is a subspace then it must be closed under addition and scalar multiplication
if: A1, A2, B1-4 all straightforward.
A3: ii) $\forall \underline{\boldsymbol{x}} \in \boldsymbol{U}, \mathbf{0}\underline{\boldsymbol{x}} \in \boldsymbol{U} \Rightarrow \underline{\mathbf{0}} \in \boldsymbol{U}$
A4: For any $\underline{\boldsymbol{x}} \in \boldsymbol{U}$, ii) $\Rightarrow (\boldsymbol{-1})\underline{\boldsymbol{x}} \in \boldsymbol{U}$
but by A4 $(\boldsymbol{-1})\underline{\boldsymbol{x}} = -\underline{\boldsymbol{x}} \in \boldsymbol{U}$ □

Examples:

i) Let $\boldsymbol{V} = \mathbb{R}^{\boldsymbol{n}}$ i.e. $\boldsymbol{V}$ is the set of $\underline{\boldsymbol{X}} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}) \quad \boldsymbol{x_i} \in \mathbb{R}$. Let $\boldsymbol{U} = \{\underline{\boldsymbol{X}} : \underline{\boldsymbol{X}} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_{n-1}}, \mathbf{0}\}$. Then $\boldsymbol{U}$ is a subspace of $\boldsymbol{V}$.

Proof. (Note that we can combine addition and scalar multiplication to save effort)
$\boldsymbol{\lambda}(\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_{n-1}}, \mathbf{0}) + \boldsymbol{\mu}(\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_{n-1}}, \mathbf{0}) = (\boldsymbol{\lambda x_1} + \boldsymbol{\mu y_1}, \boldsymbol{\lambda x_2} + \boldsymbol{\mu y_2}, \ldots, \boldsymbol{\lambda x_{n-1}} + \boldsymbol{\mu y_{n-1}}, \mathbf{0}) \in \boldsymbol{U}$ so $\boldsymbol{U}$ is closed under addition and scalar multiplication. $\therefore \boldsymbol{U}$ is a subspace. □

ii) $\boldsymbol{W} = \{\underline{\boldsymbol{X}} : \underline{\boldsymbol{X}} \in \mathbb{R}^{\boldsymbol{n}}, \underline{\boldsymbol{X}} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n})$ and $\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\lambda_i x_i} = \mathbf{0}$ for fixed $\boldsymbol{\lambda_1}, \boldsymbol{\lambda_2}, \ldots, \boldsymbol{\lambda_n}$

Proof. Take $\underline{\boldsymbol{X}}, \underline{\boldsymbol{Y}} \in \boldsymbol{W}$. Let $\underline{\boldsymbol{Z}} = \boldsymbol{\alpha}\underline{\boldsymbol{X}} + \boldsymbol{\beta}\underline{\boldsymbol{Y}} = (\boldsymbol{\alpha x_1} + \boldsymbol{\beta y_1}, \boldsymbol{\alpha x_2} + \boldsymbol{\beta y_2}, \ldots, \boldsymbol{\alpha x_n} + \boldsymbol{\beta y_n}) = (\boldsymbol{z_1}, \boldsymbol{z_2}, \ldots, \boldsymbol{z_n})$

$$\begin{aligned} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\lambda_i z_i} &= \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\lambda_i}(\boldsymbol{\alpha x_i} + \boldsymbol{\beta y_i}) \\ &= \boldsymbol{\alpha} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\lambda_i x_i} + \boldsymbol{\beta} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\lambda_i y_i} \\ &= \mathbf{0} \end{aligned}$$

and so $\underline{\boldsymbol{z}} \in \boldsymbol{W}$, and thus as $\boldsymbol{W}$ is closed under addition and scalar multiplication it is a vector space □

iii) $\tilde{\boldsymbol{W}} = \{\underline{\boldsymbol{X}} : \underline{\boldsymbol{X}} \in \mathbb{R}, \underline{\boldsymbol{X}} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}), \sum_{i=1}^{n} \boldsymbol{\lambda_i x_i} = \boldsymbol{1}$ for some $\boldsymbol{\lambda_i}$.
Take $\underline{\boldsymbol{x}} = (\frac{\boldsymbol{1}}{\boldsymbol{\lambda_1}}, \boldsymbol{0}, \ldots, \boldsymbol{0}), \underline{\boldsymbol{y}} = (\boldsymbol{0}, \frac{\boldsymbol{1}}{\boldsymbol{\lambda_2}}, \ldots, \boldsymbol{0}) \in \tilde{\boldsymbol{W}}$
$\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} = (\frac{\boldsymbol{1}}{\boldsymbol{\lambda_1}}, \frac{\boldsymbol{1}}{\boldsymbol{\lambda_2}}, \boldsymbol{0}, \ldots, \boldsymbol{0}) = \underline{\boldsymbol{z}}$, but then $\sum_{i=1}^{n} \boldsymbol{\lambda_i z_i} = \boldsymbol{2} \Rightarrow \boldsymbol{z} \notin \tilde{\boldsymbol{W}}$
Therefore $\tilde{\boldsymbol{W}}$ is not a vector space as it is not closed under addition.

3.3 Spanning Sets, Dimension, Basis

Linearly Dependent - Given $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n} \in \boldsymbol{V}$, if $\exists$ scalars $\boldsymbol{lambda_i}$ not all zero such that $\sum_{i=1}^{n} \boldsymbol{\lambda_i} \underline{v_i} = \underline{\mathbf{0}}$ then $\{\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}\}$ is linearly dependent.

Linearly Independent - If a set is not linearly dependent that it is linearly independent

Note: By definition $\{\underline{\mathbf{0}}\}$ is linearly dependent.

Spanning Sets

Spanning Set - A set of vectors $\boldsymbol{S} = \{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_n}\}$ with the $\boldsymbol{a_i} \in \boldsymbol{V}$ is a spanning set of $\boldsymbol{V}$ if $\forall \underline{x} \in \boldsymbol{V}$ $\exists \boldsymbol{\lambda_1}, \boldsymbol{\lambda_2}, \ldots, \boldsymbol{\lambda_n}$ such that $\underline{x} = \sum_{i=1}^{n} \boldsymbol{\lambda_i} \boldsymbol{a_i}$

Note: The set of all linear combinations of $\underline{a_i}$'s i.e. $\boldsymbol{U} = \{\underline{x} \in \boldsymbol{V} : \underline{x} = \sum_{i=1}^{n} \boldsymbol{\lambda_i} \underline{a_i}$ for some $\boldsymbol{\lambda_i}\}$ is a subspace as it is closed under addition and scalar multiplication.
$\boldsymbol{U}$ is called the 'span' of $\boldsymbol{S}$ - or we say $\boldsymbol{S}$ 'spans' $\boldsymbol{U}$ or $\boldsymbol{U} = \text{span}(\boldsymbol{S})$.

Example: Consider $\mathbb{R}^{\mathbf{3}}$
Let $\underline{a_1} = (1,0,0), \underline{a_2} = (0,1,0), \underline{a_3} = (0,0,1), \underline{a_4} = (1,1,1)$ and $\underline{a_5} = (0,1,1)$
Consider $\underline{a_1} + \underline{a_2} + \underline{a_3} - \underline{a_4} + 0.\underline{a_5} = \underline{\mathbf{0}}$
$\therefore \{\underline{a_1}, \underline{a_2}, \underline{a_3}, \underline{a_4}, \underline{a_5}\}$ is linearly dependent and spans $\mathbb{R}^{\mathbf{3}}$
$\{\underline{a_1}, \underline{a_2}, \underline{a_3}\}$ is linearly independent and spans $\mathbb{R}^{\mathbf{3}}$
$\{\underline{a_1}, \underline{a_2}, \underline{a_4}\}$ is linearly independent and spans $\mathbb{R}^{\mathbf{3}}$
$\{\underline{a_1}, \underline{a_4}, \underline{a_5}\}$ is linearly dependent ($\underline{a_1} - \underline{a_4} + \underline{a_5} = \underline{\mathbf{0}}$) and doesn't span $\mathbb{R}^{\mathbf{3}}$

If $\boldsymbol{S} = \{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_n}\}$ spans $\boldsymbol{V}$
then $\boldsymbol{S'} = \{(\lambda_1 \underline{a_1} + \lambda \underline{a_2} + \cdots + \boldsymbol{\lambda_n} \underline{a_n}), \underline{a_2}, \underline{a_3}, \ldots, \underline{a_n}\}$ also spans $\boldsymbol{V}$ provided that $\boldsymbol{\lambda_1} \neq \mathbf{0}$

If $\boldsymbol{S}$ is linearly dependent then so is $\boldsymbol{S'}$
If $\boldsymbol{S}$ is linearly independent then so is $\boldsymbol{S'}$

Example: $\{\underline{a_1}, \underline{a_2}, \underline{a_3}\}$ is linearly independent and spans $\mathbb{R}^{\mathbf{3}}$, then equally $\{(\underline{a_1} + \underline{a_2} + \underline{a_3}), \underline{a_2}, \underline{a_3}\}$ is linearly independent and also spans $\mathbb{R}^{\mathbf{3}}$

If we have a set $\boldsymbol{S} = \{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_n}\}$ which is linearly dependent and spans a subspace $\boldsymbol{U}$, we can reduce it to a linearly independent set that also spans $\boldsymbol{U}$

Proof. $\boldsymbol{S}$ is linearly dependent $\Rightarrow$ $\exists \boldsymbol{\lambda_j}$ such that $\sum_{j=1}^{n} \boldsymbol{\lambda_j} \underline{a_j} = \mathbf{0}$, suppose $\boldsymbol{\lambda_n} \neq \mathbf{0}$ (rearrange if necessary). Then

$$\underline{a_n} = -\frac{\mathbf{1}}{\boldsymbol{\lambda_n}} \sum_{j=1}^{n-1} \boldsymbol{\lambda_j} \underline{a_j}$$

Hence span$\{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_{n-1}}\} = U$
If $\{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_{n-1}}\}$ is linearly independent then finish. If not, repeat removing one element at a time until $\{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_p}\}$ is linearly independent for some $p < n - 1$ □

Example: Take our $\{\underline{a_1}, \underline{a_2}, \underline{a_3}, \underline{a_4}, \underline{a_5}\}$ as defined earlier.

$\{\underline{a_1}, \underline{a_2}, \underline{a_3}, \underline{a_4}\}$ also spans $\mathbb{R}^3$ and is linearly dependent
$\left.\begin{array}{l} \{\underline{a_1}, \underline{a_2}, \underline{a_3}\} \\ \{\underline{a_1}, \underline{a_2}, \underline{a_4}\} \\ \{\underline{a_1}, \underline{a_3}, \underline{a_4}\} \\ \{\underline{a_2}, \underline{a_3}, \underline{a_4}\} \end{array}\right\}$ are all linearly independent and span $\mathbb{R}^3$

Basis - A basis of a vector space V is a linearly independent spanning set (of vectors in V)

Theorem 12. Every basis of a vector space V has the same number of elements

Proof. Suppose $\{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_n}\}$ and $\{\underline{b_1}, \underline{b_2}, \ldots, \underline{b_m}\}$ are bases for V, and WLOG $m \geq n$
Since $\{\underline{a_i} : i = 1, \ldots, n\}$ is a basis $\exists \lambda$ such that

$$\underline{b_1} = \sum_{j=1}^{n} \lambda_j \underline{a_j}$$

WLOG take $\lambda_1 \neq 0$, if not then reorder. Hence

$$\underline{a_1} = \frac{\underline{b_1}}{\lambda_1} - \frac{1}{\lambda_1} \sum_{j=2}^{n} \lambda_j \underline{a_j}$$

Hence $\{\underline{b_1}, \underline{a_2}, \underline{a_3}, \ldots, \underline{a_n}\}$ is a basis. Now try to express $\underline{b_2}$ in terms of this basis

$$\underline{b_2} = \mu_1 \underline{b_1} + \sum_{j=2}^{n} \mu_j \underline{a_j}$$

with the μ_j not all 0 because the set of $\underline{b_j}$ is linearly independent. Thus

$$\underline{a_2} = \frac{\underline{b_2}}{\mu_2} - \frac{\mu_1 \underline{b_1}}{\mu_2} - \sum_{j=3}^{n} \frac{\mu_j \underline{a_j}}{\underline{\mu_2}}$$

Hence $\{\underline{b_1}, \underline{b_2}, \underline{a_3}, \ldots, \underline{a_n}\}$ is a basis.

Repeating this we eventually get that $\{\underline{b_1}, \underline{b_2}, \ldots, \underline{b_n}\}$ is a basis, and hence spans V.
If $m > n$ $\underline{b_{n+1}}$ would be linearly independent of $\{\underline{b_1}, \underline{b_2}, \ldots, \underline{b_n}\}$, which contradicts the spanning of V.
Hence we deduce that $m = n$ □

Note: The same argument shows that no linearly independent set has more members than the basis for the vector space it resides in.

Dimension - The number $\boldsymbol{k}$ of vectors in a basis of $\boldsymbol{V}$ is the dimension of $\boldsymbol{V}$, denoted $\boldsymbol{k = \dim V}$

Examples:
i) Consider $\mathbb{R}^n$. The set $\{\underline{e_1}, \underline{e_2}, \ldots, \underline{e_n}\}$ is a basis, where $\underline{e_1} = (1, 0, \ldots, 0), \underline{e_2} = (0, 1, \ldots, 0), \ldots, \underline{e_n} = (0, 0, \ldots, 1)$
Hence the dimension of $\mathbb{R}^n$ is n, or $\dim \mathbb{R}^n = n$

ii) Take the subspace $U \subset \mathbb{R}^n$ containing the set of n-tuples of the form $(x, x, \ldots, x)$ for $x \in \mathbb{R}$
A basis for U is $\{(1, 1, \ldots, 1)\}$, hence $\dim u = 1$
The set of functions on $[0, 1]$ has infinite dimension, but we will only consider vector spaces of finite dimension for this course.

Theorem 13. If U is a proper subset of V then any basis for U can be extended to a basis for V

Proof. Consider $\mathbb{R}^3$ $U = \{\underline{X} = (x_1, x_2, x_3) \in \mathbb{R}^3$ such that $x_1 + x_2 = 0\}$
U is a valid subspace (can check from earlier results)
$x_1 + x_2 = 0 \Rightarrow x_2 = -x_1$ with x_3 arbitrary

$\forall \underline{X} \in U$ $\exists \lambda_1, \lambda_2$ such that $\underline{X} = \lambda_1(1, -1, 0) + \lambda_2(0, 0, 1)$
Thus we have a basis $\{(1, -1, 0), (0, 0, 1)\}$
Now choose any vector $\underline{Y}$ in $\mathbb{R}^3$ but not in U
Then $\{(1, -1, 0), (0, 0, 1)\underline{Y}\}$ is linearly independent and as it contains ($\dim \mathbb{R}^3$) vectors it is a basis for $\mathbb{R}^3$
Hence we have extended a basis for U into a basis for $\mathbb{R}^3$
In general we can extend this same method by continually adding linearly independent vectors until we have a basis for the whole vector space □

Components

Theorem 14. Let V be a vector space with $S = \{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_n}\}$ as a basis. Each $\underline{v} \in V$ can be written as a linear combination of basis vectors i.e. $\exists \lambda_i$ such that $\underline{x} = \sum_{i=1}^{n} \lambda_i \underline{a_i}$ with the λ_i unique

Proof. Certainly if S is a basis we can find a set of λ_i satsifying the above. So just remains to prove that they are unique.
Assume we can find λ_i, μ_i such that $\underline{v} = \sum_{i=1}^{n} \lambda_i \underline{a_i} = \sum_{i=1}^{n} \mu_i \underline{a_i}$
$\Rightarrow \sum_{i=1}^{n} (\lambda_i - \mu_i)\underline{a_i} = \underline{0}$
Since the $\underline{a_i}$ are linearly independent, this means that $\lambda_i - \mu_i = 0$ $\forall i$ i.e. $\lambda_i = \mu_i$ □

We call the $\boldsymbol{\lambda_i}$ the components of vector $\underline{\boldsymbol{v}}$ with respect to the basis $\boldsymbol{\{\underline{a_1}, \underline{a_2}, \ldots, \underline{a_n}\}}$.
We usually denote them just as $\boldsymbol{v_i}$
A given basis vector $\underline{\boldsymbol{v}}$ in a vector space of dimension $\boldsymbol{n}$ is specificed uniquely by $\boldsymbol{n}$ real numbers $\boldsymbol{(v_1, v_2, \ldots, v_n)}$ i.e. $\boldsymbol{\exists!}$correspondance between $\underline{\boldsymbol{v}}$ and $\boldsymbol{(v_1, v_2, \ldots, v_n) \in \mathbb{R}^n}$

$$\text{if } \boldsymbol{w} \underbrace{\leftrightarrow}_{\text{is related to}} \boldsymbol{(w_1, w_2, \ldots, w_n)}$$

$$\text{then } \boldsymbol{\alpha \underline{v} + \beta \underline{w} \leftrightarrow (\alpha v_1 + \beta w_1, \alpha v_2 + \beta w_2, \ldots, \alpha v_n + \beta w_n)}$$

3.4 Intersection and addition of vector spaces

Let $\boldsymbol{U}, \boldsymbol{W}$ be subspaces of a vector space $\boldsymbol{V}$
Define $\boldsymbol{U} \cap \boldsymbol{W}$ (the intersection of $\boldsymbol{U}$ and $\boldsymbol{W}$) consisting of vectors $\boldsymbol{v}$ such that $\boldsymbol{v} \in \boldsymbol{U}$ and $\boldsymbol{v} \in \boldsymbol{W}$
$\boldsymbol{U} \cap \boldsymbol{W}$ is never empty because both $\boldsymbol{U}, \boldsymbol{W}$ contain $\underline{\mathbf{0}}$

Theorem 15. $\boldsymbol{U} \cap \boldsymbol{W}$ is a subspace of $\boldsymbol{V}$

Proof. Take $\underline{\boldsymbol{x}}, \underline{\boldsymbol{y}} \in \boldsymbol{U} \cap \boldsymbol{W} \Rightarrow \underline{\boldsymbol{x}}, \underline{\boldsymbol{y}} \in \boldsymbol{U}$ and $\underline{\boldsymbol{x}}, \underline{\boldsymbol{y}} \in \boldsymbol{W}$
Hence $\boldsymbol{\lambda}\underline{\boldsymbol{x}} + \boldsymbol{\mu}\underline{\boldsymbol{y}} \in \boldsymbol{U}$ for arbitrary $\boldsymbol{\lambda}, \boldsymbol{\mu} \because \boldsymbol{U}$ is a subspace
$\boldsymbol{\lambda}\underline{\boldsymbol{x}} + \boldsymbol{\mu}\underline{\boldsymbol{y}} \in \boldsymbol{W}$ for arbitrary $\boldsymbol{\lambda}, \boldsymbol{\mu} \because \boldsymbol{W}$ is a subspace
$\therefore \boldsymbol{\lambda}\underline{\boldsymbol{x}} + \boldsymbol{\mu}\underline{\boldsymbol{y}} \in \boldsymbol{U} \cap \boldsymbol{W} \therefore \boldsymbol{U} \cap \boldsymbol{W}$ is closed under addition and scalar multiplication
$\therefore \boldsymbol{U} \cap \boldsymbol{W}$ is a subspace of $\boldsymbol{V}$ (and also of $\boldsymbol{U}$ and $\boldsymbol{W}$) □

Sum of Vector Spaces - $\boldsymbol{U} + \boldsymbol{W}$ is the set of vectors of the forms $\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}}$ where $\underline{\boldsymbol{x}} \in \boldsymbol{U}$ and $\underline{\boldsymbol{y}} \in \boldsymbol{W}$. It is called the sum of $\boldsymbol{U}$ and $\boldsymbol{W}$

Theorem 16. $\boldsymbol{U} + \boldsymbol{W}$ is a subspace of $\boldsymbol{V}$

Proof. Take $\underline{\boldsymbol{z_1}}, \underline{\boldsymbol{z_2}} \in \boldsymbol{U} + \boldsymbol{W}$ such that:
$\exists \underline{\boldsymbol{x_1}}, \underline{\boldsymbol{y_1}}$ with $\boldsymbol{z_1} = \underline{\boldsymbol{x_1}} + \underline{\boldsymbol{y_1}} \quad \underline{\boldsymbol{x_1}} \in \boldsymbol{U}, \underline{\boldsymbol{y_1}} \in \boldsymbol{W}$
$\exists \underline{\boldsymbol{x_2}}, \underline{\boldsymbol{y_2}}$ with $\boldsymbol{z_2} = \underline{\boldsymbol{x_2}} + \underline{\boldsymbol{y_2}} \quad \underline{\boldsymbol{x_2}} \in \boldsymbol{U}, \underline{\boldsymbol{y_2}} \in \boldsymbol{W}$

$$\begin{aligned}\boldsymbol{\lambda}\underline{\boldsymbol{z_1}} + \boldsymbol{\mu}\underline{\boldsymbol{z_2}} &= \boldsymbol{\lambda}\underline{\boldsymbol{x_1}} + \boldsymbol{\lambda}\underline{\boldsymbol{y_1}} + \boldsymbol{\mu}\underline{\boldsymbol{x_2}} + \boldsymbol{\mu}\underline{\boldsymbol{y_2}} \\ &= \underbrace{(\boldsymbol{\lambda}\underline{\boldsymbol{x_1}} + \boldsymbol{\mu}\underline{\boldsymbol{x_2}})}_{\in \boldsymbol{U}} + \underbrace{(\boldsymbol{\lambda}\underline{\boldsymbol{y_1}} + \boldsymbol{\mu}\underline{\boldsymbol{y_2}})}_{\in \boldsymbol{W}} \in \boldsymbol{U} + \boldsymbol{W}\end{aligned}$$

□

Note that $\boldsymbol{U} \subset \boldsymbol{U} + \boldsymbol{W}$ and $\boldsymbol{W} \subset \boldsymbol{U} + \boldsymbol{W}$
$\boldsymbol{U}$ and $\boldsymbol{W}$ are both subspaces and subsets of $\boldsymbol{U} + \boldsymbol{W}$

Now let $\{\underline{\boldsymbol{u_1}}, \underline{\boldsymbol{u_2}}, \ldots, \underline{\boldsymbol{u_l}}\}$ be a basis for $\boldsymbol{U}$
$\{\underline{\boldsymbol{w_1}}, \underline{\boldsymbol{w_2}}, \ldots, \underline{\boldsymbol{w_m}}\}$ be a basis for $\boldsymbol{W}$
Then $\boldsymbol{U} + \boldsymbol{W} = \text{span}\{\underline{\boldsymbol{u_1}}, \underline{\boldsymbol{u_2}}, \ldots, \underline{\boldsymbol{u_l}}, \underline{\boldsymbol{w_1}}, \underline{\boldsymbol{w_2}}, \ldots, \underline{\boldsymbol{w_m}}\}$
$\therefore \dim(\boldsymbol{U} + \boldsymbol{W}) \leq \dim \boldsymbol{U} + \dim \boldsymbol{W}$

If $\exists \underline{\boldsymbol{x}} \neq \underline{\mathbf{0}} \in \boldsymbol{U} \cap \boldsymbol{W}$
$\boldsymbol{x} \in \boldsymbol{U} \Rightarrow \exists \boldsymbol{\lambda_i}$ such that $\underline{\boldsymbol{x}} = \sum_{i=1}^{l} \boldsymbol{\lambda_i}\underline{\boldsymbol{u_i}}$
$\boldsymbol{x} \in \boldsymbol{W} \Rightarrow \exists \boldsymbol{\mu_j}$ such that $\underline{\boldsymbol{x}} = \sum_{j=1}^{m} \boldsymbol{\mu_j}\underline{\boldsymbol{w_j}}$

Note that as $\underline{\boldsymbol{x}} \neq \underline{\mathbf{0}}, \boldsymbol{\lambda_i}$ not all zero and $\boldsymbol{\mu_i}$ not all zero.

$$\therefore \sum_{i=1}^{l} \boldsymbol{\lambda_i}\underline{\boldsymbol{u_i}} + \sum_{j=1}^{m} (-\boldsymbol{\mu_j})\underline{\boldsymbol{w_j}} = \mathbf{0}$$

hence $\{\underline{u_1}, \underline{u_2}, \ldots, \underline{u_l}, \underline{w_1}, \underline{w_2}, \ldots, \underline{w_m}\}$ is linearly dependent, so to make it a basis we have to remove at least one element.
$\therefore \dim(U+W) < \dim U + \dim W$

Theorem 17. $\dim U + \dim W = \dim(U+W) + \dim(U \cap W)$

Proof. Outline only:
Reduce $\{\underline{u_1}, \ldots, \underline{u_l}, \underline{w_1}, \ldots, \underline{w_m}\}$ to a linearly independent set. Each reduction represents another basis vector in $U \cap W$, and thus the number of reductions will be equal to $\dim(U \cap W)$ □

Example 1: Let $U = \{\underline{X} \in \mathbb{R}^4 : x_1 = 0\}$
Let $W = \{\underline{X} \in \mathbb{R}^4 : x_1 + 2x_2 = 0\}$

$U = \text{span}\{(0,1,0,0), (0,0,1,0), (0,0,0,1)\} \quad \therefore \dim U = 3$
$W = \text{span}\{(-2,1,0,0), (0,0,1,0), (0,0,0,1)\} \quad \therefore \dim W = 3$

$U \cap W = \{\underline{X} \in \mathbb{R}^4 : x_1 = 0, x_1 + 2x_2 = 0\} \Rightarrow x_1, x_2 = 0$
$\therefore U \cap W = \text{span}\{(0,0,1,0), (0,0,0,1)\} \quad \therefore \dim(U \cap W) = 2$

$U + W = \{\underline{x} + \underline{y} : \underline{x} \in U, \underline{y} \in W\}$

$$\begin{aligned} U + W &= \text{span}\{(0,1,0,0), (0,0,1,0), (0,0,0,1), (-2,1,0,0), \cancel{(0,0,1,0)}, \cancel{(0,0,0,1)}\} \\ &= \text{span}\{(0,1,0,0), (0,0,1,0), (0,0,0,1), (-2,1,0,0)\} \end{aligned}$$

Which is linearly independent and thus a basis. $\therefore \dim(U+W) = 4$
Thus for this example

$$\dim(U+W) + \dim(U \cap W) = \dim U + \dim W$$

Example 2: The vector space V of functions $f : \{1,2,3\} \to \mathbb{R}^3$ (Note that $\dim V = 3$)

We define operations of addition and scalar multiplication in an obvious way.
$U = \{f \in V : f(1) = 0\} \ \dim U = 2$
$W = \{f \in V : f(2) = 0\} \ \dim W = 2$

$U \cap W = \{f \in V : f(1) = f(2) = 0\} \ \dim(U \cap W) = 1$
$U + W = V \Rightarrow \dim(U+W) = 3$
So again here

$$\dim(U+W) + \dim(U \cap W) = \dim U + \dim W$$

3.5 Scalar product in $\mathbb{R}^n$

Define scalar product of $\underline{X}, \underline{Y} \in \mathbb{R}^n$ to be

$$\underline{X}.\underline{Y} = \sum_{i=1}^{n} x_i y_i \in \mathbb{R}$$

With the following properties:

$$(\underline{X} + \underline{Y}).\underline{Z} = \underline{X}.\underline{Z} + \underline{Y}.\underline{Z}$$

$$(\lambda\underline{X}).\underline{Y} = \lambda(\underline{X}.\underline{Y}) \quad \lambda \in \mathbb{R}$$

Define the length, or norm, of $\underline{X} \in \mathbb{R}^n$ by:

$$||\underline{X}|| = \sqrt{\underline{X}.\underline{X}} \in \mathbb{R}$$

$$= \sqrt{\sum_{i=1}^{n} x_i^2}$$

Properties:
$||\underline{X}|| = 0 \Rightarrow \underline{X} = \underline{0} = (0, 0, 0, \ldots)$
$||-\underline{X}|| = ||\underline{X}||$
$||\lambda\underline{X}|| = |\lambda|.||\underline{X}||$

In 2-D, 3-D, $||\underline{X}|| = |\underline{X}|$ as defined earlier.
For other vector spaces e..g of functions, we can find suitable definitions of scalar products and norm

Examples:
i) (hyper-sphere in $\mathbb{R}^n$, called $\mathbf{\Sigma}$, is given by
$\{\underline{X} \in \mathbb{R}^n : ||\underline{X} - \underline{A}|| = r > 0, \underline{A} \in \mathbb{R}^n\}$, centre $\underline{A}$, radius r

ii) hyperplane in $\mathbb{R}^n$, called $\mathbf{\Pi}$ given by
$\{\underline{X} \in \mathbb{R}^n : (\underline{X} - \underline{B}).\underline{N} = 0, \underline{B} \in \mathbb{R}^n, \underline{N} \in \mathbb{R}^n, ||\underline{N}|| = 1\}$
A hyperplane through B normal to N

$\mathbf{\Sigma}$ is never a subspace of $\mathbb{R}^n$ but $\mathbf{\Pi}$ is a subspave iff $\underline{B}.\underline{N} = \underline{0}$ i.e. it passes through the origin. It has dimension $(n-1)$ - to prove this just show that $\mathbf{\Pi}$ is closed under addition and scalar multiplication.

Theorem 18. If $\underline{X}, \underline{Y} \in \mathbb{R}^n$ then $|\underline{X}.\underline{Y}| \leq (||\underline{X}||.||\underline{Y}||)$

Proof. Assume $\underline{Y} \neq \underline{0}$ (otherwise trivial)

$$\begin{aligned} ||\underline{X} + t\underline{Y}||^2 &= (\underline{X} + t\underline{Y}.(\underline{X} + t\underline{Y}) \\ &= \underline{X}.\underline{X} + 2t\underline{X}.\underline{Y} + t^2\underline{X}.\underline{Y} \\ &= ||\underline{X}||^2 + t^2||\underline{Y}||^2 + 2t\underline{X}.\underline{Y} \geq 0 \ \forall t \in \mathbb{R} \end{aligned}$$

Consider this as a quadratic, $\boldsymbol{f(t) = 0}$ has at most 1 root

$$\therefore \mathbf{(2(\underline{X}.\underline{Y}))^2 \leq 4||\underline{X}||^2||\underline{Y}||^2}$$
$$\therefore \mathbf{|\underline{X}.\underline{Y}| \leq ||\underline{X}||.||\underline{Y}||}$$

Explicit from:

$$|\sum_{i=1}^{n} \boldsymbol{x_i y_i}| \leq \sqrt{\sum_{i=1}^{n} \boldsymbol{x_i^2}} \sqrt{\sum_{i=1}^{n} \boldsymbol{y_i^2}}$$

□

Corollary 19 (Triangle inequality)**.** For all $\underline{\boldsymbol{X}}, \underline{\boldsymbol{Y}} \in \mathbb{R}^{\boldsymbol{n}}$

$$||\underline{\boldsymbol{X}} + \underline{\boldsymbol{Y}}|| \leq ||\underline{\boldsymbol{X}}|| + ||\underline{\boldsymbol{Y}}||$$

Proof.

$$\begin{aligned} ||\underline{\boldsymbol{X}} + \underline{\boldsymbol{Y}}||^2 &= (\underline{\boldsymbol{X}} + \underline{\boldsymbol{Y}}).(\underline{\boldsymbol{X}} + \underline{\boldsymbol{Y}}) \\ &= \underbrace{\underline{\boldsymbol{X}}.\underline{\boldsymbol{X}}}_{||\underline{\boldsymbol{X}}||^2} + \underbrace{\mathbf{2}\underline{\boldsymbol{X}}.\underline{\boldsymbol{Y}}}_{\leq \mathbf{2}||\underline{\boldsymbol{X}}||.||\underline{\boldsymbol{Y}}||} + \underbrace{\underline{\boldsymbol{Y}}.\underline{\boldsymbol{Y}}}_{||\underline{\boldsymbol{Y}}||^2} \qquad \leq (||\underline{\boldsymbol{X}}|| + ||\underline{\boldsymbol{Y}}||)^2 \end{aligned}$$

Hence the result □

Similarly $||\underline{\boldsymbol{X}} - \underline{\boldsymbol{Y}}|| \geq \big|\, ||\underline{\boldsymbol{X}}|| - ||\underline{\boldsymbol{Y}}|| \,\big|$

Final note on $\mathbb{R}^{\boldsymbol{n}}$
We've written $\underline{\boldsymbol{X}} = \boldsymbol{(x_1, x_2, \ldots, x_n)} \in \mathbb{R}^{\boldsymbol{n}}$, where $\boldsymbol{(x_1, x_2, \ldots, x_n)}$ are the components of $\boldsymbol{x}$ with respect to a basis:
$\underline{\boldsymbol{e_1}} = \boldsymbol{(1, 0, \ldots, 0)}$
$\underline{\boldsymbol{(e_2)}} = \boldsymbol{(0, 1, \ldots, 0)}$
$\vdots$
$\underline{\boldsymbol{e_n}} = \boldsymbol{(0, 0, \ldots, 1)}$

$\therefore \underline{\boldsymbol{X}} = \boldsymbol{x_1}\underline{\boldsymbol{e_1}} + \boldsymbol{x_2}\underline{\boldsymbol{e_2}} + \cdots + \underline{\boldsymbol{e_n}}$
But with a different basis $\underline{\boldsymbol{X}}$ would have different components

4 Linear Maps to Matrices

4.1 Introduction

Let $\boldsymbol{A}, \boldsymbol{B}$ be sets, a map $\boldsymbol{f}$ from $\boldsymbol{A}$ into $\boldsymbol{B}$ is a rule which assigns a member $\boldsymbol{x'} = \boldsymbol{f}(\boldsymbol{x})$ to each member of $\boldsymbol{A}$. $\boldsymbol{x'}$ is unique for any given $\boldsymbol{x}$.
We write:

$$\boldsymbol{f} : \boldsymbol{A} \to \boldsymbol{B} \qquad \boldsymbol{x'} = \boldsymbol{f}(\boldsymbol{x})$$

$\boldsymbol{A}$ is called the domain of $\boldsymbol{f}$
$\boldsymbol{B}$ is called the co-domian or range of $\boldsymbol{f}$
$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x'}$ is the image of $\boldsymbol{x}$
$\boldsymbol{f}(\boldsymbol{A})$ is the image of $\boldsymbol{A}$ (under $\boldsymbol{f}$), $\boldsymbol{f}(\boldsymbol{A}) = \{\boldsymbol{x'} : \boldsymbol{x'} = \boldsymbol{f}(\boldsymbol{x}) \; \forall \boldsymbol{x} \in \boldsymbol{A}, \boldsymbol{x'} \in \boldsymbol{B}\}$
Note that $\boldsymbol{f}(\boldsymbol{A}) \subset \boldsymbol{B}$, but it is not required that $\boldsymbol{f}(\boldsymbol{A}) = \boldsymbol{B}$

Examples:
Möbius maps of the complex plane
Translation
Inversion with respect to a sphere

We will look at maps from a vector space $\boldsymbol{V}$ to another vector space $\boldsymbol{W}$, focusing on $\boldsymbol{V} = \mathbb{R}^{\boldsymbol{n}}, \boldsymbol{W} = \mathbb{R}^{\boldsymbol{n}}$
Recall that this is a unique correspondance between any members of a vector space $\boldsymbol{V}$ of dimension $\boldsymbol{n}$ and a member of $\mathbb{R}^{\boldsymbol{n}}$

Linear Maps - Let $\boldsymbol{V}, \boldsymbol{W}$ be vector spaces over $\mathbb{R}$. The map $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$ is a linear map (or linear transformation) if:

i) $\boldsymbol{T}(\underline{\boldsymbol{a}} + \underline{\boldsymbol{b}}) = \boldsymbol{T}(\underline{\boldsymbol{a}}) + \boldsymbol{T}(\underline{\boldsymbol{b}}) \quad \forall \underline{\boldsymbol{a}}, \underline{\boldsymbol{b}} \in \boldsymbol{V}$
ii) $\boldsymbol{T}(\boldsymbol{\lambda}\underline{\boldsymbol{a}}) = \boldsymbol{\lambda}\boldsymbol{T}(\underline{\boldsymbol{a}}) \quad \forall \underline{\boldsymbol{a}} \in \boldsymbol{V}, \boldsymbol{\lambda} \in \mathbb{R}$

equivalently $\boldsymbol{T}(\boldsymbol{\lambda}\underline{\boldsymbol{a}} + \boldsymbol{\mu}\underline{\boldsymbol{b}}) = \boldsymbol{\lambda}\boldsymbol{T}(\underline{\boldsymbol{a}}) + \boldsymbol{\mu}\boldsymbol{T}(\underline{\boldsymbol{b}}) \; \forall \boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}, \underline{\boldsymbol{a}}, \underline{\boldsymbol{b}} \in \boldsymbol{V}$

Thus $\boldsymbol{T}(\boldsymbol{V})$ is a subspace of $\boldsymbol{W}$, since
$\boldsymbol{T}(\underline{\boldsymbol{a}}), \boldsymbol{T}(\underline{\boldsymbol{b}}) \in \boldsymbol{T}(\boldsymbol{V})$
$\boldsymbol{\lambda}\boldsymbol{T}(\underline{\boldsymbol{a}}) + \boldsymbol{\mu}\boldsymbol{T}(\underline{\boldsymbol{b}}) = \boldsymbol{T}(\boldsymbol{\lambda}\underline{\boldsymbol{a}} + \boldsymbol{\mu}\underline{\boldsymbol{b}}) \in \boldsymbol{T}(\boldsymbol{V}) \quad \because \underline{\boldsymbol{a}}, \underline{\boldsymbol{b}} \in \boldsymbol{V}, \boldsymbol{\lambda} \in \mathbb{R}$

If $\underline{\boldsymbol{b}} = \underline{\boldsymbol{0}} \in \boldsymbol{V}$ then $\boldsymbol{T}(\underline{\boldsymbol{a}}) + \boldsymbol{T}(\underline{\boldsymbol{b}}) = \boldsymbol{T}(\underline{\boldsymbol{a}} + \underline{\boldsymbol{b}}) = \boldsymbol{T}(\underline{\boldsymbol{a}}) \; \forall \underline{\boldsymbol{a}} \in \boldsymbol{V}$
Hence $\boldsymbol{T}(\underline{\boldsymbol{b}}) = \underline{\boldsymbol{0}} \in \boldsymbol{W}$
$\therefore \boldsymbol{T}(\underline{\boldsymbol{0}}_{\boldsymbol{v}}) = \underline{\boldsymbol{0}}_{\boldsymbol{W}}$

$\boldsymbol{T}(\underline{\boldsymbol{b}}) = \underline{\boldsymbol{0}} \in \boldsymbol{W} \not\Rightarrow \underline{\boldsymbol{b}} = \underline{\boldsymbol{0}} \in \boldsymbol{V}$ (we could have multiple maps to it)

Examples:
i) Translations in $\mathbb{R}^{\boldsymbol{3}}$ (isometries)

$$\underline{\boldsymbol{x}} \mapsto \underline{\boldsymbol{x'}} = \underline{\boldsymbol{x}} + \underline{\boldsymbol{a}} = \boldsymbol{T}(\underline{\boldsymbol{x}}) \; (\underline{\boldsymbol{a}} \in \mathbb{R}^{\boldsymbol{3}}, \underline{\boldsymbol{a}} \neq \underline{\boldsymbol{0}})$$

This is not a linear map:

$$\boldsymbol{T}(\underline{\boldsymbol{x}}) + \boldsymbol{T}(\underline{\boldsymbol{y}}) = \underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} + \mathbf{2}\underline{\boldsymbol{a}}$$
$$\boldsymbol{T}(\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}}) = \underline{\boldsymbol{x}} + \underline{\boldsymbol{y}} + \underline{\boldsymbol{a}}$$

Which are not the same as $\underline{\boldsymbol{a}} \neq \underline{\mathbf{0}}$

ii) Reflection in the plane $\boldsymbol{\Pi} = \{\underline{\boldsymbol{x}} : \underline{\boldsymbol{x}}.\underline{\boldsymbol{n}} = \mathbf{0} \quad \underline{\boldsymbol{x}}, \underline{\boldsymbol{n}} \in \mathbb{R}^{\mathbf{3}} \, |\underline{\boldsymbol{n}}| = \mathbf{1}\}$

$$\boldsymbol{R}_{\boldsymbol{\Pi}} : \mathbb{R}^{\mathbf{3}} \to \mathbb{R}^{\mathbf{3}} \qquad \underline{\boldsymbol{x}} \mapsto \underline{\boldsymbol{x}}' = \boldsymbol{R}_{\boldsymbol{\Pi}}(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}} - \mathbf{2}(\underline{\boldsymbol{x}}.\underline{\boldsymbol{n}})\underline{\boldsymbol{n}}$$

$\boldsymbol{R}_{\boldsymbol{\Pi}}$ is a linear map from $\mathbb{R}^{\mathbf{3}}$ to $\mathbb{R}^{\mathbf{3}}$ $\because \forall \underline{\boldsymbol{x_1}}, \underline{\boldsymbol{x_2}} \in \mathbb{R}^{\mathbf{3}}$

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{\Pi}}(\boldsymbol{\lambda}\underline{\boldsymbol{x_1}} + \boldsymbol{\mu}\underline{\boldsymbol{x_2}}) &= \boldsymbol{\lambda}\underline{\boldsymbol{x_1}} + \boldsymbol{\mu}\underline{\boldsymbol{x_2}} - \mathbf{2}((\boldsymbol{\lambda}\underline{\boldsymbol{x_1}} + \boldsymbol{\mu}\underline{\boldsymbol{x_2}}).\underline{\boldsymbol{n}})\underline{\boldsymbol{n}} \\ &= \boldsymbol{\lambda}(\underline{\boldsymbol{x_1}} - \mathbf{2}(\underline{\boldsymbol{x_1}}.\underline{\boldsymbol{n}})\underline{\boldsymbol{n}}) + \boldsymbol{\mu}(\underline{\boldsymbol{x_2}} - \mathbf{2}(\underline{\boldsymbol{x_2}}.\underline{\boldsymbol{n}})\underline{\boldsymbol{n}}) \\ &= \boldsymbol{\lambda}\boldsymbol{R}_{\boldsymbol{\Pi}}(\underline{\boldsymbol{x_1}}) + \boldsymbol{\mu}\boldsymbol{R}_{\boldsymbol{\Pi}}(\underline{\boldsymbol{x_2}}) \; \forall \boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}\end{aligned}$$

hence the result

iii) Projection onto a plane (not an isometry)

$$\boldsymbol{P} : \mathbb{R}^{\mathbf{3}} \to \mathbb{R}^{\mathbf{3}} \qquad \underline{\boldsymbol{x}} \mapsto \underline{\boldsymbol{x}}' = \boldsymbol{P}(\underline{\boldsymbol{x}}) = (\underline{\boldsymbol{x}}.\underline{\boldsymbol{n}})\underline{\boldsymbol{n}}$$

Where $\underline{\boldsymbol{n}}$ is a unit vector. To check that this is linear we must check that $\boldsymbol{P}(\boldsymbol{\lambda}\underline{\boldsymbol{x_1}} + \boldsymbol{\mu}\underline{\boldsymbol{x_2}}) = \boldsymbol{\lambda}\boldsymbol{P}(\underline{\boldsymbol{x_1}} + \boldsymbol{\mu}\boldsymbol{P}(\underline{\boldsymbol{x_2}}))$ - easily verified.
Note: $\boldsymbol{P}(\mathbb{R}^{\mathbf{3}}) = \{\underline{\boldsymbol{x}} \in \mathbb{R}^{\mathbf{3}} : \underline{\boldsymbol{x}} = \boldsymbol{\lambda}\underline{\boldsymbol{n}}$ for some $\boldsymbol{\lambda} \in \mathbb{R}\}$, a 1-D subspace = span$(\underline{\boldsymbol{n}})$

iv)$\boldsymbol{T} : \mathbb{R}^{\mathbf{3}} \to \mathbb{R}^{\mathbf{2}}$ (domain is different from range)
$(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \mapsto (\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{T}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x} + \mathbf{2}\boldsymbol{y}, \mathbf{2}\boldsymbol{x} - \boldsymbol{z})$

Easily check that this is a linear map

Take our standard basis for $\mathbb{R}^{\mathbf{3}}$
$\underline{\boldsymbol{e_1}} = (\mathbf{1}, \mathbf{0}, \mathbf{0}) \qquad \boldsymbol{T}(\underline{\boldsymbol{e_1}}) = (\mathbf{1}, \mathbf{2})$
$\underline{\boldsymbol{e_2}} = (\mathbf{0}, \mathbf{1}, \mathbf{0}) \qquad \boldsymbol{T}(\underline{\boldsymbol{e_2}}) = (\mathbf{1}, \mathbf{0}) \qquad \mathbb{R}^{\mathbf{3}} = \text{span}\{\underline{\boldsymbol{e_1}}, \underline{\boldsymbol{e_2}}, \underline{\boldsymbol{e_3}}\}$
$\underline{\boldsymbol{e_3}} = (\mathbf{0}, \mathbf{0}, \mathbf{1}) \qquad \boldsymbol{T}(\underline{\boldsymbol{e_3}}) = (\mathbf{0}, -\mathbf{1})$
$\therefore \boldsymbol{T}(\mathbb{R}^{\mathbf{3}}) = \text{span}\{\boldsymbol{T}(\underline{\boldsymbol{e_1}}), \boldsymbol{T}(\underline{\boldsymbol{e_2}}), \boldsymbol{T}(\underline{\boldsymbol{e_3}})\} = \mathbb{R}^{\mathbf{2}}$

Note that $\boldsymbol{T}(\underline{\boldsymbol{e_1}}) - \boldsymbol{T}(\underline{\boldsymbol{e_2}}) + \mathbf{2}\boldsymbol{T}(\underline{\boldsymbol{e_3}}) = \underline{\mathbf{0}} \in \mathbb{R}^{\mathbf{2}}$
$\Rightarrow \boldsymbol{T}(\underline{\boldsymbol{e_1}} - \underline{\boldsymbol{e_2}} + \mathbf{2}\underline{\boldsymbol{e_3}}) = \underline{\mathbf{0}} \in \mathbb{R}^{\mathbf{2}}$
Consider the subspace of $\mathbb{R}^{\mathbf{3}} : \{\underline{\boldsymbol{x}} \in \mathbb{R}^{\mathbf{3}} : \underline{\boldsymbol{x}} = \boldsymbol{\lambda}(\underline{\boldsymbol{e_1}} - \underline{\boldsymbol{e_2}} + \mathbf{2}\underline{\boldsymbol{e_3}}), \boldsymbol{\lambda} \in \mathbb{R}\} =$ span$\{\underline{\boldsymbol{e_1}} - \underline{\boldsymbol{e_2}} + \mathbf{2}\underline{\boldsymbol{e_3}}\}$
This whole 1-D subspace maps into $\underline{\mathbf{0}} \in \mathbb{R}^{\mathbf{2}}$

v) $\boldsymbol{T : \mathbb{R}^2 \to \mathbb{R}^4}$ (so the range has a larger dimension that the domain)
$\boldsymbol{(x, y) \mapsto (s, t, u, v) = T(x, y) = (x + y, x, y - 3x, y)}$

Again easily verified that this is a linear map

$$\left.\begin{array}{l} \boldsymbol{T(\underline{e_1}) = T(1, 0) = (1, 1, -3, 0)} \\ \boldsymbol{T(\underline{e_2}) = T(0, 1) = (1, 0, 1, 1)} \end{array}\right\} \text{Linearly independent}$$

$\therefore \boldsymbol{T(\mathbb{R}^3)} = \text{span}\boldsymbol{\{(1, 1, -3, 0), (1, 0, 1, 1)\}}$, a 2-D subspace of $\mathbb{R}^4$

4.2 Rank, Nullity and Kernel

Take $\boldsymbol{T : V \to W}$ a linear map

$\boldsymbol{T(V)}$ is the image of $\boldsymbol{V}$ under $\boldsymbol{T}$ and is a subspace of $\boldsymbol{W}$

Rank - The rank is the dimension of the image of $\boldsymbol{V}$
$\boldsymbol{r(T) = \dim T(V)}$

Kernel - The kernel is the set of elements of $\boldsymbol{V}$ which map to the zero vector in $\boldsymbol{W}$
$\boldsymbol{\ker(T) - \{\underline{a} \in V : T(\underline{a}) = \underline{0} \in W\}}$
$\boldsymbol{\ker(T)}$ is non-empty as it contains at least $\boldsymbol{\underline{0} \in V}$ - and because it is closed with respect to addition and scalar multiplication it is a subspace of $\boldsymbol{V}$

Nullity - The nullity of $\boldsymbol{T}$ is the dimension of the kernel
$\boldsymbol{n(T) = \dim \ker(T)}$
Note that if the kernel only has one elements, $\underline{\mathbf{0}}$, then it is said to have 0 dimension

Examples:
i) $\boldsymbol{T : \mathbb{R}^2 \to \mathbb{R}^3}$
$\boldsymbol{(x, y) \mapsto T(x, y) = (2x + 3y, 4x + 6y, -2x - 3y) = (2x + 3y)(1, 2, -1)}$
$\boldsymbol{T(\mathbb{R}^2)}$ is a line $\boldsymbol{\{\underline{x} \in \mathbb{R}^2 : \underline{x} = \lambda(1, 2, 01), \lambda \in \mathbb{R}\}}$
$\boldsymbol{\therefore r(T) = 1}$
$\boldsymbol{\ker(T) = \{\underline{x} \in \mathbb{R}^2 : \underline{x} = (x, y), 2x + 3y = \underline{0}\}}$
$\boldsymbol{\Rightarrow \underline{x} = s(-3, 2), s \in \mathbb{R}}$ a line in $\boldsymbol{\mathbb{R}^2}$
$\boldsymbol{\therefore n(T) = 1}$
Note that $\boldsymbol{r(T) + n(T) = \dim(\mathbb{R}^2)}$
ii) Projection onto a line:

$$\boldsymbol{P : \mathbb{R}^3 \to \mathbb{R}^3 \qquad \underline{x} \mapsto \underline{x'} = (\underline{x}.\underline{n})\underline{n}, |\underline{n}| = 1}$$

$\boldsymbol{P(\mathbb{R}^3) = \{\underline{x} \in \mathbb{R}^3 : \underline{x} = \lambda\underline{n}, \lambda \in \mathbb{R}\}}$, so $\boldsymbol{r(P) = 1}$
$\boldsymbol{\ker(P) = \{\underline{x} \in \mathbb{R}^3 : (\underline{x}.\underline{n})\underline{n} = \underline{0}\} \Rightarrow (\underline{x}.\underline{n}) = 0}$
which is a plane through the origin $\perp$ to $\boldsymbol{\underline{n}}$.
$\boldsymbol{\therefore \dim(\ker(P)) = n(P) = 2}$
Note again that $\boldsymbol{r(P) + n(P) = 3 = \dim(\mathbb{R}^3)}$

4.3 Composition of linear maps

Let $\boldsymbol{S, T}$ be two linear maps $\begin{array}{cc} \boldsymbol{S: U \to V} & \boldsymbol{T: V \to W} \\ \boldsymbol{\underline{u} \mapsto \underline{v} = S(\underline{u})} & \boldsymbol{\underline{v} \mapsto \underline{w} = T(\underline{v})} \end{array}$

Composite map - The composite or product map $\boldsymbol{TS}$ is given by:

$$\boldsymbol{TS: U \to W} \qquad \boldsymbol{\underline{u} \mapsto \underline{w} = T(S(\underline{u}))}$$

> Note that to define $\boldsymbol{TS}$ we need the domain of $\boldsymbol{T}$ to include the image of $\boldsymbol{S}$

Examples:
i) $\boldsymbol{R_\Pi : \mathbb{R}^3 \to \mathbb{R}^3}$ (reflection in the plane $\boldsymbol{\Pi}$.
The range is contained in the domain, so can form the composite $\boldsymbol{R_\Pi R_\Pi = R_\Pi^2 = I}$ the identity map

ii) $\boldsymbol{P : \mathbb{R}^3 \to \mathbb{R}^3}$ (projection onto a line)
As before, can form the composite, $\boldsymbol{PP = P}$

iii) Take the pair of maps

$$\begin{array}{cc} \boldsymbol{S : \mathbb{R}^2 \to \mathbb{R}^3} & \boldsymbol{T : \mathbb{R}^3 \to \mathbb{R}} \\ \boldsymbol{(u, v) \mapsto S(u, v) = (-v, u, u + v)} & \boldsymbol{(x, y, z) \mapsto T(x, y, z) = x + y + z} \end{array}$$

$\boldsymbol{TS}$ is thus defined as follows:

$$\boldsymbol{TS : \mathbb{R}^2 \to \mathbb{R}} \qquad \boldsymbol{(u, v) \mapsto T(S(u, v)) = 2u}$$

But $\boldsymbol{ST}$ is not defined

4.4 Bases and Matrix description of linear maps

Let $\underline{e_i}, i = 1, 2, 3$ be the standard basis for $\mathbb{R}^3$
Every $\underline{a} \in \mathbb{R}^3$ has a unique expression as some linear combination of those vectors, $\underline{a} = \sum_{i=1}^{3} a_i \underline{e_i} \quad q_i \in \mathbb{R}$

Consider a linear map $M : \mathbb{R}^3 \to \mathbb{R}^3$
$M(\underline{a}) = \sum_{i=1}^{3} a_i M(\underline{e_i})$ since M is linear

Consider $M(\underline{e_j}) = \underline{e'_j} = \sum_{i=1}^{3} M_{ij} \underline{e_i}$ where $M_{ij} \in \mathbb{R}$
M_{ij} is the i^{th} component of $\underline{e'_j} = M(\underline{e_j})$ with respect to the standard basis
$M_{ij} = (\underline{e'_j})_i$ which for an orthonormal basis is given by $\underline{e_i}.\underline{e'_j}$

$\therefore$ for a general $\underline{a} \in \mathbb{R}^3$

$$\begin{aligned} \underline{a} \mapsto \underline{a'} = M(\underline{a}) &= \sum_{i=1}^{3} a_j M(\underline{e_j}) \\ &= \sum_{j=1}^{3} a_j \sum_{i=1}^{3} M_{ij} \underline{e_i} \end{aligned}$$

Thus $\underline{a'} = \sum_{i=1}^{3} a'_i \underline{e_i}$ with $a'_i = \sum_{j=1}^{3} M_{ij} a_j$ a mapping $M : \mathbb{R}^3 \to \mathbb{R}^3$, once the basis vectors are chosen, is completely specified by M_{ij} (which has 9 quantities for $i = 1, 2, 3, j = 1, 2, 3$)

Under the summation convention $a'_i = M_{ij} a_j$

Or explicitly:
$a'_1 = M_{11} a_1 + M_{12} a_2 + M_{13} a_3$
$a'_2 = M_{21} a_1 + M_{22} a_2 + M_{23} a_3$
$a'_3 = M_{31} a_1 + M_{32} a_2 + M_{33} a_3$

or

$$\underbrace{\begin{pmatrix} a'_1 \\ a'_2 \\ a'_3 \end{pmatrix}}_{\underline{a'}} = \underbrace{\begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}}_{\underline{a}} \quad \underline{a'} = M\underline{a} \text{ in matrix algebra}$$

i^{th} row of $\underline{a'} = \underbrace{(i^{textth} \text{ row of } M)}_{\text{row vector}} . \underline{a}$

$\therefore \underline{a'} = M\underline{a}$ as $a'_i = M_{ij} a_j$

We call the M_{ij} the elements of M (and sometimes write $M = \{M_{ij}\}$).

Lebelling convention is that the first number is the row number i and the second it the column number j

$\underline{e'_j} = \sum_{i=1}^{3} M_{ij}\underline{e_i}$ so the j^{th} column contains components of $\underline{e'_j}$

Can write $M = (\underbrace{\underline{e'_1}, \underline{e'_2}, \underline{e'_3}}_{\text{column vectors}})$

Examples:

i) Reflection in the plane $\mathbf{\Pi}$
Recall this is given by the map $\underline{x} \mapsto \underline{x} - 2(\underline{x}.\underline{n})n$ where $|\underline{n}| = 1$
Consider the effect of $R_{\mathbf{\Pi}}$ on each member of the standard basis:

$$\begin{aligned}
\underline{e_1} : R_{\mathbf{\Pi}}(\underline{e_1}) &= \underline{e_1} - 2\underbrace{(\underline{e_1}.\underline{n})}_{n_1}\underline{n} \\
&= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 2n_1 \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \\
&= \begin{pmatrix} 1 - 2n_1^2 \\ -2n_1n_2 \\ -2n_1n_2 \end{pmatrix}
\end{aligned}$$

Which is the first column of the matrix H representing $R_{\mathbf{\Pi}}$. Similarly from considering $R_{\mathbf{\Pi}}(\underline{e_2})$ and $R_{\mathbf{\Pi}}(\underline{e_3})$:

$$H = \begin{pmatrix} 1 - 2n_1^2 & -2n_1n_2 & -2n_1n_3 \\ -2n_1n_2 & 1 - 2n_2^2 & -2n_2n_3 \\ -2n_1n_3 & -2n_2n_3 & 1 - 2n_3^2 \end{pmatrix}$$

For a different approach we can use suffix notation:

$$\begin{aligned}
(R_{\mathbf{\Pi}}(\underline{x}))_i = x'_i &= \underbrace{x_i}_{\delta_{ij}} - 2(x_jn_j)n_i \\
&= (\delta_{ij} - 2n_in_j)x_j \\
&= H_{ij}x_j
\end{aligned}$$

Hence $(H)_{ij} = H_{ij} = \delta_{ij} - 2n_in_j$

ii) $P_{\underline{b}} : \mathbb{R}^3 \to \mathbb{R}^3 \qquad \underline{x} \mapsto \underline{b} \times \underline{x}$

$$\begin{aligned}\underline{e_1'} &= \underline{b} \times \underline{e_1} \\ &= \underline{b} \times (1, 00) \\ &= \underbrace{(0, b_3, -b_2)}_{1^{\text{st}} \text{ column of } P_{\underline{b}}}\end{aligned}$$

Hence the matrix $\boldsymbol{G}$ is given by

$$\boldsymbol{G} = \begin{pmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{pmatrix}$$

Also $x_i' = \underbrace{\epsilon_{ijk} b_j}_{ij^{\text{th}} \text{ element of } \boldsymbol{G}} x_k$

$\therefore G_{ik} = \epsilon_{ijk} b_j$

iii) Rotation by $\boldsymbol{\theta}$ about $\boldsymbol{z}$ axis (no effect on 3-D direction, only on x-y plane)

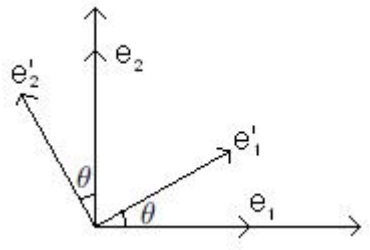

Figure 36: Rotation about the z-axis

$$\begin{aligned}\underline{e_1'} &= \cos\theta \underline{e_1} + \sin\theta \underline{e_2} + 0.\underline{e_3} \\ \underline{e_2'} &= -\sin\theta \underline{e_1} + \cos\theta \underline{e_2} + 0.\underline{e_3} \\ \underline{e_3'} &= 0.\underline{e_1} + 0.\underline{e_2} + \underline{e_3}\end{aligned}$$

The matrix is given by:

$$\boldsymbol{G} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

iv) Dilation
Consider the map $\mathbb{R}^3 \to \mathbb{R}^3 \quad (x, y, z) \mapsto (\lambda x, \mu y, \nu z) \quad \lambda, \mu, \nu > 0$

$\underline{e'_1} = (\lambda, 0, 0) = \lambda \underline{e_1} \quad \underline{e'_2} = (0, \mu, 0) = \mu \underline{e_2} \quad \underline{e'_3} = (0, 0, \nu) = \nu \underline{e_3}$

Therefore the matrix for this is given by:

$$\boldsymbol{D} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{pmatrix}$$

If $\lambda = \mu = \nu$ we have 'pure dilation'

$$\boldsymbol{D} = \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

v) Shear
A simple shear in a plane displaces points in a direction (e.g. the $\boldsymbol{x}$ direction) by an amount proportional to the perpendicular distance from the origin (i.e. $\boldsymbol{y}$)

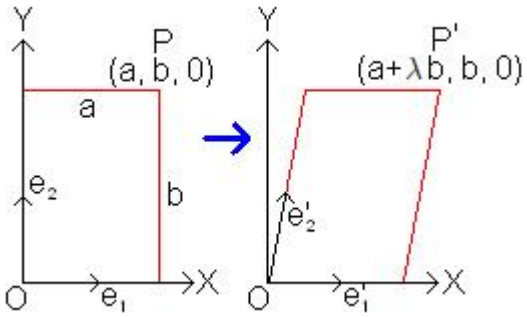

Figure 37: Simple 2-D shear

$$\underline{e'_1} = \underline{e_1}$$
$$\underline{e'_2} = \underline{e_2} + \lambda \underline{e_1}$$
$$\underline{e'_3} = \underline{e_3}$$

The matrix is given by:

$$\boldsymbol{S} = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Matrices for maps $\boldsymbol{V} \rightarrow \boldsymbol{W}$ with $\boldsymbol{W}$ not the same as $\boldsymbol{V}$

For $\boldsymbol{M : \mathbb{R}^3 \to \mathbb{R}^3}$ we have a 3x3 matrix
Now consider $\boldsymbol{A : \mathbb{R}^n \to \mathbb{R}^m}$ $(\boldsymbol{m \neq n})$
Let $\{\underline{\boldsymbol{e_k}}\}$ be the standard basis of $\mathbb{R}^n$, then $\underline{\boldsymbol{x}} \in \mathbb{R}^n \Rightarrow \exists\{\boldsymbol{x_k}\}$ such that $\sum_{k=1}^{n} \boldsymbol{x_k}\underline{\boldsymbol{e_k}} = \underline{\boldsymbol{x}}$

$$\boldsymbol{A}\underline{\boldsymbol{x}} = \sum_{k=1}^{n} \boldsymbol{x_k A}\underline{\boldsymbol{e_k}}$$

Let $\{\underline{\boldsymbol{f_j}}\}$ be the standard basis for $\mathbb{R}^m$

$$\therefore\ \exists \boldsymbol{a_{jk}} \in \mathbb{R} \text{ such that } \boldsymbol{A}\underline{\boldsymbol{e_k}} = \sum_{j=1}^{m} \boldsymbol{a_{jk}}\underline{\boldsymbol{f_j}}$$

Hence:

$$\begin{aligned}
\boldsymbol{A}\underline{\boldsymbol{x}} &= \sum_{k=1}^{n} \boldsymbol{x_k A}\underline{\boldsymbol{e_k}} \\
&= \sum_{k=1}^{n}\sum_{j=1}^{m} \boldsymbol{x_k a_{jk}}\underline{\boldsymbol{f_j}} \\
&= \sum_{j=1}^{m} \underbrace{\sum_{k=1}^{n} \boldsymbol{a_{jk} x_k}}_{\text{components of } \boldsymbol{A}\underline{\boldsymbol{x}} \text{ wrt } \{\underline{\boldsymbol{f_j}}\}} \underline{\boldsymbol{f_j}}
\end{aligned}$$

Therefore:

$$\underline{\boldsymbol{x'}} = \boldsymbol{A}\underline{\boldsymbol{x}} \quad \boldsymbol{x'_j = a_{jk} x_k} \text{ using suffix notation}$$

In explicit matrix form:

$$\underbrace{\begin{pmatrix} \boldsymbol{x'_1} \\ \boldsymbol{x'_2} \\ \vdots \\ \boldsymbol{x'_m} \end{pmatrix}}_{\text{column vector with } \boldsymbol{m} \text{ rows}} = \underbrace{\begin{pmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{12}} & \cdots & \boldsymbol{a_{1n}} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} & \cdots & \boldsymbol{a_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a_{m1}} & \boldsymbol{a_{m2}} & \cdots & \boldsymbol{a_{mn}} \end{pmatrix}}_{\boldsymbol{m}\text{x}\boldsymbol{n} \text{ matrix}} \underbrace{\begin{pmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \vdots \\ \boldsymbol{x_n} \end{pmatrix}}_{\text{column vector with } \boldsymbol{n} \text{ rows}}$$

Using the same rules for multiplication as before

4.5 Algebra of Matrices

Multiplication by scalars

Take a linear map $\boldsymbol{A} : \mathbb{R}^n \to \mathbb{R}^m$ and $\boldsymbol{\lambda} \in \mathbb{R}$
Define $(\boldsymbol{\lambda A})$ such that $(\boldsymbol{\lambda A})\underline{\boldsymbol{x}} = \boldsymbol{\lambda}(\boldsymbol{A}\underline{\boldsymbol{x}})\ \forall \underline{\boldsymbol{x}} \in \mathbb{R}^n$

If the matrix with respect to given bases of $\mathbb{R}^m, \mathbb{R}^n$ has elements $\{\boldsymbol{a_{ij}}$ then the matrix of $(\boldsymbol{\lambda A})$ with respect to the same basis has elements $\{\boldsymbol{\lambda a_{ij}}\}$

Addition
Similarly given matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, both $\boldsymbol{m}$x$\boldsymbol{n}$ matrices with elements $\{\boldsymbol{a_{ij}}\}, \{\boldsymbol{b_{ij}}\}$
we define $\boldsymbol{A + B} = \{\boldsymbol{a_{ij} + b_{ij}}\}$ - also an $\boldsymbol{m}$x$\boldsymbol{n}$ matrix

Matrix Multiplication
Let $\boldsymbol{S} : \mathbb{R}^n \to \mathbb{R}^m, \boldsymbol{T} : \mathbb{R}^m \to \mathbb{R}^l$ be linear maps.

Let the matrix of $\boldsymbol{S}$ wrt the standard basis be $\{\boldsymbol{a_{ij}}\}$ - A - $\boldsymbol{m}$x$\boldsymbol{n}$
and the matrix of $\boldsymbol{T}$ wrt the standard basis be $\{\boldsymbol{b_{ij}}\}$ - B - $\boldsymbol{l}$x$\boldsymbol{m}$

and let he composite map $\boldsymbol{TS} : \mathbb{R}^n \to \mathbb{R}^l$ have matrix $\boldsymbol{C} = \{\boldsymbol{c_{ij}}\}$ - $\boldsymbol{l}$x$\boldsymbol{n}$

We have $\underline{\boldsymbol{x'}} = \boldsymbol{S}(\underline{\boldsymbol{x}}) = \boldsymbol{A}\underline{\boldsymbol{x}}$ with $\boldsymbol{x'_i = a_{ij}x_j}$
$\underline{\boldsymbol{x''}} = \boldsymbol{T}(\underline{\boldsymbol{x'}}) = \boldsymbol{B}\underline{\boldsymbol{x'}}$ with $\boldsymbol{x''_i = b_{ij}x'_j}$
Combining these gives $\boldsymbol{x''_i = b_{ij}a_{jk}x_k}$
But also $\underline{\boldsymbol{x''}} = (\boldsymbol{TS})(\underline{\boldsymbol{x}}) = \boldsymbol{C}\underline{\boldsymbol{x}} = \boldsymbol{c_{ij}x_k}$

Hence $\boldsymbol{c_{ik} = b_{ij}a_{jk}}$

We can interpret this as the elements of the matrix product $\boldsymbol{BA}$

The $\boldsymbol{ik}^{\text{th}}$ element of $\boldsymbol{BA}$ is the scalar product of the $\boldsymbol{i}^{\text{th}}$ row of $\boldsymbol{B}$ and the $\boldsymbol{k}^{\text{th}}$ column of $\boldsymbol{A}$
This requires the number of columns of $\boldsymbol{B}$ to be equal to the number of rows of $\boldsymbol{A}$

If $\boldsymbol{A}$ is $\boldsymbol{m}$x$\boldsymbol{n}$, and $\boldsymbol{B}$ is $\boldsymbol{p}$x$\boldsymbol{q}$ then
$\boldsymbol{BA}$ exists only if $\boldsymbol{q = m}$, then $\boldsymbol{BA}$ is $\boldsymbol{p}$x$\boldsymbol{n}$
$\boldsymbol{AB}$ exists only if $\boldsymbol{n = p}$, then $\boldsymbol{AB}$ is $\boldsymbol{m}$x$\boldsymbol{q}$
If $\boldsymbol{m = n = p = q}$ then $\boldsymbol{AB}$ and $\boldsymbol{BA}$ exist, but in general $\boldsymbol{AB \neq BA}$

Examples:

i) $\begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{-1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}$
$\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{-1} & \mathbf{0} \end{pmatrix}$

ii) $\begin{pmatrix} p & q & r \\ s & t & u \end{pmatrix} \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \begin{pmatrix} pa+qc+re & pb+qd+rf \\ sa+tc+ue & sb+td+uf \end{pmatrix}$

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} \begin{pmatrix} p & q & r \\ s & t & u \end{pmatrix} = \begin{pmatrix} ap+bs & aq+bt & ar+bu \\ cp+ds & cq+dt & cr+du \\ ep+fs & eq+ft & er+uf \end{pmatrix}$$

Multiplication of matrices is associative
i.e. if $\boldsymbol{A, B, C}$ are matrices such that $\boldsymbol{AB, BC}$ exist then $\boldsymbol{(AB)C = A(BC)}$ (and both exist!)

We can verify this from suffix notation:

$$\begin{aligned} (A(BC))_{ij} &= A_{ik}(BC)_{kj} \\ &= A_{ik}B_{kl}C_{lj} \\ &= (AB)_{il}C_{lj} \\ &= ((AB)C)_{ij} \end{aligned}$$

Transpose - If $\boldsymbol{A = \{a_{ij}\}}$ is an $\boldsymbol{m}$x$\boldsymbol{n}$ matrix then the transpose $\boldsymbol{A^T}$ is $\boldsymbol{n}$x$\boldsymbol{m}$ with components $\boldsymbol{(A^T)_{ij} = A_{ji} = a_{ji}}$
$\boldsymbol{((A^T)^T) = A}$

Examples:

i) $\begin{pmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{3} & \mathbf{2} \\ \mathbf{4} & \mathbf{3} \end{pmatrix}^{\boldsymbol{T}} = \begin{pmatrix} \mathbf{2} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{2} & \mathbf{3} \end{pmatrix}$

ii) If $\underline{\boldsymbol{a}} = \begin{pmatrix} \boldsymbol{a_1} \\ \boldsymbol{a_2} \\ \vdots \\ \boldsymbol{a_n} \end{pmatrix}$ is a column vector then $\underline{\boldsymbol{a}}^{\boldsymbol{T}} = \begin{pmatrix} \boldsymbol{a_1} & \boldsymbol{a_2} & \boldsymbol{\ldots} & \boldsymbol{a_n} \end{pmatrix}$ is a row vector

$\boldsymbol{(AB)^T = B^T A^T}$ if these products exist, since if $\boldsymbol{C = AB = \{c_{ij}\}}$

$$\begin{aligned} c_{ij} &= a_{ik}b_{kj} \\ (C^T)_{ij} = c_{ji} &= a_{jk}b_{ki} \\ &= (A^T)_{kj}(B^T)_{ik} \\ &= (B^T A^T)_{ij} \end{aligned}$$

Example: $\underline{\boldsymbol{x}}, \underline{\boldsymbol{y}}$ are 3x1 column vectors
If $\boldsymbol{A}$ is a 3x3 matrix, then $\underline{\boldsymbol{x}}^{\boldsymbol{T}} \boldsymbol{A} \underline{\boldsymbol{y}}$ is a 1x1 matrix (i.e. a single number)

$\underline{x}^t \boldsymbol{A} \underline{y} = x_i A_{ij} y_j$

$(\underline{x}^t \boldsymbol{A} \underline{y})^T = (\underline{x}^T \boldsymbol{A} \underline{y}) \because (\underline{x}^T \boldsymbol{A} \underline{y})^T = y_i A_{ji} x_j$, which is the same as above

Symmetry

Symmetric Matrix - A square nxn matrix $\boldsymbol{A}$ is symmetric if $\boldsymbol{A} = \boldsymbol{A}^T$ i.e. $a_{ij} = a_{ji}$

Antisymmetric Matrix - It is antisymmetric if $\boldsymbol{A} = -\boldsymbol{A}^T$ i.e. $a_{ij} = -a_{ji}$
Hence for an antisymmetric matrix $a_{11} = a_{22} = a_{33} = \ldots = a_{nn} = 0$

Examples:

i) 3x3 symmetric matrix $\boldsymbol{S} = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ i.e. 6 independent elements

ii) 3x3 antisymmetric matrix $\boldsymbol{A} = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$ i.e. 3 independent elements

If we related $a = v_3, b = -v_2, c = v_1$
$\boldsymbol{A} = \{a_{ij} = \{\epsilon_{ijk} v_k\} \leftrightarrow \underline{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$

Trace - The trace of a square nxn matrix $\boldsymbol{A} = \{a_{ij}\}$ is $\mathbf{tr}\, \boldsymbol{A} = a_{ii}$, the sum of the diagonal elements

Example: Let $\boldsymbol{B} = \{b_{ij}\}$ be mxn
$\boldsymbol{C} = \{c_{ij}\}$ be nxm

$\boldsymbol{BC}, \boldsymbol{CB}$ exist, not usually the same
$\left.\begin{array}{l} \mathbf{tr}(\boldsymbol{BC}) = (\boldsymbol{BC})_{ii} = b_{ik} c_{ki} \\ \mathbf{tr}(\boldsymbol{CB}) = (\boldsymbol{CB})_{ii} = c_{ik} b_{ki} \end{array}\right\}$ equal

$\Rightarrow \mathbf{tr}(\boldsymbol{BC}) = \mathbf{tr}(\boldsymbol{CB})$

Identity/Unit Matrix $\boldsymbol{I} = \{\delta_{ij}\} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ for the 3x3 example
It is defined such that $\boldsymbol{IA} = \boldsymbol{AI} = \boldsymbol{A} \forall$ square $\boldsymbol{A}$

Invertible Matrix - A (square) matrix $\boldsymbol{A}$ is invertible if $\exists \boldsymbol{B}$ such that $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$. If $\boldsymbol{B}$ exists then we write $\boldsymbol{B} = \boldsymbol{A}^{-1}$, the inverse of $\boldsymbol{A}$

Theorem 20. If $\boldsymbol{A}^{-1}$ exists then it is unique

Proof. Suppose both $\boldsymbol{B}$ and $\boldsymbol{A}$ are inverses
$\therefore \boldsymbol{AB = BA = I = AC = CA}$
$\boldsymbol{B = IB = (CA)B = C(AB) = CI = C}$
$\therefore \boldsymbol{B = C}$ □

It follows that
i) $\boldsymbol{(A^{-1})^{-1} = A}$
ii) $\boldsymbol{(AB)^{-1} = B^{-1}A^{-1}}$
Provided both $\boldsymbol{A}$, $\boldsymbol{B}$ are both invertible and the same size

For a proof of ii): $\boldsymbol{B^{-1}A^{-1}(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I}$
Determinant
Consider the effect of linear maps on volume. The volume of a parallelipiped with sides $\underline{\boldsymbol{a}}, \underline{\boldsymbol{b}}, \underline{\boldsymbol{c}}$ is $\underline{\boldsymbol{a}}.(\underline{\boldsymbol{b}} \times \underline{\boldsymbol{c}})$, so the volume of the unit cube (sides $\underline{\boldsymbol{e_1}}, \underline{\boldsymbol{e_2}}, \underline{\boldsymbol{e_3}}$) is $\underline{\boldsymbol{e_1}}.(\underline{\boldsymbol{e_2}} \times \underline{\boldsymbol{e_3}}) = \boldsymbol{1}$

Now consider the cube under the mapping $\boldsymbol{A : \underline{e_1} \mapsto A\underline{e_1} = \underline{e'_1}}$. The volume of the mapped cube is $\underline{\boldsymbol{e'_1}}.(\underline{\boldsymbol{e'_2}} \times \underline{\boldsymbol{e'_3}})$.
This is called the determinant of $\boldsymbol{A}$, and is given explicitly by the formula

$$\begin{aligned}\boldsymbol{\det A} &= \boldsymbol{a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{21}(a_{32}a_{13} - a_{12}a_{33}) + a_{31}(a_{12}a_{23} - a_{22}a_{13})} \\ &= \boldsymbol{\epsilon_{ijk}\underline{e'_1}_i\ \underline{e'_2}_j\ \underline{e'_3}_k} \\ &= \boldsymbol{\epsilon_{ijk}a_{1i}a_{2j}a_{3k}}\end{aligned}$$

A linear map $\boldsymbol{A : \mathbb{R}^3 \to \mathbb{R}^3}$ is therefore volume preserving iff $\boldsymbol{\det A = \pm 1}$ where $\boldsymbol{A}$ is the matrix with respect to any basis

$\det\boldsymbol{A}$ is $\boldsymbol{> 0}$ if $\boldsymbol{A}$ sends a right-handed set to a right-handed set, and $\boldsymbol{< 0}$ is $\boldsymbol{A}$ sends a right-handed set to a left-handed set

In 2-D, we have $\boldsymbol{\det A = a_{11}a_{22} - a_{12}a_{21}}$, and the map is area preserving if $\boldsymbol{\det A = \pm 1}$

Example: Rotation matrix

$$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{\cos\theta} & \boldsymbol{-\sin\theta} & \boldsymbol{0} \\ \boldsymbol{\sin\theta} & \boldsymbol{\cos\theta} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} \end{pmatrix} = \boldsymbol{R} \qquad \boldsymbol{\det R = \pm 1}$$

For reflection in a plane, the determinant is -1 as it sends right-handed to left-handed sets

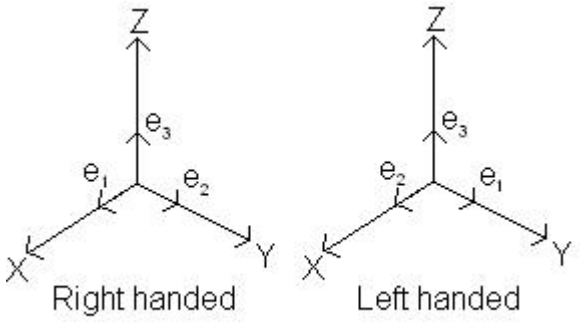

Figure 38: Right and left handed sets of basis vectors

4.6 Orthogonal matrices

Orthogonal Matrix - A nxn matrix $\boldsymbol{A} = \{a_{ij}\}$ is orthogonal if $\boldsymbol{A^T A} = \boldsymbol{AA^T} = \boldsymbol{I}$ i.e. $\boldsymbol{A}$ is invertible and $\boldsymbol{A^T} = \boldsymbol{A^{-1}}$
$\boldsymbol{AA^T} = \boldsymbol{I} \Rightarrow \boldsymbol{A)ik(A^T)_{kj}} = \boldsymbol{\delta_{ij}} \Rightarrow \boldsymbol{a_{ik}a_{jk}} = \boldsymbol{\delta_{ij}}$
i.e. ($\boldsymbol{i}^{\text{th}}$ row of $\boldsymbol{A}$) . ($\boldsymbol{j}^{\text{th}}$ row of $\boldsymbol{A}$) = $\boldsymbol{\delta_{ij}}$
i.e. the rows form an orthonormal set

If $\boldsymbol{A : \mathbb{R}^3 \to \mathbb{R}^3}$ has matrix $\boldsymbol{A}$ with respect to the standard basis, and with $\boldsymbol{A}$ orthogonal
$\underline{\boldsymbol{e_1}} \mapsto \boldsymbol{A}\underline{\boldsymbol{e_1}} = \boldsymbol{A}\underline{\boldsymbol{e_1}}$ ($\boldsymbol{1}^{\text{st}}$ column of matrix $\boldsymbol{A}$)
$\underline{\boldsymbol{e_2}} \mapsto \boldsymbol{A}\underline{\boldsymbol{e_2}} = \boldsymbol{A}\underline{\boldsymbol{e_2}}$ ($\boldsymbol{2}^{\text{nd}}$ column of matrix $\boldsymbol{A}$)
$\underline{\boldsymbol{e_3}} \mapsto \boldsymbol{A}\underline{\boldsymbol{e_3}} = \boldsymbol{A}\underline{\boldsymbol{e_3}}$ ($\boldsymbol{3}^{\text{rd}}$ column of matrix $\boldsymbol{A}$)

hence $\{\underline{\boldsymbol{e_1}}, \underline{\boldsymbol{e_2}}, \underline{\boldsymbol{e_3}}\}$ transforms to a new orthonormal set according to $\det \boldsymbol{A}$

Examples:
i) Reflection in the plane $\boldsymbol{\Pi}$
We noted that $\boldsymbol{R^2_\Pi} = \boldsymbol{I}$ and labelled the matrix associated with the map $\boldsymbol{H} = \{\boldsymbol{H_{ij}}\}$
Hence $\boldsymbol{H^2} = \boldsymbol{I}$
Recalled that $\boldsymbol{H_{ij}} = \boldsymbol{\delta_{ij}2n_in_j}$, and thus $\boldsymbol{H}$ is symmetric. Hence $\boldsymbol{H^TH} = \boldsymbol{HH^T} = \boldsymbol{I}$ and so the matrix $\boldsymbol{H}$ is orthogonal

ii) $\boldsymbol{R} = \begin{pmatrix} \boldsymbol{\cos\theta} & \boldsymbol{-\sin\theta} & \boldsymbol{0} \\ \boldsymbol{\sin\theta} & \boldsymbol{\cos\theta} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} \end{pmatrix}$

The columns and rows form and orthonormal set, so hence $\boldsymbol{R}$ is orthogonal

Scalar product is preserved by orthogonal matrix multiplication
e.g. $\underline{\boldsymbol{x}} \mapsto \underline{\boldsymbol{x'}} = \boldsymbol{A}\underline{\boldsymbol{x}}$ $\underline{\boldsymbol{y}} \mapsto \underline{\boldsymbol{y'}} = \boldsymbol{A}\underline{\boldsymbol{y}}$ with $\boldsymbol{A}$ orthogonal

Consider $\underline{\boldsymbol{x}}.\underline{\boldsymbol{y}} = \underline{\boldsymbol{x}}^{\boldsymbol{T}}\underline{\boldsymbol{y}}$

$$\begin{aligned}\underline{\boldsymbol{x}'}.\underline{\boldsymbol{y}'} &= (\boldsymbol{A}\underline{\boldsymbol{x}})^{\boldsymbol{T}}.(\boldsymbol{A}\underline{\boldsymbol{y}}) \\ &= \underline{\boldsymbol{x}}^{\boldsymbol{T}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\underline{\boldsymbol{y}} \\ &= \underline{\boldsymbol{x}}.\underline{\boldsymbol{y}}\end{aligned}$$

And thus as $\underline{\boldsymbol{x}}.\underline{\boldsymbol{x}} = \underline{\boldsymbol{x}'}.\underline{\boldsymbol{x}'}$ we can see that length is preserved, and thus if a transformation is represented by an orthogonal matrix it is an isometry. For $\mathbb{R}^{\mathbf{3}} \longrightarrow \mathbb{R}^{\mathbf{3}}$ the only isometries are rotations and reflections (and combinations) which have matrices that are orthogonal

4.7 Change of Basis

Consider changing from our standard basis of $\mathbb{R}^3$ to a new basis $\underline{\eta_1}, \underline{\eta_2}, \underline{\eta_3}$, linearly independent but not necessarily orthonormal (or even orthogonal)

Let $\underline{x}$ be any vector in $\mathbb{R}^3$, then

$$\underline{x} = \sum_{i=1}^{3} x_1 \underline{e_i} = \sum_{k=1}^{3} \xi_k \underline{\eta_k}$$

where $\{\xi_k\}$ are the components of $\underline{x}$ with respect to the new basis.
Consider $\underline{x}.\underline{e_j}$:

$$\underline{x}.\underline{e_j} = x_j = \sum_{k=1}^{3} \xi_k \underline{\eta_k}.\underline{e_j} = P_{jk}\xi_k$$

Where P_{jk} is the j^{th} component of $\underline{\eta_k}$ (with respect to the standard basis).

We wrtie $\underline{x} = P\underline{\eta}$ (where $\underline{x}$ and $\underline{\eta}$ are to be interpreted as column vectors whose elements are the x_i and η_i) where the matrix P is

$P = (\ \underline{\eta_1} \quad \underline{\eta_2} \quad \underline{\eta_3}\ )$ matrix with columns components of new basis vectors $\underline{\eta_k}$

Matrices are therefore a convenient way of expressing the changes in componenets due to a change of basis.

Since the $\underline{\eta_k}$ are a basis, there exist $E_{ki} \in \mathbb{R}$ such that $\underline{e_i} = \sum_{i=1}^{3} E_{ki}\underline{\eta_k}$. Hence

$$\underline{x} = \sum_{i=1}^{3} x_i (\sum_{k=1}^{3} E_{ki}\underline{\eta_k}) = \sum_{k=1}^{3} (\sum_{i=1}^{3} x_i E_{ki})\underline{\eta_k} = \sum_{k=1}^{3} \xi_k \underline{\eta_k}$$

By uniqueness of componets with respect to a given basis $E_{ki}x_i = \xi_k$

Thus we have $P\underline{\xi} = \underline{x}$ and $E\underline{x} = \underline{\xi}$ for all $\underline{x} \in \mathbb{R}^3$, so $P\underline{\xi} = \underline{x} = PE\underline{x}$ for all $\underline{x} \in \mathbb{R}^3$, hence $PE = I$. Similarly $EP = I$ and hence $E = P^{-1}$, so P is invertible.

Now consider a linear map $\mathcal{M} : \mathbb{R}^3 \to \mathbb{R}^3$ under which $\underline{x} \mapsto \underline{x'} = \mathcal{M}(\underline{x})$ and (in terms of column vectors $\underline{x'} = M\underline{x}$ where $\{x'_i\}$ and $\{x_i\}$ are components with respect to the standard basis $\{\underline{e_i}\}$). M is the matrix of $\mathcal{M}$ with respect to the standard basis.

From above $\underline{x'} = P\underline{\xi'}$ and $\underline{x} = P\underline{\xi}$ where $\{\xi'_j\}$ and $\{\xi_j\}$ are components with respect to the new basis $\{\underline{\eta_j}\}$

Thus $\boldsymbol{P\underline{\xi'} = MP\underline{\xi}}$, hence $\boldsymbol{\underline{\xi'} = (P^{-1}MP)\underline{\xi}}$

$\boldsymbol{P^{-1}MP}$ is the matrix of $\boldsymbol{\mathcal{M}}$ with respect to the new basis $\boldsymbol{\{\underline{\eta_j}\}}$, where $\boldsymbol{P = (\ \underline{\eta_1}\ \ \underline{\eta_2}\ \ \underline{\eta_3}\ )}$, i.e. the columns of $\boldsymbol{P}$ are the components of new basis vectors with respect to the old basis (and in this case, the old basis is the standard basis).

A similar approach may be used to deduce the matrix of the map $\boldsymbol{\mathcal{N} : \mathbb{R}^n \to \mathbb{R}^m}$ (where $\boldsymbol{m \neq n}$) with respect to new bases of both $\mathbb{R}^n$ and $\mathbb{R}^m$.

Suppose $\boldsymbol{\{\underline{e_i}\}}$ is the standard basis of $\mathbb{R}^n$ and $\boldsymbol{\{\underline{f_i}\}}$ is the standard basis of $\mathbb{R}^m$, and $\boldsymbol{N}$ is the matrix of $\boldsymbol{\mathcal{N}}$ with respect to these two bases, so $\boldsymbol{\underline{x} \mapsto \underline{x'} = N\underline{x}}$ (where $\boldsymbol{\underline{x}}$ and $\boldsymbol{\underline{x'}}$ are interpreted as column vectors of components).

Now consider new basis $\boldsymbol{\{\underline{\eta_i}\}}$ of $\mathbb{R}^n$ and $\boldsymbol{\{\underline{\phi_i}\}}$ of $\mathbb{R}^m$, with $\boldsymbol{P = (\ \underline{\eta_1}\ \ \dots\ \ \underline{\eta_n}\ )}$ (an $\boldsymbol{n}$x$\boldsymbol{n}$ matrix) and $\boldsymbol{Q = (\ \underline{\phi_1}\ \ \dots\ \ \underline{\phi_m}\ )}$ (an $\boldsymbol{m}$x$\boldsymbol{m}$ matrix).

Then $\boldsymbol{\underline{x} = P\underline{\xi}, \underline{x'} = Q\underline{\xi'}}$ where $\boldsymbol{\underline{\xi}}$ and $\boldsymbol{\underline{\xi'}}$ are column vectors of componenets with respect to bases $\boldsymbol{\{\underline{\eta_i}\}}$ and $\boldsymbol{\{\underline{\phi_i}\}}$ respectively.

Hence $\boldsymbol{Q\underline{\xi'} = NP\underline{\xi}}$, implying $\boldsymbol{\underline{\xi'} = Q^{-1}NP\underline{\xi}}$. So $\boldsymbol{Q^{-1}NP}$ is a matrix of transformation with respect to new bases (of $\mathbb{R}^n$ and $\mathbb{R}^m$).

Example: Consider a simple shear in the $\boldsymbol{x}$ direction within the $\boldsymbol{x, y}$ plane, with magnitude $\boldsymbol{\gamma}$
The matrix with respect to the standard basis $\boldsymbol{\{\underline{e_1}, \underline{e_2}, \underline{e_3}\}}$ is

$$\begin{pmatrix} \mathbf{1} & \boldsymbol{\gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix} = \boldsymbol{M}$$

Now consider the matrix of this transformation with respect to the basis $\boldsymbol{\{\underline{\eta_1}, \underline{\eta_2}, \underline{\eta_3}\}}$, where

$$\begin{aligned} \boldsymbol{\underline{\eta_1}} &= \boldsymbol{\cos\psi \underline{e_1} + \sin\psi \underline{e_2}} \\ \boldsymbol{\underline{\eta_2}} &= \boldsymbol{-\sin\psi \underline{e_1} + \cos\psi \underline{e_2}} \\ \boldsymbol{\underline{\eta_3}} &= \boldsymbol{\underline{e_3}} \end{aligned}$$

Then

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{\cos\psi} & \boldsymbol{-\sin\psi} & \mathbf{0} \\ \boldsymbol{\sin\psi} & \boldsymbol{\cos\psi} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix} \quad \text{which is orthogonal, so} \quad \boldsymbol{P^{-1}} = \begin{pmatrix} \boldsymbol{\cos\psi} & \boldsymbol{\sin\psi} & \mathbf{0} \\ \boldsymbol{-\sin\psi} & \boldsymbol{\cos\psi} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix}$$

The matrix with respect to the new basis is $\boldsymbol{P^{-1}MP =}$

$$\begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\psi + \gamma\sin\psi & -\sin\psi + \gamma\cos\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 + \gamma\sin\psi\cos\psi & \gamma\cos^2\psi & 0 \\ -\gamma\sin^2\psi & 1 - \gamma\sin\psi\cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

5 Determinants, Matrix Inverses and Linear Equations

5.1 Introduction

Consider linear equations in two unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= d_1 \\ a_{21}x_1 + a_{22}x_2 &= d_2 \end{aligned}$$

or equivalently, $A\underline{x} = \underline{d}$ where:

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \underline{d} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \text{ and } A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Now solve by forming suitable linear combinations of the two equations

$$\begin{aligned} (a_{11}a_{22} - a_{21}a_{12})x_1 &= a_{22}d_1 - a_{12}d_2 \\ (a_{21}a_{12} - a_{22}a_{21})x_2 &= a_{21}d_1 - a_{11}d_2 \end{aligned}$$

We identify $a_{11}a_{22} - a_{21}a_{12}$ as $\det A$ (defined earlier)

Thus if $\det A \neq 0$, the equations have a unique solution:

$$\begin{aligned} x_1 &= \frac{(a_2 2 d_1 - a_{12}d_2)}{\det A} \\ x_2 &= \frac{(-a_{21}d_1 + a_{11}d_2)}{\det A} \end{aligned}$$

Returning to matrix form, $A\underline{x} = \underline{d}$ implies $\underline{x} = A^{-1}\underline{d}$ (if A^{-1} exists). Thus we have that

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Check that $AA^{-1} = A^{-1}A = I$

5.2 Determinants for 3x3 and larger

For a 3x3 matrix, we write

$$\begin{aligned}\det \boldsymbol{A} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{33} & a_{32} & a_{33} \end{vmatrix} \\ &= a_1 1(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \\ &= a_1 1 \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_2 1 \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_3 1 \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\end{aligned}$$

We may use this as a way of defining (and evaluation) determinants of larger nxn matrices.

Properties of determinants

i) $\det \boldsymbol{A} = \det \boldsymbol{A}^T$ (follows from definition). Note that the expansion of 3x3 (or larger) determinants therefore works using rows as well as columns
ii) We noted earlier that

$$\det \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix} = \epsilon_{ijk}\alpha_I\beta_j\gamma_k = \underline{alpha}.(\underline{\beta} \times \underline{\gamma})$$

Now write $\alpha_i = a_{i1}, \beta_j = a_{j2}, \gamma_k = a_{k3}$. Then if $\boldsymbol{A} = \{a_{ij}\}, \det \boldsymbol{A} = \epsilon_{ijk}a_{i1}a_{j2}a_{k3}$
iii) (Following the triple product analogy) $\underline{alpha}.(\underline{\beta} \times \underline{\gamma}) = 0$ if and only if $\underline{\alpha}, \underline{\beta}, \underline{\gamma}$ are coplanar i.e. $\underline{\alpha}, \underline{\beta}, \underline{\gamma}$ are linearly dependent. Similarly, $\det \boldsymbol{A} = 0$ if and only if there is linear dependence between the columns of $\boldsymbol{A}$ (or, from i), the rows of $\boldsymbol{A}$)
iv) If we interchange any two of $\underline{\alpha}, \underline{\beta}$ and $\underline{\gamma}$ we change the sign of $\underline{\alpha}.(\underline{\beta} \times \underline{\gamma})$. Hence if we interchange any two columns of $\boldsymbol{A}$ we change the sign of $\det \boldsymbol{A}$. Similarly from i) we get the same effect if we interchange any two rows.
v) Add to any column of $\boldsymbol{A}$ linear combinations of other columns to give $\tilde{\boldsymbol{A}}$. Then $\det \tilde{(\boldsymbol{A})} = \det \boldsymbol{A}$ (Proved by considering $(\underline{\alpha} + \lambda\underline{\beta} + \mu\underline{\gamma}).(\underline{\beta} \times \underline{\gamma})$. A similar results applies to rows.
vi) Multiply any single row of column of $\boldsymbol{A}$ to give $\tilde{\boldsymbol{A}}$. Then again $\det \tilde{\boldsymbol{A}} = \det \boldsymbol{A}$
vii) $\det(\lambda \boldsymbol{A}) = \lambda^3 \det \boldsymbol{A}$ (or in general $\det(\lambda \boldsymbol{A}) = \lambda^n \det \boldsymbol{A}$ for nxn

Theorem 21. If $\boldsymbol{A} = \{a_{ij}\}$ is 3x3, then $\epsilon_{pqr} \det \boldsymbol{A} = \epsilon_{ijk}a_{pi}a_{qj}a_{rk}$

Proof. Use ii) above if $p = 1, q = 2, r = 3$
If p and q are swapped then sign of the LHS reverse, and

$$\epsilon_{ijk}a_{qi}a_{pj}a_{rk} = \epsilon_{jik}a_{qj}a_{pi}a_{rk} = -\epsilon_{ijk}a_{pi}a_{qj}a_{rk}$$

so sign of RHS also reverses. Similar for swaps for $\boldsymbol{p}$ and $\boldsymbol{r}$ or $\boldsymbol{q}$ and $\boldsymbol{r}$. Hence results holds for $\boldsymbol{\{pqr\}}$ any permutation of $\boldsymbol{\{123\}}$
If $\boldsymbol{p = q = 1}$, say, then LHS is also 0 and

$$\boldsymbol{\epsilon_{ijk}a_{1i}a_{1j}a_{rk} = \epsilon_{jik}a_{1j}a_{1i}a_{rk} = -\epsilon_{ijk}a_{1i}a_{1j}a_{rk}}$$

hence RHS is also 0. Similarly for any case where any pair of $\boldsymbol{p, q}$ and $\boldsymbol{r}$ are equal. Hence the result. □

Theorem 22. $\boldsymbol{\det AB = (\det A)(\det B)}$ with $\boldsymbol{A}$ and $\boldsymbol{B}$ both 3x3 matrices

Proof.

$$\begin{aligned}\boldsymbol{\det AB} &= \boldsymbol{\epsilon_{ijk}(AB)_{i1}(AB)_{j2}(AB)_{k3}} \\ &= \boldsymbol{\epsilon_{ijk}a_{ip}b_{p1}a_{j1}b_{q2}a_{kr}b_{r3}} \\ &= \boldsymbol{\epsilon_{pqr}\det Ab_{p1}b_{q2}b_{r3}} \text{ by Theorem 21} \\ &= \boldsymbol{\det A.\det B}\end{aligned}$$

□

Theorem 23. If $\boldsymbol{A}$ is orthogonal then $\boldsymbol{\det A = \pm 1}$

Proof. $\boldsymbol{AA^T = I}$ implies $\boldsymbol{\det(AA^T) = \det I = 1}$, which implies $\boldsymbol{\det A \det A^T = (\det A)^2 = 1}$, hence $\boldsymbol{\det A = \pm 1}$ □

5.3 Inverse of a 3x3 matrix

Define the cofactor $\boldsymbol{\Delta_{ij}}$ of the $\boldsymbol{ij}^{\text{th}}$ element of a square matrix $\boldsymbol{A}$ as

$$\boldsymbol{\Delta_{ij} = (-1)^{i+j} \det M_{ij}}$$

where $\boldsymbol{M_{ij}}$ is the square matrix obtained by eliminating the $\boldsymbol{i}^{\text{th}}$ row and the $\boldsymbol{j}^{\text{th}}$ column of A.
We have

$$\begin{aligned}\det \boldsymbol{A} = a_1 1 \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_1 2 \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_1 3 \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ = a_{11}\Delta_{11} + a_{12}\Delta_{12} + a_{13}\Delta_{13} \\ = a_{1j}\Delta_{1j}\end{aligned}$$

Similarly, noting that

$$\det \boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{vmatrix}$$

we have

$$\begin{aligned}\det \boldsymbol{A} = a_2 1 \begin{vmatrix} a_{32} & a_{33} \\ a_{12} & a_{13} \end{vmatrix} - a_2 2 \begin{vmatrix} a_{31} & a_{33} \\ a_{11} & a_{13} \end{vmatrix} + a_2 3 \begin{vmatrix} a_{31} & a_{32} \\ a_{11} & a_{12} \end{vmatrix} \\ = a_{21}\Delta_{21} + a_{22}\Delta_{22} + a_{23}\Delta_{23} = a_{2j}\Delta_{2j} \\ = a_{31}\Delta_{31} + a_{32}\Delta_{32} + a_{33}\Delta_{33} = a_{3j}\Delta_{3j}\end{aligned}$$

Similarly $\boldsymbol{\det A = a_{j1}\Delta_{j1} = a_{j2}\Delta_{j2} = a_{j3}\Delta_{j3}}$, but

$$a_{2j}\Delta_{1j} = \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0$$

Since the rows are linearly dependent

Theorem 24. $\boldsymbol{a_{ji}\Delta_{ki} = \det A \delta_{jk}}$ by the above

Theorem 25. Given a 3x3 matrix $\boldsymbol{A}$ with $\boldsymbol{\det A \neq 0}$, define $\boldsymbol{B}$ by

$$\boldsymbol{(B)_{ki} = (\det)^{-1}\Delta_{ik}}$$

Then $\boldsymbol{AB = BA = 1}$

Proof.

$$\boldsymbol{(AB)_{ij} = a_{ik}(B)_{kj} = (\det A)^{-1} a_{ik}\Delta_{jk} = (\det A)^{-1} \det A \delta_{ij} = \delta_{ij}}$$

Hence $\boldsymbol{AB = I}$, Similarly $\boldsymbol{BA = I}$. It follows that $\boldsymbol{B = A^{-1}}$ and $\boldsymbol{A}$ is invertible. □

The above is the formula for the inverse. A similar result holds for $\boldsymbol{n}$x$\boldsymbol{n}$ matrices, including 2x2

Example: Consider:

$$\boldsymbol{S} = \begin{pmatrix} \mathbf{1} & \boldsymbol{\gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix} \text{ representing a simple shear}$$

Then $\mathbf{det}\, \boldsymbol{S} = \mathbf{1}$ and

$$\begin{array}{ccc} \boldsymbol{\Delta_{11}} = \mathbf{1} & \boldsymbol{\Delta_{12}} = \mathbf{0} & \boldsymbol{\Delta_{13}} = \mathbf{0} \\ \boldsymbol{\Delta_{21}} = -\boldsymbol{\gamma} & \boldsymbol{\Delta_{22}} = \mathbf{1} & \boldsymbol{\Delta_{23}} = \mathbf{0} \\ \boldsymbol{\Delta_{31}} = \mathbf{0} & \boldsymbol{\Delta_{32}} = \mathbf{0} & \boldsymbol{\Delta_{33}} \end{array}$$

Hence

$$\boldsymbol{S^{-1}} = \begin{pmatrix} \boldsymbol{\Delta_{11}} & \boldsymbol{\Delta_{21}} & \boldsymbol{\Delta_{31}} \\ \boldsymbol{\Delta_{12}} & \boldsymbol{\Delta_{22}} & \boldsymbol{\Delta_{32}} \\ \boldsymbol{\Delta_{13}} & \boldsymbol{\Delta_{23}} & \boldsymbol{\Delta_{33}} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & -\boldsymbol{\gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix}$$

The effect of the shear is revered by changing the sign of $\boldsymbol{\gamma}$

5.4 Solving linear equations: Gaussian elimination

One approach to solving equations $\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{d}}$ (with $\boldsymbol{A}$ and $\boldsymbol{n}$x$\boldsymbol{n}$ matrix, $\underline{\boldsymbol{x}}$ and $\underline{\boldsymbol{d}}$ $\boldsymbol{n}$x1 column vectors of unknowns and right-hand sides respectively) numerically would be to calculate $\boldsymbol{A^{-1}}$ using the method given previously (extended to $\boldsymbol{n}$x$\boldsymbol{n}$) and thus $\boldsymbol{A^{-1}}\underline{\boldsymbol{d}}$. This is actually very inefficient.
An alternative is Gaussian elimination, illustrated here for the 3x3 case.

We have

$$\boldsymbol{a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = d_1} \quad (1)$$
$$\boldsymbol{a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = d_2} \quad (2)$$
$$\boldsymbol{a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = d_3} \quad (3)$$

Assume $\boldsymbol{a_{11} \neq 0}$, otherwise re-order, otherwise stop (since no unique solution). Then (1) may be used to eliminate $\boldsymbol{x_1}$:

$$\boldsymbol{x_1 = \frac{d_1 - a_{12}x_2 - a_{13}x_3}{a_{11}}}$$

Now (2) becomes

$$\boldsymbol{(a_{22} - \frac{a_{21}}{a_{11}})x_2 + (a_{23} - \frac{a_{21}}{a_{11}}a_{13})x_3 = d_2 - \frac{a_{21}}{a_{11}}d_1}$$
$$\boldsymbol{a'_{22}x_2 + a'_{23}x_3 = d'_2} \quad (2')$$

And (3) becomes

$$\boldsymbol{(a_{32} - \frac{a_{31}}{a_{11}})x_2 + (a_{33} - \frac{a_{31}}{a_{11}}a_{13})x_3 = d_3 - \frac{a_{31}}{a_{11}}d_1}$$
$$\boldsymbol{a'_{32}x_2 + a'_{33}x_3 = d'_2} \quad (3')$$

Assume $\boldsymbol{a'_{22} \neq 0}$, otherwise reorder, otherwise stop. Use (2') to eliminate $\boldsymbol{x'_2}$ from (3') to give

$$\boldsymbol{(a'_{33} - \frac{a'_{32}}{a'_{22}}a'_{23})x_3 = a''_{33}x_3 = d'_3 - \frac{a'_{32}}{a'_{22}}d'_2} \quad (3'')$$

Now, providing $\boldsymbol{a''_{33} \neq 0}$, (3'') gives $\boldsymbol{x_3}$, then (2') gives $\boldsymbol{x_2}$, then (1) gives $\boldsymbol{x_1}$. This method fails only if $\boldsymbol{A}$ is not invertible, i.e. only if $\boldsymbol{\det A = 0}$

5.5 Solving linear equations

If $\det \boldsymbol{A} \neq \mathbf{0}$ then the equations $\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{d}}$ have a unique solution $\underline{\boldsymbol{x}} = \boldsymbol{A}^{-1}\underline{\boldsymbol{d}}$. This is a corollary to Theorem 25.

What can we say about the solution if $\det \boldsymbol{A} = \mathbf{0}$? (As usual we consider $\boldsymbol{A}$ to be 3x3)
$\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{d}}$ $(\underline{\boldsymbol{d}} \neq \underline{\mathbf{0}})$ is a set of inhomogenous equations
$\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$ is the corresponding set of homogenous equation (with the unique solution $\boldsymbol{A}^{-1}\underline{\mathbf{0}} = \underline{\boldsymbol{x}}$ if $\det \boldsymbol{A} \neq \mathbf{0}$
We first consider the homogenous equations and then return to the inhomogenous equations.

a) Geometric view

Write $\underline{\boldsymbol{r_i}}$ as the vector with components equal to the elements of the $\boldsymbol{i}^{\text{th}}$ row of $\boldsymbol{A}$, for $\boldsymbol{i} = \mathbf{1, 2, 3}$. Then the above equations may be expressed as

$$\begin{aligned} \underline{\boldsymbol{r_i}}.\underline{\boldsymbol{x}} &= \boldsymbol{d_i} \quad (\boldsymbol{i} = \mathbf{1, 2, 3}) \quad \text{Inhomogenous equations} \\ \underline{\boldsymbol{r_i}}\underline{\boldsymbol{x}} &= \mathbf{0} \quad (\boldsymbol{i} = \mathbf{1, 2, 3}) \quad \text{Homogenous equations} \end{aligned}$$

Each individual equation represents a plane in $\mathbb{R}^{\mathbf{3}}$. The solution of each set of 3 equations is the intersection of 3 planes.
For the homogenous equations the three planes each pass through O. There are three possibilities:

i) intersection only at O
ii) three planes have a common line (including O)
iii) three planes coincide

If $\det \boldsymbol{A} \neq \mathbf{0}$ then $\underline{\boldsymbol{r_1}}.(\underline{\boldsymbol{r_2}} \times \underline{\boldsymbol{r_3}}) \neq \mathbf{0}$ and the set $\{\underline{\boldsymbol{r_1}}, \underline{\boldsymbol{r_2}}, \underline{\boldsymbol{r_3}}\}$ is linearly independent, with $\text{span}(\{\underline{\boldsymbol{r_1}}, \underline{\boldsymbol{r_2}}, \underline{\boldsymbol{r_3}}\}) = \mathbb{R}^{\mathbf{3}}$. The intersection of the planes $\underline{\boldsymbol{r_1}}.\underline{\boldsymbol{x}} = \mathbf{0}$ and $\underline{\boldsymbol{r_2}}.\underline{\boldsymbol{x}} = \mathbf{0}$ is the line $\{\underline{\boldsymbol{x}} \in \mathbb{R}^{\mathbf{3}} : \underline{\boldsymbol{x}} = \boldsymbol{\lambda}\underline{\boldsymbol{k}}, \boldsymbol{\lambda} \in \mathbb{R}, \underline{\boldsymbol{k}} = \underline{\boldsymbol{r_1}} \times \underline{\boldsymbol{r_2}}\}$. Then $\underline{\boldsymbol{r_3}}.\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$ implies $\boldsymbol{\lambda} = \mathbf{0}$, hence $\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$ and the three planes intersect only at the origin i.e case i)
If $\det = \mathbf{0}$ then the set $\{\underline{\boldsymbol{r_1}}, \underline{\boldsymbol{r_2}}, \underline{\boldsymbol{r_3}}\}$ is linearly dependent, with $\dim(\text{span}(\{\underline{\boldsymbol{r_1}}, \underline{\boldsymbol{r_2}}, \underline{\boldsymbol{r_3}}\})) = \mathbf{2 or 1}$. Assume 2. Then WLOG $\underline{\boldsymbol{r_1}}$ and $\underline{\boldsymbol{r_2}}$ are linearly independent and the intersection of the planes $\underline{\boldsymbol{r_1}}.\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$ and $\underline{\boldsymbol{r_2}}.\underline{\boldsymbol{x}} = \underline{\mathbf{0}}$ is the line $\{\underline{\boldsymbol{x}} \in \mathbb{R}^{\mathbf{3}} : \underline{\boldsymbol{x}} = \boldsymbol{\lambda}\underline{\boldsymbol{k}}, \boldsymbol{\lambda} \in \mathbb{R}, \underline{\boldsymbol{k}} = \underline{\boldsymbol{r_1}} \times \underline{\boldsymbol{r_2}}\}$. Since $\underline{\boldsymbol{r_1}}.(\underline{\boldsymbol{r_2}} \times \underline{\boldsymbol{r_3}}) = \mathbf{0}$, all points on this line satisfy $\underline{\boldsymbol{r_3}}.\underline{\boldsymbol{x}} = \mathbf{0}$ and the intersection of the three planes is a line i.e. case ii)
Otherwise $\dim(\text{span}(\{\underline{\boldsymbol{r_1}}, \underline{\boldsymbol{r_2}}, \underline{\boldsymbol{r_3}}\})) = \mathbf{1}$, $\underline{\boldsymbol{r_1}}, \underline{\boldsymbol{r_2}}$ and $\underline{\boldsymbol{r_3}}$ are all parallel, and $\underline{\boldsymbol{r_1}}.\underline{\boldsymbol{x}} = \mathbf{0}$ implies $\underline{\boldsymbol{r_2}}.\underline{\boldsymbol{x}} = \mathbf{0}$ and $\underline{\boldsymbol{r_3}}.\underline{\boldsymbol{x}} = \mathbf{0}$, so the intersection of the three planes is a plane i.e. case iii). (we may write the plane as $\{\underline{\boldsymbol{x}} \in \mathbb{R}^{\mathbf{3}} : \underline{\boldsymbol{x}} = \boldsymbol{\lambda}\underline{\boldsymbol{k}} + \boldsymbol{\mu}\underline{\boldsymbol{l}}, \boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}\}$ for any two linearly independent vectors $\underline{\boldsymbol{k}}$ and $\underline{\boldsymbol{l}}$ such that $\underline{\boldsymbol{k}}.\underline{\boldsymbol{r_i}} = \underline{\boldsymbol{l}}.\underline{\boldsymbol{r_i}}, \boldsymbol{i} = \mathbf{1, 2, 3}$).

b) Linear mapping view

Consider the linear map $\boldsymbol{T_A} : \mathbb{R}^3 \to \mathbb{R}^3$ such that $\underline{\boldsymbol{x}} \mapsto \underline{\boldsymbol{x'}} = \boldsymbol{A}\underline{\boldsymbol{x}}$. ($\boldsymbol{A}$ is the matrix of $\boldsymbol{T_A}$ with respect to the standard basis).
The kernel $\mathbf{ker}(\boldsymbol{T_A}) = \{\underline{\boldsymbol{x}} \in \mathbb{R}^3 : \boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}\}$, thus $\mathbf{ker}(\boldsymbol{T_A})$ is the 'solution space' of $\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}$ with dimension $\boldsymbol{nT_A}$

Case i) applies if $\boldsymbol{n(T_A) = 0}$
Case ii) applies if $\boldsymbol{n(T_A) = 1}$
Case iii) applies if $\boldsymbol{n(T_A) = 2}$

We now use the fact that if $\{\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}}, \underline{\boldsymbol{w}}\}$ is a basis for $\mathbb{R}^3$, then $\boldsymbol{I(T_A)} = \text{span}\{\boldsymbol{T_A}(\underline{\boldsymbol{u}}), \boldsymbol{T_A}(\underline{\boldsymbol{v}}), \boldsymbol{T_A}(\underline{\boldsymbol{w}})\}$. Consider the different cases:

i) $\boldsymbol{\lambda T_A}(\underline{\boldsymbol{u}}) + \boldsymbol{\mu T_A}(\underline{\boldsymbol{v}}) + \boldsymbol{\nu T_A}(\underline{\boldsymbol{w}}) = \underline{\boldsymbol{0}}$ implies that $\boldsymbol{T_A}(\boldsymbol{\lambda}\underline{\boldsymbol{u}} + \boldsymbol{\mu}\underline{\boldsymbol{v}} + \boldsymbol{\nu}\underline{\boldsymbol{w}}) = \underline{\boldsymbol{0}}$, which implies $\boldsymbol{\lambda}\underline{\boldsymbol{u}} + \boldsymbol{\mu}\underline{\boldsymbol{v}} + \boldsymbol{\nu}\underline{\boldsymbol{w}} = \underline{\boldsymbol{0}}$, hence $\boldsymbol{\lambda = \mu = \nu = 0}$. Hence $\{\boldsymbol{T_A}(\underline{\boldsymbol{u}}), \boldsymbol{T_A}(\underline{\boldsymbol{v}}), \boldsymbol{T_A}(\underline{\boldsymbol{w}})\}$ is linearly independent and $\boldsymbol{r(T_A) = 3}$
ii) WLOG choose $\underline{\boldsymbol{u}} \in \mathbf{ker}(\boldsymbol{T_A})$, then $\boldsymbol{T_A}(\underline{\boldsymbol{u}}) = \underline{\boldsymbol{0}}$. Consider $\text{span}(\{\boldsymbol{T_A}(\underline{\boldsymbol{v}}), \boldsymbol{T_A}(\underline{\boldsymbol{w}})\})$. $\boldsymbol{\mu T_A}(\underline{\boldsymbol{v}}) + \boldsymbol{\nu T_A}(\underline{\boldsymbol{w}}) = \underline{\boldsymbol{0}}$ implies $\boldsymbol{T_A}(\boldsymbol{\mu}\underline{\boldsymbol{v}} + \boldsymbol{\nu}\underline{\boldsymbol{w}}) = \underline{\boldsymbol{0}}$, hence $\exists \boldsymbol{\alpha} \in \mathbb{R}$ such that $\boldsymbol{\mu}\underline{\boldsymbol{v}} + \boldsymbol{\nu}\underline{\boldsymbol{w}} = \boldsymbol{\alpha}\underline{\boldsymbol{u}}$. Hence $-\boldsymbol{\alpha}\underline{\boldsymbol{u}} + \boldsymbol{\mu}\underline{\boldsymbol{v}} + \boldsymbol{\nu}\underline{\boldsymbol{w}} = \underline{\boldsymbol{0}}$ and hence $\boldsymbol{\alpha = \mu = \nu = 0}$ and $\boldsymbol{T_A}(\underline{\boldsymbol{v}})$ and $\boldsymbol{T_A}(\underline{\boldsymbol{w}})$ are linearly independent. Thus $\mathbf{dim}(\text{span}(\{\boldsymbol{T_A}(\underline{\boldsymbol{u}}), \boldsymbol{T_A}(\underline{\boldsymbol{v}}), \boldsymbol{T_A}(\underline{\boldsymbol{w}})\})) = \boldsymbol{r(T_A) = 2}$
iii) WLOG choose linearly independent $\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}} \in \mathbf{ker}(\boldsymbol{T_A})$, then $\boldsymbol{A}\underline{\boldsymbol{w}} \neq \underline{\boldsymbol{0}}$ and $\mathbf{dim}(\text{span}(\{\boldsymbol{T_A}(\underline{\boldsymbol{u}}), \boldsymbol{T_A}(\underline{\boldsymbol{v}}), \boldsymbol{T_A}(\underline{\boldsymbol{w}})\})) = \boldsymbol{r(T_A) = 1}$

Remarks:
a) In each of the above cases we have $\boldsymbol{n(T_A) + r(T_A) = 3}$, further examples of the 'rank-nullity' formula
b) In each case we also have $\boldsymbol{r(T_A)} = \mathbf{dim}(\text{span}(\{\underline{\boldsymbol{r_1}}, \underline{\boldsymbol{r_2}}, \underline{\boldsymbol{r_3}}\})) = \boldsymbol{r(T_A)} =$ number of linearly independent rows of $\boldsymbol{A}$ ('row rank'). But $\boldsymbol{r(T_A)} = \mathbf{dim}(\text{span}(\{\boldsymbol{A}\underline{\boldsymbol{e_1}}, \boldsymbol{A}\underline{\boldsymbol{e_2}}, \boldsymbol{A}\underline{\boldsymbol{e_3}}\}))$ (choosing standard basis) = number of linearly independent columns of $\boldsymbol{A}$ ('column rank')

Implication for inhomogenous equation $\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{d}}$

If $\mathbf{det}\, \boldsymbol{A} \neq \boldsymbol{0}$ then $\boldsymbol{r(T_A) = 3}$ and $\boldsymbol{I(T_A)} = \mathbb{R}^3$. Since $\underline{\boldsymbol{d}} \in \mathbb{R}^3$, $\exists \underline{\boldsymbol{x}} \in \mathbb{R}^3$ for which $\underline{\boldsymbol{d}}$ is the image under $\boldsymbol{T_A}$ i.e. $\underline{\boldsymbol{x}} = \boldsymbol{A^{-1}}\underline{\boldsymbol{d}}$ exists and is unique.

If $\mathbf{det}\, \boldsymbol{A} = \boldsymbol{0}$ then $\boldsymbol{r(T_A) < 3}$ and $\boldsymbol{I(T_A)}$ is a proper subspace of $\mathbb{R}^3$. Then either $\underline{\boldsymbol{d}} \notin \boldsymbol{I(T_A)}$, in which case there are no solutions and the equations are inconsistent, or $\underline{\boldsymbol{d}} \in \boldsymbol{I(T_A)}$, in which case there is at least one solution and the equations are consistent. The latter case is described by the next theorem:

Theorem 26. If $\underline{\boldsymbol{d}} \in \boldsymbol{I(T_A)}$ then the general solution to $\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{d}}$ can be written as $\underline{\boldsymbol{x}} = \underline{\boldsymbol{x_0}} + \underline{\boldsymbol{y}}$ where $\underline{\boldsymbol{x_0}}$ is a particular fixed solution of $\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{d}}$ and $\underline{\boldsymbol{y}}$ is the general solution of $\boldsymbol{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}$

Proof. $\boldsymbol{A\underline{x_0} = \underline{d}}$ and $\boldsymbol{A\underline{y} = \underline{0}}$, hence $\boldsymbol{A(\underline{x_0} + \underline{y}) = \underline{d} + \underline{0} = \underline{d}}$. If:

i) $\boldsymbol{n(T_A) = 0, r(T_A) = 3}$, then $\boldsymbol{\underline{y} = \underline{0}}$ and the solution is unique
ii) $\boldsymbol{n(T_A) = 1, r(T_A) = 2}$, then $\boldsymbol{\underline{y} = \lambda\underline{k}}$ and $\boldsymbol{\underline{x} = \underline{x_0} + \lambda\underline{k}}$ representing a line
iii) $\boldsymbol{n(T_A) = 2, r(T_A) = 1}$, then $\boldsymbol{y = \lambda\underline{k} + \mu\underline{l}}$ and $\boldsymbol{\underline{x} = \underline{x_0} + \lambda\underline{k} + \mu\underline{l}}$ representing a plane

□

Example: 2x2 case

$$\boldsymbol{A\underline{x} = \underline{d}}$$

$$\boldsymbol{\begin{pmatrix} 1 & 1 \\ a & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ b \end{pmatrix}}$$

$\boldsymbol{\det A = 1 - a}$
If $\boldsymbol{a \neq 1}$ then $\boldsymbol{\det A \neq 0}$ and so $\boldsymbol{A^{-1}}$ exists and is unique

$$\boldsymbol{A^{-1} = \frac{1}{1-a}\begin{pmatrix} 1 & -1 \\ -a & 1 \end{pmatrix}}$$ and the unique solution is $\boldsymbol{A^{-1}\begin{pmatrix} 1 \\ b \end{pmatrix}}$

If $\boldsymbol{a = 1}$ then $\boldsymbol{\det A = 0}$

$$\boldsymbol{A\underline{x} = \begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix}}, \boldsymbol{I(T_A)} = \text{span}\left\{\boldsymbol{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}\right\} \text{ and } \boldsymbol{\ker(T_A)} = \text{span}\left\{\boldsymbol{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}\right\}$$

sp $\boldsymbol{r(T_A) = 1}$ and $\boldsymbol{n(T_A) = 1}$

If $\boldsymbol{b \neq 1}$ then $\boldsymbol{\begin{pmatrix} 1 \\ b \end{pmatrix} \notin I(T_A)}$ and there are no solutions (equations inconsistent)
If $\boldsymbol{b = 1}$ then $\boldsymbol{\begin{pmatrix} 1 \\ b \end{pmatrix} \in I(T_A)}$ and solutions exist (equations consistent)
A particular solution is $\boldsymbol{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}}$
The general solution is $\boldsymbol{\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \underline{y}}$ where $\boldsymbol{y}$ is any vector in $\boldsymbol{\ker(T_A)}$
Hence the general solution is $\boldsymbol{\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda\begin{pmatrix} 1 \\ -1 \end{pmatrix}}$ where $\boldsymbol{\lambda \in \mathbb{R}}$

6 Complex Vector Spaces $\mathbb{C}^n$

6.1 Introduction

We have considered vector spaces with scalars $\in \mathbb{R}$
Now generalise to vector spaces with scalars $\in \mathbb{C}$

$\mathbb{C}^n$ is the set of n-tuples of complex numbers i.e. for $z \in \mathbb{C}^n, z = (z_1, z_2, \ldots, z_n), z_i \in \mathbb{C}, i = 1, 2, \ldots, n$

By extension from $\mathbb{R}^n$, define vector addition and scalar multiplication: for $z = (z_1, z_2, \ldots, z_n), \zeta = (\zeta_1, \zeta_2, \ldots, \zeta_n) \in \mathbb{C}^n$ and $c \in \mathbb{C}$

$$z + \zeta = (z_1 + \zeta_1, z_2 + \zeta_2, \ldots, z_n + \zeta_n)$$
$$cz = (cz_1, cz_2, \ldots, cz_n)$$

Check of A1-A4, B1-B4 from §3 shows that $\mathbb{C}^n$ is a vector space over $\mathbb{C}$
Note that $\mathbb{R}^n$ is a subset of $\mathbb{C}^n$, but not a subspace of $\mathbb{C}^n$ (as a vector space over $\mathbb{C}$) since $\mathbb{R}^n$ is not closed under multiplication by an arbitrary complex number.
$\mathbb{C}^n$ has dimension n as a vector space over $\mathbb{C}$ since the standard basis of $\mathbb{R}^n$ (as a vector space over $\mathbb{R}$) is also a basis of $\mathbb{C}^n$ (as a vector space over $\mathbb{C}$).

6.2 Linear Mappings

Consider $\boldsymbol{\mathcal{M}} : \mathbb{C}^{\boldsymbol{n}} \rightarrow \mathbb{C}^{\boldsymbol{m}}$
Let $\{\underline{\boldsymbol{e_i}}\}$ be the standard basis of $\mathbb{C}^{\boldsymbol{n}}$ and let $\{\underline{\boldsymbol{f_i}}\}$ be the standard basis of $\mathbb{C}^{\boldsymbol{m}}$
Then, under $\boldsymbol{\mathcal{M}}$

$$\underline{\boldsymbol{e_j}} \mapsto \underline{\boldsymbol{e'_j}} = \boldsymbol{\mathcal{M}}\underline{\boldsymbol{e_j}} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{3}} \boldsymbol{M_{ij}}\underline{\boldsymbol{f_i}}$$

where $\boldsymbol{M_{ij}} \in \mathbb{C}$. As before this defines the matrix of $\boldsymbol{\mathcal{M}}$ with respect to bases $\{\underline{\boldsymbol{e_i}}\}$ of $\mathbb{C}^{\boldsymbol{n}}$ and $\{\underline{\boldsymbol{f_i}}\}$ of $\mathbb{C}^{\boldsymbol{m}}$. $\boldsymbol{M}$ is a complex $\boldsymbol{m}$x$\boldsymbol{n}$ matrix.

Hermitian matrix - A square complex matrix $\boldsymbol{M}$ is Hermitian if $\overline{\boldsymbol{M^T}} = \boldsymbol{M}$

Unitary matrix - A square complex matrix $\boldsymbol{M}$ is unitary if $\overline{\boldsymbol{M^T}} = \boldsymbol{M^{-1}}$

6.3 Scalar product for $\mathbb{C}^n$

It we retain $\underline{z}.\underline{\zeta} = \sum_{i=1}^{n} z_i\zeta_i \in \mathbb{C}^n$ then we lose $\underline{z}.\underline{z} \in \mathbb{R}$ and hence $\underline{z}.\underline{z} \geq \mathbf{0}$. The natural extension of scalar products is to define it as

$$\underline{z}.\underline{\zeta} = \langle \underline{z}, \underline{\zeta} \rangle = \sum_{i=1}^{n} \overline{z_i}\zeta_i$$

Note that $\langle \underline{\zeta}, \underline{z} \rangle = \overline{\langle \underline{z}, \underline{\zeta} \rangle} \neq \langle \underline{z}, \underline{\zeta} \rangle$ (in general), but $\langle \underline{z}, \underline{z} \rangle \in \mathbb{R}$ and $\langle \underline{z}, \underline{z} \rangle > \mathbf{0}$ if $\underline{z} \neq \underline{\mathbf{0}}$
Hence we can use this new scalar product as a definition of length or norm:

$$||\underline{z}|| = \sqrt{\langle \underline{z}, \underline{z} \rangle} = \sqrt{\sum_{i=1}^{n} |z_i|^2} \text{ for } z \in \mathbb{C}^n$$

Note: We can also consider $\mathbb{C}^n$ as a vector space over $\mathbb{R}^n$. Note that $\mathbb{R}^n$ is a subspace and $\mathbf{dim}(\mathbb{C}^n) = \mathbf{2}n$

7 Revision of Part I

If $\boldsymbol{F}$ is a field, say $\mathbb{R}$ or $\mathbb{C}$ and $\boldsymbol{n} \in \mathbb{C}$, $\boldsymbol{V} = \boldsymbol{F^n}$ is an $\boldsymbol{n}$-dimensional vector space over $\boldsymbol{F}$ consisting of $\boldsymbol{n}$-tuples
Mostly we deal with $\boldsymbol{n = 2, 3}$

Let $\boldsymbol{B} = \{\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}\}$ be a fixed basis.
$\therefore$ any $\boldsymbol{v} \in \boldsymbol{V}$ can be written uniquely as a combination of these $\boldsymbol{v_i}$

$$\underline{v} = \sum_{i=1}^{n} c_i \underline{v_i} \quad \text{with } c_i \in F$$

Write $[\underline{v}]_B = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ - the column of coefficients of $\underline{v}$ with respect to $\boldsymbol{B}$

let $\boldsymbol{\alpha : V \to V}$ be a linear transformation, so $\boldsymbol{\alpha(a\underline{v} + b\underline{w}) = a\alpha(\underline{v}) + b\alpha(\underline{w}) \quad a, b \in F, \underline{v}, \underline{w} \in V}$. This can be represented by a matrix: Having fixed $\boldsymbol{B}$ as above

$$\begin{aligned} \boldsymbol{A} &= [\boldsymbol{\alpha}]_{\boldsymbol{B}} \text{ the matrix of } \boldsymbol{\alpha} \text{ with respect to } \boldsymbol{B} \\ &= (\boldsymbol{a_{ij}}) - \boldsymbol{n}\text{x}\boldsymbol{n} \end{aligned}$$

Which is give by $\boldsymbol{\alpha(\underline{v_j}) = \alpha_{ij} v_i}$, and thus

$$\boldsymbol{A} = ([\boldsymbol{\alpha}(\underline{v_1})]_{\boldsymbol{B}} \mid [\boldsymbol{\alpha}(\underline{v_2})]_{\boldsymbol{B}} \mid \ldots \mid [\boldsymbol{\alpha}(\underline{v_n})]_{\boldsymbol{B}}) \tag{4}$$

Thus for $\underline{v} \in \boldsymbol{V}$

$$[\boldsymbol{\alpha}(\underline{v})]_{\boldsymbol{B}} = [\boldsymbol{\alpha}]_{\boldsymbol{B}} [\underline{v}]_{\boldsymbol{B}} = \boldsymbol{A}[\underline{v}]_{\boldsymbol{B}} \tag{5}$$

that is

$$\text{if } \boldsymbol{\alpha}(\underline{v}) = \sum_{i} \boldsymbol{d_i} \underline{v_i} \text{ then } \boldsymbol{d_i} = \sum_{i} \boldsymbol{a_{ij} c_j} \tag{6}$$

Recall that $\underline{v} = \sum_j \boldsymbol{c_j v_j}$

To check this $\forall \underline{v} \in \boldsymbol{V}$ it is enough to check it $\forall \underline{v_i} \in \boldsymbol{B}$

$[\underline{v_i}]_{\boldsymbol{B}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ with the 1 in the $\boldsymbol{i}^{\text{th}}$ place, and $\boldsymbol{A} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ is the $\boldsymbol{i}^{\text{th}}$ column of $\boldsymbol{A}$

This depends on the choice of $\boldsymbol{B}$. Given $\boldsymbol{\alpha}$, often want to choose $\boldsymbol{B}$ so that $[\boldsymbol{\alpha}]_{\boldsymbol{B}}$ is 'nice'

What happens under a change of basis? Given

$$\boldsymbol{B} = \{\underline{\boldsymbol{v_1}}, \underline{\boldsymbol{v_2}}, \ldots, \underline{\boldsymbol{v_n}}\} \text{ an old basis for } \boldsymbol{V}$$
$$\boldsymbol{C} = \{\underline{\boldsymbol{w_1}}, \underline{\boldsymbol{w_2}}, \ldots, \underline{\boldsymbol{w_n}}\} \text{ a new basis for } \boldsymbol{V}$$

Write $\boldsymbol{P} = (\boldsymbol{p_{ij}})$ for the $\boldsymbol{n}$x$\boldsymbol{n}$ matrix obtained as $\boldsymbol{w_j} = \sum_{\boldsymbol{i}} \boldsymbol{p_{ij}} \boldsymbol{v_i}$
So $\boldsymbol{P} = ([\underline{\boldsymbol{w_1}}]_{\boldsymbol{B}} \mid [\underline{\boldsymbol{w_2}}]_{\boldsymbol{B}} \mid \ldots \mid [\underline{\boldsymbol{w_n}}]_{\boldsymbol{B}})$, which is the change of basis matrix from $\boldsymbol{B}$ to $\boldsymbol{C}$
This is invertible, with the inverse going from $\boldsymbol{C}$ to $\boldsymbol{B}$. Then

$$[\underline{\boldsymbol{v}}]_{\boldsymbol{B}} = \boldsymbol{P}[\underline{\boldsymbol{v}}]_{\boldsymbol{C}} \; \forall \underline{\boldsymbol{v}} \in \boldsymbol{V} \tag{7}$$

Theorem 27. If $\boldsymbol{\alpha} : \; \rightarrow \boldsymbol{V}$ is linear and $\boldsymbol{B}$ and $\boldsymbol{C}$ are bases for $\boldsymbol{V}$, let $\boldsymbol{P}$ be the change of basis from $\boldsymbol{B} \rightarrow \boldsymbol{C}$. Then

$$[\boldsymbol{\alpha}]_{\boldsymbol{C}} = \boldsymbol{P^{-1}}[\boldsymbol{\alpha}]_{\boldsymbol{B}}\boldsymbol{P}$$

Proof.

$$[\boldsymbol{\alpha}(\underline{\boldsymbol{v}})]_{\boldsymbol{B}} = [\boldsymbol{\alpha}]_{\boldsymbol{B}}[\underline{\boldsymbol{v}}]_{\boldsymbol{B}} \text{ by 2}$$
$$\text{so } \boldsymbol{P}[\boldsymbol{\alpha}(\underline{\boldsymbol{v}})]_{\boldsymbol{C}} = [\boldsymbol{\alpha}]_{\boldsymbol{B}}\boldsymbol{P}[\underline{\boldsymbol{v}}]_{\boldsymbol{C}} \text{ by 4}$$
$$\text{so } [\boldsymbol{\alpha}(\underline{\boldsymbol{v}})]_{\boldsymbol{C}} = \boldsymbol{P^{-1}}[\boldsymbol{\alpha}]_{\boldsymbol{B}}\boldsymbol{P}[\underline{\boldsymbol{v}}]_{\boldsymbol{C}} \; \forall \underline{\boldsymbol{v}} \in \boldsymbol{V}$$

But $[\boldsymbol{alpha}]_{\boldsymbol{C}}$ is the unique matrix for which this holds, so

$$[\boldsymbol{\alpha}]_{\boldsymbol{C}} = \boldsymbol{P^{-1}}[\boldsymbol{\alpha}]_{\boldsymbol{B}}\boldsymbol{P}$$

as claimed □

Conjugate matrices - The $\boldsymbol{n}$x$\boldsymbol{n}$ matrices $\boldsymbol{C}$, $\boldsymbol{D}$ are conjugate (or similar) if $\boldsymbol{C} = \boldsymbol{P^{-1}DP}$ for some invertible matrix $\boldsymbol{P}$. Equivalently $\boldsymbol{PC} = \boldsymbol{DP}$ and $\boldsymbol{P}$ is invertible

Lemma 28. i) $\boldsymbol{\det(P^{-1}DP) = \det D}$ so conjugate matrices have the same determinant
ii) Recall that trace$(\boldsymbol{D}) = \mathbf{tr}\, \boldsymbol{D} = \sum_{\boldsymbol{i}} \boldsymbol{d_{ii}}$.
tr$(\boldsymbol{P^{-1}DP})$ = tr$(\boldsymbol{D})$ so conjugate matrices have the same trace
Hence we can define $\mathbf{det}\, \boldsymbol{\alpha}$ and $\mathbf{tr}\, \boldsymbol{\alpha}$ for any linear transformation $\boldsymbol{\alpha} : \boldsymbol{V} \rightarrow \boldsymbol{V}$ to be

$$\left.\begin{array}{r} \mathbf{det}\, \boldsymbol{\alpha} = \mathbf{det}[\boldsymbol{\alpha}]_{\boldsymbol{B}} \\ \mathbf{tr}\, \boldsymbol{\alpha} = \mathbf{tr}[\boldsymbol{\alpha}]_{\boldsymbol{B}} \end{array}\right\} \text{ for any valid basis } \boldsymbol{B}$$

Proof. i)

$$\begin{aligned}\det(P^{-1}DP) &= \det P^{-1} \det D \det P \\ &= (\det P)^{-1} \det P \det D \\ &= \det D\end{aligned}$$

ii) $\operatorname{tr} AB = \operatorname{tr} BA$ as $a_{ij}b_{ji} = b_{ij}a_{ji}$, so

$$\operatorname{tr}(P^{-1}DP) = \operatorname{tr}(DPP^{-1}) = \operatorname{tr}(D)$$

□

8 Eigenvalues and Eigenvectors

Let $\boldsymbol{V}$ be a vector space over $\boldsymbol{F}$, with $\boldsymbol{\alpha : V \to V}$ linear

eigenvalue - $\boldsymbol{\lambda \in F}$ is an eigenvalue of $\boldsymbol{\alpha}$ if there is some non-zero vector $\underline{\boldsymbol{v}}$ such that $\boldsymbol{\alpha(\underline{v}) = \lambda \underline{v}}$

eigenvector - The vector referred to above is called an eigenvector

Notes:
i) For $\boldsymbol{L =< \underline{v} >= \text{span}\{\underline{v}\}}$ with $\boldsymbol{\alpha(\underline{v}) = \lambda \underline{v}}$ then $\boldsymbol{\alpha(L) \subset L}$ so $\boldsymbol{L}$ is a line fixed by $\boldsymbol{\alpha}$
ii) $\boldsymbol{\alpha(\underline{v}) = \lambda \underline{v} \Rightarrow \alpha^n(\underline{v}) = \lambda^n \underline{v}}$ and $\boldsymbol{(c_m \alpha^m + \cdots + c_1 \alpha + c_0 I)(\underline{v}) = (c_m \lambda^m + \cdots + c_1 \lambda + c_0)\underline{v}}$, so $\boldsymbol{p(\alpha)\underline{v}) = p(\lambda)\underline{v}}$ for any polynomial $\boldsymbol{p(x)}$

Lemma 29. If $\boldsymbol{\lambda}$ is an eigenvalue of $\boldsymbol{\alpha}$ $\boldsymbol{V_\lambda = \{\underline{v} : \alpha(\underline{v}) = \lambda \underline{v}\}}$
Then $\boldsymbol{V_\lambda}$ is a subspace of $\boldsymbol{V}$ - the $\boldsymbol{\lambda}$-eigenspace.

Proof. $\boldsymbol{0 \in V_\lambda}$
if $\boldsymbol{\underline{v}, \underline{w} \in V_\lambda}$ and $\boldsymbol{a, b \in F}$ then

$$\begin{aligned} \boldsymbol{\alpha(a\underline{v} + b\underline{w})} &= \boldsymbol{a\alpha(\underline{v}) + b\alpha(\underline{w})} \\ &= \boldsymbol{a\lambda\underline{v} + b\lambda\underline{w}} \\ &= \boldsymbol{\lambda(a\underline{v} + b\underline{w})} \end{aligned}$$

so $\boldsymbol{a\underline{v} + b\underline{w}}$ lies in $\boldsymbol{V_\lambda}$ so it a subspace (closed under linear combination) □

Lemma 30. $\boldsymbol{\lambda}$ is an e-value of $\boldsymbol{\alpha}$ iff $\boldsymbol{\det(\alpha - \lambda I) = 0}$

Proof.

$\boldsymbol{\lambda}$ is an evalue	iff $\boldsymbol{(\alpha - \lambda I)}$ is singular
	iff $\boldsymbol{\det(\alpha - \lambda I) = 0}$
	iff $\boldsymbol{\ker(\alpha - \lambda I) \neq 0}$

□

We can thus define the eigenvalue of an $\boldsymbol{n}$x$\boldsymbol{n}$ matrix as a value $\boldsymbol{\lambda}$ for which $\boldsymbol{\det(A - \lambda I) = 0}$, or equivalently, for some $\boldsymbol{\underline{v} \in F^n, \underline{v} \neq 0}$ $\boldsymbol{A\underline{v} = \lambda \underline{v}}$

The characteristic polynomial The characteristic polynomial of the linear map $\boldsymbol{\alpha : V \to V}$ is the polynomial $\boldsymbol{\chi_\alpha(t) = \det(\alpha - tI)}$. The characteristic polynomial of the $\boldsymbol{n}$x$\boldsymbol{n}$ matrix $\boldsymbol{A}$ is

$$\boldsymbol{\chi_\alpha(t) = \det(A - tI)}$$

Note: The eigenvalues are the roots of the polynomial

Lemma 31. Conjugate matrices have the same characteristic polynomial

Proof. If $\boldsymbol{C = P^{-1}DP}$ then

$$\begin{aligned}\boldsymbol{\chi_C(t) = \det(C - tI)} &= \boldsymbol{\det(P^{-1}DP - tP^{-1}P)} \\ &= \cancel{\boldsymbol{\det P^{-1}}}\,\boldsymbol{\det(D - tI)}\cancel{\boldsymbol{\det P}} \\ &= \boldsymbol{\chi_D(t)}\end{aligned}$$

□

Example: Let $\boldsymbol{\alpha : \mathbb{R}^2 \to \mathbb{R}^2}$ be given by

$$\boldsymbol{(x, y) \mapsto (x + 2y, -x + 4y)}$$

with respect to the standard basis $\boldsymbol{B}$ we have:

$$\boldsymbol{A : [\alpha]_B = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}}$$

What are the e-values and e-vectors of $\boldsymbol{\alpha}$?

$\boldsymbol{\chi_\alpha(t) = \det(A - tI) = \det \begin{pmatrix} 1-t & 2 \\ -1 & 4-t \end{pmatrix} = t^2 - 5t + 6 = (t-2)(t-3)}$
so the e-values are 2, 3
To find the e-vectors consider the two cases separately:
$\boldsymbol{\lambda = 2}$:

$$\boldsymbol{(A - 2I)\underline{v_2} = \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}}$$

$$\boldsymbol{\therefore \underline{v_2} = \text{span}\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}}$$

$\boldsymbol{\lambda = 3}$:

$$\boldsymbol{(A - 3I)\underline{v_3} = \begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}}$$

$$\boldsymbol{\therefore \underline{v_3} = \text{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}}$$

Put $\boldsymbol{C = \left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}}$, and $\boldsymbol{P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}}$, the change of basis matrix from the standard $\boldsymbol{B}$ to eigen $\boldsymbol{C}$. Then

$$\boldsymbol{P^{-1}AP = \underbrace{\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}}_{\text{e-values}} = [\alpha]_C}$$

Recall: $\boldsymbol{p(t) = a_k t^k + a_{k-1}t^{k-1} + \cdots a_1 t + a_0}$ is a polynomial of degree $\boldsymbol{k}$ if $\boldsymbol{a_k \neq 0}$

Remark: If $\boldsymbol{A}$ is $\boldsymbol{n}$x$\boldsymbol{n}$, then $\boldsymbol{\chi_A(t)}$ is a polynomial of degree $\boldsymbol{n}$
In fact, $\boldsymbol{\chi(t) = (-1)^n t^n + \cdots}$
The eigenvalues are precisely the roots of $\boldsymbol{\chi_A(t) = 0}$

Theorem 32. Fundamental Theorem of Algebra If $f(z) = a_n z^n + \cdots + a_1 z + a_0$ is a polynomial of degree $n/geq1$, with the coefficients $a_i \in \mathbb{C}$, then f has a root in $\mathbb{C}, \exists z_0 \in \mathbb{C}$ with $f(z_0) = 0$

Proof. See Further analysis course □

Theorem 33. Let V be an n-dimensional complex vector space, $n/geq1$. If $\alpha : V \to V$ is a linear map, then α has an eigenvector

Proof. (From the Fundamental Theorem of Algebra) The characteristic polynomial of α is a polynomial of degree $n \geq 1$ with complex coefficients, so it has a root $\lambda \in \mathbb{C}$ that is an eigenvalue, so there is a corresponding eigenvector. □

Example 34. It is not true in general for real vector spaces of even dimension. eg. $V = \mathbb{R}^2$, let α be a rotation of $\mathbb{R}^2$ about the origin by θ there are no real eigenvectors unless $\theta = 0, \pi$ with associated $\alpha = I$ or $\alpha = -I$
This is geometrically clear or, by considering its matrix (with respect to the standard basis)

$$\rho_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

Which has characteristic polynomial $t^2 - 2\cos\theta t + 1$, which has no real roots unless $\theta = 0, \pi$

We can consider this instead on the complex vector space $\mathbb{C}^2$. Then we do have eigenvalues $e^{i\theta}$, complex e-values. Interesting maybe, but not relevant to real life...

Lemma 35. Let $p(z)$ be a polynomial of degree n with coefficients in F, let z_0 be a vector in F, so $p(z_0) = 0$. Then there is a polynomial q of degree $n - 1$ with coefficients in F such that $p(z) = q(z)(z - z_0)$

Proof. If $p(z) = a_n z^n + \cdots + a_1 z + a_0$ then $p(z) = p(z) - \underbrace{p(z_0)}_{0} =$
$a_n(z^n - z_0^n) + \cdots + a_1(z - z_0)$
since $(z^k - z_0^k) = (z - z_0)\underbrace{(z^{k-1} + z^{k-2}z_0 + \cdots + z_0^{k-1})}_{q_k(z)}$
we have $p(z) = (z - z_0)\sum_{k=1}^{n} a_k q_k(z)$ □

Corollary 36. A polynomial of degree n over any field F has at most n roots

Proof. Induction on n, using Lemma 35 □

Multiplicity of roots - For the polynomial $p(z)$ the multiplicity of the root z_0 is defined to be the highest e with $(z - z_0)^e$ dividing $p(z)$ - so $p(z) = (z - z_0)^e q(z)$ with $q(z)$ a polynomial of $\deg(n - e), q(z_0) \neq 0$

Corollary 37. A complex polynomial of degree n has precisely n roots each counted with its multiplicity.
If $p(z) = c_n z^n + \cdots c_1 z + c_0$, then $p(z) = c_n \prod_{i=1}^{k} (z - z_i)^{e_i}$ where $z_1, \ldots, z_k$ are the distinct roots of p and $e_1, \ldots, e_k$ are the multiplicities with $n = \sum_i e_i$

Lemma 38. If $p(t)$ is a real polynomial of odd degree n then it has at least one real root

Proof. This follows from the Intermediate Value Theorem in Analysis I. WLOG We may assume that the leading coefficient is positive. $\therefore$ as $t \to \pm\infty$ $p(t) \to \pm\infty$, and as it is continuous it must therefore cross the axis in between. □

Proof. - Alternatively if z_0 is a complex non-real root of $p(z)$ so is $\overline{z_0}$ (as $p(\overline{z_0}) = \overline{p(z_0)} = \overline{0} = 0$), so $p(z) = \underbrace{[(z - z_0)(z - \overline{z_0})]}_{\text{real coeff's}} \underbrace{q(z)}_{\text{real coeff's}}$ with $\deg(q) = \deg(p) - 2$, and use induction until $\deg q = 1$, which has 1 real root. □

Corollary 39. If V is a real vector space of odd dimension then any linear map $\alpha : V \to V$ has a real eigenvalue

Proof. The characteristic polynomial has odd degree n and is real, and so has a real root, so α has a real eigenvector □

Remark 40. Recall $\chi_A(t) = \det(A - tI)$
Assume A is an nxn matrix with n eigenvalues $\lambda_1, \ldots, \lambda_n$ (not necessarily distinct, they can be repeated according to multiplicities). Let $\chi_A(t) = c_n t^n + \cdots + c_1 t + c_0$ so $c_n = (-1)^n$
Then i) $c_0 = \det A = \lambda_1 \lambda_2 \ldots \lambda_n$
ii) $(-1)^{n-1} c_{n-1} = \operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Proof. i) $c_0 = \chi_A(0) = \det(A - 0I) = \det A$
and as $\chi_A(t) = (\lambda_1 - t)(\lambda_2 - t) \ldots (\lambda_n - t)$
$\chi_A(0) = \lambda_1 \lambda_2 \ldots \lambda_n$
ii)

$$\chi_A(t) = \det \begin{pmatrix} a_{11} - t & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - t & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - t \end{pmatrix}$$

$= (a_{11} - t)$(polynomial of degree $< n - 1) - a_{21}$(polynomial of degree $< n - 1) + a_{31}$(polynomi

so the coefficient of t^{n-1} comes from $9a_{11} - t)|$smaller det$| = (a_{11} - t)(a_{22} - t) \ldots (a_{nn} - t)$
So it is $(-1)^{n-1} \operatorname{tr} A$ by multiplying out the above □

Proof. On the other hand, the coefficient of t^{n-1} is $(\lambda_1 - t)(\lambda_2 - t) \ldots (\lambda_n - t) = (-1)^{n-1}(\lambda_1 + \cdots + \lambda_n)$ □

9 Diagonal and Upper Triangular Matrices

Diagonal Matrix - The nxn matrix $A = (a_{ij})$ is diagonal if $a_{ij} = 0$ for $i \neq j$

Upper Triangular Matrix - It is upper triangular if $a_{ij} = 0$ for $i > j$

Remark 41. The eigenvalues of an upper triangular matrix are precisely its diagonal entries

$$\chi_A(t) = \det \begin{pmatrix} a_{11} - t & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} - t \end{pmatrix} = (a_{11} - t).(\text{something smaller, same shape})$$
$$= (a_{11} - t)(a_{22} - t) \dots (a_{nn} - t)$$

Lemma 42. Let $\alpha : V \to V$ be a linear map, let B be a basis for the vector space.
Then the matrix $[\alpha]_B$ is diagonal iff B consists of eigenvectors

Proof. Let $B = \{\underline{v_1}, \underline{v_2}, \dots, \underline{v_n}\}$
Then $\underline{v_i}$ is an eigenvector of α, say $\alpha \underline{v_i} = \lambda \underline{v_i}$ iff the i-column of $[\alpha]_B$ is $\begin{pmatrix} 0 \\ \vdots \\ \lambda \\ \vdots \\ 0 \end{pmatrix}$ (λ in the i^{th} row) □

Diagonalisable matrix - We have that the linear map $\alpha : V \to V$ is diagonalisable if V has a basis of eigenvectors of α, or equally that the nxn matrix A over F is diagonalisable over F if A is conjugate over F to a diagonal matrix, so $P^{-1}AP$ is diagonal for some invertible matrix P with entries in F

Example 43. How do matrices/maps fail to be diagonalisable?
i) The real matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is not diagonalisable over $\mathbb{R}$ as it has no eigenvalues in $\mathbb{R}$.
Considered over $\mathbb{C}$ it becomes conjugate to $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ with characteristic polynomial $t^2 = -1$ and so e-values $i, -i$
ii) The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalisable over any field. The eigenvalues are 1 and 1, the eigenspace V_1 is one dimensional, so cannot diagonalise
(Equally it would have to be conjugate to the identity, as it has the same e-values, but $P^{-1}IP = I \; \forall P \neq \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ so not conjugate)

So not all matrices can be diagonalised.

Proposition 44. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of the linear map $\alpha : V \to V$ and if $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_k}$ are any (non-zero) eigenvectors with $\alpha \underline{v_i} = \lambda_i \underline{v_i}$ (not using summation convection) for each i, then $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_k}$ are linearly independent.

Proof. Suppose this is false, let $c_1\underline{v_1} + \cdots, c_j\underline{v_j} = \underline{0}$ be the shortest non-trivial linear combination giving $\underline{0}$. Apply α to both sides to get

$$c_1\lambda_1\underline{v_1} + \cdots + c_j\lambda_j\underline{v_j} = \underline{0}$$

Subtract λ_j times the original expression. This gives

$$c_1(\lambda_1 - \lambda_j)\underline{v_1} + \cdots + c_{j-1}(\lambda_{j-1} - \lambda_j)\underline{v_{j-1}} = \underline{0}$$

\# as we chose the shortest LD set, and $\lambda_i \neq \lambda_j$ as we said they were distinct. So the proposition is true □

Theorem 45. Let $\alpha : V \to V$ be a linear map with $\dim V = n$ and assume that α has n distinct eigenvalues. Then α is diagonalisable

Proof. Let $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}$ be non-$\underline{0}$ with $\alpha\underline{v_i} = \lambda_i\underline{v_i}$ (so $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the distinct eigenvalues and $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}$ are the corresponding eigenvectors). Then $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}$ are linearly independent by proposition 18, so they form a basis B (as $n = \dim V$). Then

$$[\alpha]_B = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

□

Remark - Theorem 19 gives a sufficient but not necessary condition. For example, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is diagonalisable, but $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not, although both have evalues 1 and 1

Proposition 46. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct eigenvalues of the linear map $\alpha : V \to V$, let B_i be a basis of the corresponding eigenspaces $V_{\lambda_i} = \{\underline{v} \in V : \alpha\underline{v} = \lambda_i\underline{v}\}$. Then $B_1 \cup B_2 \cup \ldots \cup B_k$ is linearly independent. If $\sum |B_i| = \dim V$, then B is a basis and $[\alpha]_B$ is diagonal

Proof. Suppose $\sum_{1 \leq i \leq k} \sum_{\mu \in B_i} c_\mu \mu = \underline{0}$
Put $\underline{v_i} = \sum_{\mu \in B_i} c_\mu \mu$. Then $\underline{v_1} + \cdots + \underline{v_k} = \underline{0}$
But by Proposition 44, each $\underline{v_i} = \underline{0}$, so $\sum_{\mu \in B_i} c_\mu \mu = \underline{0}$
But B_i is a basis, so $c_\mu = 0 \; \forall \mu \in B_1 \cup B_2 \cup \ldots \cup B_k$ □

Note 47. Let V be a vector space over F, and $\alpha : V \to V$ a linear map. A procedure to find out whether or not α is diagonalisable, and to diagonalise where possible is as follows:

> Work out $\chi_\alpha(t) = \det(\alpha - tI) = \det(A - tI)$ for $A = [\alpha]_B$ for some basis B.
> Find all distinct roots of χ_α in F, say $\lambda_1, \ldots, \lambda_k$.
> For each λ_i, find a basis B_i of the eigenspace V_{lambda_i}.
> If $\sum |B_i| = \dim_F V$, then $B = \bigcup_{i=1}^{k} B_i$ is a basis of V and $[\alpha]_B$ is diagonal.
> If $\sum |B_i| < \dim V$ then α is not diagonalisable.

Theorem 48. Let $\alpha : \mathbb{C}^2 \to \mathbb{C}^2$ be linear. For some basis B, the matrix $[\alpha]_B$ is one of:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

for $\lambda, \lambda_i \in \mathbb{C}$. No two of these are conjugate, except $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ is conjugate to $\begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}$ $(\lambda_1 \neq \lambda_2)$

Proof. If α has distinct eigenvalues $\lambda_1 \neq \lambda_2$, then α is diagonalisable, by Theorem 45
So assume $\chi_\alpha(z) = (\lambda - z)^2$, so the eigenvalue is λ with multiplicity 2.
If $\dim V_\lambda = 2$ then $V = V_\lambda$ so second case, and any basis B for V will do.
Assume $\dim V_\lambda = 1$. let $\underline{0} \neq \underline{v} \in V_\lambda$. Let $B' = \{\underline{v}, \underline{w}\}$ with any $\underline{w} \in V \setminus V_\lambda$.
Then $[\alpha]_{B'} = \begin{pmatrix} \lambda & b \\ 0 & \lambda \end{pmatrix} = A$
The first column comes from v being $\in V_\lambda$, and the second as the bottom right has to be an evalue of α, so has to be λ, and so the matrix is upper triangular.
Observe: $(\alpha - \lambda I)^2$ is the 0 map, since $(A - \lambda I)^2$ is the 0 matrix.

$$(A - \lambda I)^2 = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Now replace by $\underline{v}$ by $\underline{u} = (\alpha - \lambda I)\underline{w}$ to get $B = \{\underline{u}, \underline{w}\}$

$$(\alpha - \lambda I)\underline{u} = (\alpha - \lambda I)^2 \underline{w} = 0 \Rightarrow \underline{u} \in V_\lambda$$

and $(\alpha - \lambda I)\underline{w} = \underline{u}$, so $\alpha \underline{u} = \lambda \underline{u}$, $\alpha \underline{w} = \underline{u} + \lambda \underline{w}$
So $[\alpha]_B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$
Finally, no two of the matrices are conjugate, except $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ to $\begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}$ □

The first type $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ are determined by the distinct eigenvalues λ_1 and λ_2 and map to their order.

The second type $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ are determined by the double eigenvalue λ and being diagonal (each matrix here is in fact self-conjugate)
The third type $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ are determined by the double eigenvalue λ and are not diagonalisable.
We can also rephrase Theorem 48 as **Theorem 48′** - Any 2x2 complex matrix is conjugate to one of

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

Which is unique except that $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}$ conjugate

Remarks
① - Canonical (Jordan Normal) Form exists for any dimension. Given $\alpha : V \to V$ linear, with V a complex space of dimension n, there is a basis B of V such that:

$$\begin{pmatrix} a_{11} & * & 0 & \dots & \dots & 0 \\ 0 & a_{22} & * & \ddots & \dots & 0 \\ 0 & 0 & a_{33} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & a_{(n-1)(n-1)} & * \\ 0 & 0 & 0 & \dots & \ddots & a_{nn} \end{pmatrix} \text{ with } * = 1, 0$$

$$[\alpha]_B = A = (a_{ij}) \text{ with } \begin{matrix} a_{ii} & \text{eigenvalues} \\ a_{i(i+1)} & \in \{0, 1\} \\ a_{ij} & 0 \text{ otherwise} \end{matrix}$$

Over $\mathbb{R}$ and even $\mathbb{R}^2$ it is more complicated.
② It follows from 48' that for $n = 2, \chi_A(A) = 0$ for all 2x2 complex matrices A
$\chi_A(z) = z^2 + c_1 z + c_0 \quad \chi_A(A) = A^2 + c_1 A + c_0 I = 0$ matrix

③ Over $\mathbb{C}$, $\chi_\alpha(z) = \prod_{i=1}^{k} (\lambda_i - z)^{e_i}$
$\lambda_1, \lambda_2, \dots, \lambda_k$ are all the distinct evalues of α
$e_1, e_2, \dots, e_k$ are their algebraic multiplicities.
The geometric multiplicity of λ_j is $\dim V_{\lambda_j}$. In general

$$1 \leq g_j \leq e_j \ \forall j$$

By Proposition 46 α is diagonalisable iff $e_j = g_j \ \forall j$

Example 49. $\boldsymbol{A} = \begin{pmatrix} \mathbf{0} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} \end{pmatrix}$ over $\mathbb{R}$ i.e. $\boldsymbol{a_{ij}} = \mathbf{1} - \boldsymbol{\delta_{ij}}$. Now $\boldsymbol{A} + \boldsymbol{I}$ is of rank 1, so -1 is an eigenvalue (of multiplicity 2

$$\boldsymbol{V_{-1}} = \left\langle \begin{pmatrix} \mathbf{1} \\ \mathbf{-1} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{-1} \end{pmatrix} \right\rangle$$

As $\mathbf{tr}\, \boldsymbol{A} = \mathbf{0}$ $\sum$evalues $= \mathbf{0}$ $\therefore$ remaining eigenvalue$= \mathbf{2}$

$$\boldsymbol{V_{-1}} = \left\langle \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} \right\rangle$$

Put $\boldsymbol{P} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{1} \\ \mathbf{-1} & \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{-1} & \mathbf{1} \end{pmatrix}$, the change of basis matrix from the standard to the eigenvectors then $\boldsymbol{P^{-1}AP} = \begin{pmatrix} \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{2} \end{pmatrix}$

9.1 Real symmetrix and orthogonal matrices

Let $\boldsymbol{A}$ be a real $\boldsymbol{n}$x$\boldsymbol{n}$ matrix. It is symmetric if $\boldsymbol{a_{ij} = a_{ji}}\ \forall \boldsymbol{i \neq j}$ so $\boldsymbol{A^T = A}$
$\boldsymbol{A}$ is orthogonal if $\boldsymbol{A^T = A^{-1}}$, so $\boldsymbol{AA^T = A^TA = I}$
The main theorem of this section is

Theorem 50. A real symmetric matrix $\boldsymbol{A}$ is diagonalisable over $\mathbb{R}$. Moreover, there is a (real) orthogonal matrix $\boldsymbol{P}$ such that $\boldsymbol{P^{-1}AP}$ is diagonal - or equally $\boldsymbol{P^TAP}$ which is easier to find

Remark Scalar product (or dot product) of $\underline{\boldsymbol{x}}, \underline{\boldsymbol{y}} \in \mathbb{R}^{\boldsymbol{n}}$ with coordinates $\boldsymbol{x_i}, \boldsymbol{y_i}$

$$\underline{\boldsymbol{x}}.\underline{\boldsymbol{y}} = \sum \boldsymbol{x_i y_i}$$

or on $\mathbb{C}^{\boldsymbol{n}}$

$$\underline{\boldsymbol{x}}.\underline{\boldsymbol{y}} = \sum \overline{\boldsymbol{x_i}} \boldsymbol{y_i}$$

Note 51. In terms of column matrices $\underline{\boldsymbol{x}}.\underline{\boldsymbol{y}} = \boldsymbol{x^T y}$, and the length of $\underline{\boldsymbol{x}}$ is $\sqrt{\underline{\boldsymbol{x}}.\underline{\boldsymbol{x}}}$. $\underline{\boldsymbol{x}}$ is a unit vector if $|\underline{\boldsymbol{x}}| = \boldsymbol{1}$
$\underline{\boldsymbol{v_1}}, \underline{\boldsymbol{v_2}}, \ldots, \underline{\boldsymbol{v_k}}$ is an orthogonal set of vectors if $\underline{\boldsymbol{v_i}}.\underline{\boldsymbol{v_j}} = \boldsymbol{0}\ \forall \boldsymbol{i \neq j}$
It is orthonormal if each vector is a unit vector.

Lemma 52. Any set of $\boldsymbol{n}$ orthonormal vectors in $\mathbb{R}^{\boldsymbol{n}}$ is a basis for $\mathbb{R}^{\boldsymbol{n}}$

Proof. Enugh to show that any set $\underline{\boldsymbol{v_1}}, \underline{\boldsymbol{v_2}}, \ldots, \underline{\boldsymbol{v_k}}$ of non-zero orthonormal vectors is linearly independent.
If $\sum \boldsymbol{c_j} \underline{\boldsymbol{v_j}} = \underline{\boldsymbol{0}}$ then $\boldsymbol{0} = (\sum \boldsymbol{c_j} \underline{\boldsymbol{v_j}}.\underline{\boldsymbol{v_i}}) = \boldsymbol{c_i}(\underline{\boldsymbol{v_i}}.\underline{\boldsymbol{v_i}})$
Now $\underline{\boldsymbol{v_i}}.\underline{\boldsymbol{v_i}} \neq \boldsymbol{0}$ as $\underline{\boldsymbol{v_i}} \neq \underline{\boldsymbol{0}}$ so $\boldsymbol{c_i = 0}\ \forall \boldsymbol{i}$ □

Example: The standard basis for $\mathbb{R}^{\boldsymbol{n}}$ is orthonormal.
Note: A matrix is orthonomal iff its column are orthonormal

Let $\boldsymbol{B' = \{\underline{v}, \underline{w}\}}, \underline{\boldsymbol{v}} \in \boldsymbol{V_\lambda}\ \underline{\boldsymbol{w}} \in \boldsymbol{V \setminus V_\lambda}$
$[\boldsymbol{\alpha}]_{\boldsymbol{b'}} = \begin{pmatrix} \boldsymbol{\lambda} & \boldsymbol{b} \\ \boldsymbol{0} & \boldsymbol{\lambda} \end{pmatrix}, \boldsymbol{b \neq 0}, \boldsymbol{\alpha}(\underline{\boldsymbol{w}}) = \boldsymbol{b}\underline{\boldsymbol{v}} + \boldsymbol{\lambda}\underline{\boldsymbol{w}}$
Replace $\underline{\boldsymbol{v}}$ by $\underline{\boldsymbol{u}} = \boldsymbol{b}\underline{\boldsymbol{v}}$
put $\boldsymbol{B = \{\underline{u}, \underline{w}\}}$ then $[\boldsymbol{\alpha}]_{\boldsymbol{B}} = \begin{pmatrix} \boldsymbol{\lambda} & \boldsymbol{1} \\ \boldsymbol{0} & \boldsymbol{\lambda} \end{pmatrix}$

9.2 Real Symmetric Matrices

Lemma 53. The eigenvalues of a real symmetric matrix are real

Proof. Let $\boldsymbol{A}$ be a real symmetric matrix. Let $\boldsymbol{\lambda}$ be an eigenvalue of $\boldsymbol{A}$ in $\mathbb{C}$. Will show that $\boldsymbol{\lambda} \in \mathbb{R}$ by showing $\overline{\boldsymbol{\lambda}} = \boldsymbol{\lambda}$
Let $\underline{\boldsymbol{v}} \in \mathbb{C}^{\boldsymbol{n}}$ be a complex (column) eigenvector of $\boldsymbol{A}$ with $\boldsymbol{A}\underline{\boldsymbol{v}} = \boldsymbol{\lambda}\underline{\boldsymbol{v}}$. Now because $\boldsymbol{A}$ is a real

$$\boldsymbol{A}\overline{\underline{\boldsymbol{v}}} = \overline{\boldsymbol{A}}\,\overline{\underline{\boldsymbol{v}}} = \overline{\boldsymbol{A}\underline{\boldsymbol{v}}} = \overline{\boldsymbol{\lambda}\underline{\boldsymbol{v}}} = \overline{\boldsymbol{\lambda}}\,\overline{\underline{\boldsymbol{v}}}$$

Hence

$$\begin{aligned}\boldsymbol{\lambda}\underbrace{\overline{\underline{\boldsymbol{v}}}^{\boldsymbol{T}}\underline{\boldsymbol{v}}}_{\neq \boldsymbol{0}} &= \overline{\underline{\boldsymbol{v}}}^{\boldsymbol{T}}(\boldsymbol{\lambda}\underline{\boldsymbol{v}}) = \overline{\underline{\boldsymbol{v}}}^{\boldsymbol{T}}(\boldsymbol{A}\underline{\boldsymbol{v}}) = (\overline{\underline{\boldsymbol{v}}}^{\boldsymbol{T}}\boldsymbol{A})\underline{\boldsymbol{v}} = (\boldsymbol{A}^{\boldsymbol{T}}\overline{\underline{\boldsymbol{v}}})^{\boldsymbol{T}}\underline{\boldsymbol{v}} \\ &= (\boldsymbol{A}\overline{\underline{\boldsymbol{v}}})^{\boldsymbol{T}}\underline{\boldsymbol{v}} \because \boldsymbol{A} \text{ symmetric from above} \\ &= (\overline{\boldsymbol{\lambda}}\overline{\underline{\boldsymbol{v}}})^{\boldsymbol{T}}\underline{\boldsymbol{v}} = \overline{\boldsymbol{\lambda}}\,\overline{\underline{\boldsymbol{v}}}^{\boldsymbol{T}}\underline{\boldsymbol{v}}\end{aligned}$$

Hence $\boldsymbol{\lambda} = \overline{\boldsymbol{\lambda}}$ since $\overline{\underline{\boldsymbol{v}}}^{\boldsymbol{T}}\underline{\boldsymbol{v}} \neq \boldsymbol{0}$ as $\underline{\boldsymbol{v}} \neq \boldsymbol{0}$ □

Lemma 54. The eigenvectors corresponding to distinct eigenvalues of the real symmetrix matrix $\boldsymbol{A}$ are orthogonal

Proof. Let $\boldsymbol{\lambda} \neq \boldsymbol{\mu}$ be distinct eigenvalues of $\boldsymbol{A}$ with $\underline{\boldsymbol{v}}, \underline{\boldsymbol{w}}$ be corresponding eigenvectors such that $\boldsymbol{A}\underline{\boldsymbol{v}} = \boldsymbol{\lambda}\underline{\boldsymbol{v}}, \boldsymbol{A}\underline{\boldsymbol{w}} = \boldsymbol{\lambda}\underline{\boldsymbol{w}}$. Then

$$\begin{aligned}\underline{\boldsymbol{v}}^{\boldsymbol{T}}(\boldsymbol{A}\underline{\boldsymbol{w}}) &= \underline{\boldsymbol{v}}^{\boldsymbol{T}}(\boldsymbol{\mu}\underline{\boldsymbol{w}}) = \boldsymbol{\mu}\underline{\boldsymbol{v}}^{\boldsymbol{T}}\underline{\boldsymbol{w}} \\ &= (\underline{\boldsymbol{v}}^{\boldsymbol{T}}\boldsymbol{A})\underline{\boldsymbol{w}} = (\boldsymbol{A}^{\boldsymbol{T}}\underline{\boldsymbol{v}})^{\boldsymbol{T}}\underline{\boldsymbol{w}} = (\boldsymbol{A}\underline{\boldsymbol{v}})^{\boldsymbol{T}}\underline{\boldsymbol{w}} = (\boldsymbol{\lambda}\underline{\boldsymbol{v}})^{\boldsymbol{T}}\underline{\boldsymbol{w}} = \boldsymbol{\lambda}\underline{\boldsymbol{v}}^{\boldsymbol{T}}\underline{\boldsymbol{w}}\end{aligned}$$

But $\boldsymbol{\mu} \neq \boldsymbol{\lambda} \therefore \underline{\boldsymbol{v}}^{\boldsymbol{T}}\underline{\boldsymbol{w}} = \boldsymbol{0} \therefore \underline{\boldsymbol{v}}, \underline{\boldsymbol{w}}$ are orthogonal □

Proposition 55. (Special case of Theorem 50, p. 105, compare with Theorem 45, p. 101)
If the real symmetric $\boldsymbol{n}$x$\boldsymbol{n}$ matrix $\boldsymbol{A}$ has $\boldsymbol{n}$ distinct eigenvalues there exists an orthogonal amtrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-\boldsymbol{1}}\boldsymbol{A}\boldsymbol{P}$ is diagonal $(=\boldsymbol{P}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{P})$

Proof. Let $\underline{\boldsymbol{v_1}}, \underline{\boldsymbol{v_2}}, \ldots, \underline{\boldsymbol{v_n}}$ be a set of (non-$\underline{\boldsymbol{0}}$) e-vectors corresponding to the $\boldsymbol{n}$ distinct e-values $\boldsymbol{\lambda_1}, \boldsymbol{\lambda_2}, \ldots, \boldsymbol{\lambda_n}$
Normalise each $\underline{\boldsymbol{v_i}}$, so replace each $\underline{\boldsymbol{v_i}}$ by $\frac{\underline{\boldsymbol{v_i}}}{|\underline{\boldsymbol{v_i}}|}$ so each $\underline{\boldsymbol{v_i}}$ is a unit vector. Then $\boldsymbol{B} = \{\underline{\boldsymbol{v_i}}\}$ is an orthonormal basis. If $\boldsymbol{P}$ is obtained by taking $\underline{\boldsymbol{v_i}}$ as the $\boldsymbol{i}^{\text{th}}$ column then $\boldsymbol{P}^{\boldsymbol{T}}\boldsymbol{P} = \boldsymbol{I}$ as $\underline{\boldsymbol{v_i}}^{\boldsymbol{T}}\underline{\boldsymbol{v_j}} = \boldsymbol{\delta_{ij}}$
And $\boldsymbol{P}^{-\boldsymbol{1}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{\lambda_1} & \ldots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \ldots & \boldsymbol{\lambda_n} \end{pmatrix}$ diagonal
Also can check directly:

$$\begin{aligned}\boldsymbol{A}\boldsymbol{P} &= \boldsymbol{A}(\underline{\boldsymbol{v_1}} \mid \underline{\boldsymbol{v_2}} \mid \ldots \mid \underline{\boldsymbol{v_n}}) \\ &= (\boldsymbol{A}\underline{\boldsymbol{v_1}} \mid \boldsymbol{A}\underline{\boldsymbol{v_2}} \mid \ldots \mid \boldsymbol{A}\underline{\boldsymbol{v_n}}) \\ &= (\boldsymbol{\lambda_1}\underline{\boldsymbol{v_1}} \mid \boldsymbol{\lambda_2}\underline{\boldsymbol{v_2}} \mid \ldots \mid \boldsymbol{\lambda_n}\underline{\boldsymbol{v_n}})\end{aligned}$$

So

$$\boldsymbol{P^TAP} = \begin{pmatrix} \underline{\boldsymbol{v_1}}^{\boldsymbol{T}} \\ \hline \\ \vdots \\ \hline \\ \underline{\boldsymbol{v_n}}^{\boldsymbol{T}} \end{pmatrix} (\boldsymbol{\lambda_1}\underline{\boldsymbol{v_1}} \mid \boldsymbol{\lambda_2}\underline{\boldsymbol{v_2}} \mid \ldots \mid \boldsymbol{\lambda_n}\underline{\boldsymbol{v_n}})$$

$$= \begin{pmatrix} \boldsymbol{\lambda_1} & \ldots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \ldots & \boldsymbol{\lambda_n} \end{pmatrix}$$

as $\underline{\boldsymbol{v_i}}^{\boldsymbol{T}}\underline{\boldsymbol{v_j}} = \boldsymbol{\delta_{ij}}$ □

Back to Example 49: $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{1} & \boldsymbol{1} \\ \boldsymbol{1} & \boldsymbol{0} & \boldsymbol{1} \\ \boldsymbol{1} & \boldsymbol{1} & \boldsymbol{0} \end{pmatrix}, \boldsymbol{V_{-1}} = \left\langle \begin{pmatrix} \boldsymbol{1} \\ \boldsymbol{-1} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{1} \\ \boldsymbol{0} \\ \boldsymbol{-1} \end{pmatrix} \right\rangle, \boldsymbol{V_2} = \left\langle \begin{pmatrix} \boldsymbol{1} \\ \boldsymbol{1} \\ \boldsymbol{1} \end{pmatrix} \right\rangle$ Note: $\underline{\boldsymbol{v_1}} \perp \underline{\boldsymbol{v_2}}$.

To find $\boldsymbol{P}$ orthogonal such that $\boldsymbol{P^{-1}AP} = \begin{pmatrix} \boldsymbol{-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{2} \end{pmatrix}$ need to find an orthonormal basis of eigenvectors of $\boldsymbol{A}$

In $\boldsymbol{V_2}$ just normalise to get $\begin{pmatrix} \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{3}}} \\ \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{3}}} \\ \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{3}}} \end{pmatrix}$

In $\boldsymbol{V_{-1}}$ we can use the GRAMM-SCHMIDT process, a very special case!

$\underline{\boldsymbol{u_1}} = \begin{pmatrix} \boldsymbol{1} \\ \boldsymbol{-1} \\ \boldsymbol{0} \end{pmatrix}, \underline{\boldsymbol{u_2}} = \begin{pmatrix} \boldsymbol{1} \\ \boldsymbol{0} \\ \boldsymbol{-1} \end{pmatrix}, \underline{\boldsymbol{v_1}} = \underline{\boldsymbol{u_1}}$ and $\underline{\boldsymbol{v_2}} = \underline{\boldsymbol{u_2}} + \boldsymbol{a_1}\underline{\boldsymbol{v_1}}$

But $\underline{\boldsymbol{v_1}}.\underline{\boldsymbol{v_2}} = (\underline{\boldsymbol{v_1}}.\underline{\boldsymbol{u_2}}) + \boldsymbol{a_1}(\underline{\boldsymbol{v_1}}.\underline{\boldsymbol{v_1}})$, so as we want $\underline{\boldsymbol{v_1}}.\underline{\boldsymbol{v_2}}$ set $\boldsymbol{a} = -\frac{\underline{\boldsymbol{v_1}}.\underline{\boldsymbol{u_2}}}{\underline{\boldsymbol{v_1}}.\underline{\boldsymbol{v_1}}} = -\frac{\boldsymbol{1}}{\boldsymbol{2}}$.

Now after normalising

$$\boldsymbol{P} = \begin{pmatrix} \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{2}}} & \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{6}}} & \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{3}}} \\ -\frac{\boldsymbol{1}}{\sqrt{\boldsymbol{2}}} & \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{6}}} & \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{3}}} \\ \boldsymbol{0} & -\frac{\boldsymbol{2}}{\sqrt{\boldsymbol{6}}} & \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{3}}} \end{pmatrix}$$

and therefore $\boldsymbol{P^TP} = \boldsymbol{I}$

$$\boldsymbol{P^TAP} = \begin{pmatrix} \boldsymbol{-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{2} \end{pmatrix}$$

Theorem 56. Let $\boldsymbol{A}$ be a real $\boldsymbol{n}$x$\boldsymbol{n}$ matrix. The following are equivalent:
i) $\boldsymbol{A}$ is an orthogonal matrix - so $\boldsymbol{A^TA} = \boldsymbol{I} = \boldsymbol{AA^T}$

ii) $|\boldsymbol{A}\underline{\boldsymbol{v}}| = |\underline{\boldsymbol{v}}|\ \forall \underline{\boldsymbol{v}}$ = column vectors (so $\boldsymbol{A}$ is a linear isometry)
iii) $\boldsymbol{A}\underline{\boldsymbol{v}}.\boldsymbol{A}\underline{\boldsymbol{w}} = \underline{\boldsymbol{v}}.\underline{\boldsymbol{w}}\ \forall \underline{\boldsymbol{v}}, \underline{\boldsymbol{w}}$ column vectors i.e. angles are preserved
iv) If $\underline{\boldsymbol{v_1}}, \underline{\boldsymbol{v_2}}, \ldots, \underline{\boldsymbol{v_n}}$ are orthonormal, then so are $\boldsymbol{A}\underline{\boldsymbol{v_1}}, \boldsymbol{A}\underline{\boldsymbol{v_2}}, \ldots, \boldsymbol{A}\underline{\boldsymbol{v_n}}$ i.e. $\boldsymbol{A}$ takes one orthonormal basis to another.
v) The columns of $\boldsymbol{A}$ are orthonormal

Proof. i) $\Rightarrow$ ii) Assume $\boldsymbol{AA^T} = \boldsymbol{I} = \boldsymbol{A^TA}$
$|\boldsymbol{A}\underline{\boldsymbol{v}}|^2 = (\boldsymbol{A}\underline{\boldsymbol{v}})^T(\boldsymbol{A}\underline{\boldsymbol{v}}) = \underline{\boldsymbol{v}}^T\boldsymbol{A}^T\boldsymbol{A}\underline{\boldsymbol{v}} = \underline{\boldsymbol{v}}^T\boldsymbol{I}\underline{\boldsymbol{v}} = \underline{\boldsymbol{v}}^T\underline{\boldsymbol{v}} = |\underline{\boldsymbol{v}}|^2$
$\therefore |\boldsymbol{A}\underline{\boldsymbol{v}}| = |\underline{\boldsymbol{v}}|$ since both are non-negative
ii) $\Rightarrow$ iii) Assume $|\boldsymbol{A}\underline{\boldsymbol{v}}| = |\underline{\boldsymbol{v}}|\ \forall \underline{\boldsymbol{v}}$

$$|\underline{\boldsymbol{v}} + \underline{\boldsymbol{w}}|^2 = (\underline{\boldsymbol{v}} + \underline{\boldsymbol{w}})^T(\underline{\boldsymbol{v}} + \underline{\boldsymbol{w}}) = \underline{\boldsymbol{v}}^T\underline{\boldsymbol{v}} + \underline{\boldsymbol{v}}^T\underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}^T\underline{\boldsymbol{v}} + \underline{\boldsymbol{w}}^T\underline{\boldsymbol{w}} = |\underline{\boldsymbol{v}}|^2 + 2\underline{\boldsymbol{v}}^T\underline{\boldsymbol{w}} + |\underline{\boldsymbol{w}}|^2$$
$$= |\boldsymbol{A}(\underline{\boldsymbol{v}} + \underline{\boldsymbol{w}})|^2 = |\boldsymbol{A}\underline{\boldsymbol{v}} + \boldsymbol{A}\underline{\boldsymbol{w}}|^2 = |\boldsymbol{A}\underline{\boldsymbol{v}}|^2 + 2(\boldsymbol{A}\underline{\boldsymbol{v}})^T\boldsymbol{A}\underline{\boldsymbol{w}} + |\boldsymbol{A}\underline{\boldsymbol{w}}|^2$$

(we have used that, although normally $(\underline{\boldsymbol{w}}^T\underline{\boldsymbol{v}})^T = \underline{\boldsymbol{v}}^T\underline{\boldsymbol{w}}$, as $\underline{\boldsymbol{w}}^T\underline{\boldsymbol{v}}$ is only a number $(\underline{\boldsymbol{w}}^T\underline{\boldsymbol{v}})^T = \underline{\boldsymbol{w}}^T\underline{\boldsymbol{v}}$)
So therefore, as $|\boldsymbol{A}\underline{\boldsymbol{v}}| = |\underline{\boldsymbol{v}}|, \boldsymbol{A}\underline{\boldsymbol{v}}.\boldsymbol{A}\underline{\boldsymbol{w}} = (\boldsymbol{A}\underline{\boldsymbol{v}})^T\boldsymbol{A}\underline{\boldsymbol{w}} = \underline{\boldsymbol{v}}^T\underline{\boldsymbol{w}} = \underline{\boldsymbol{v}}.\underline{\boldsymbol{w}}$
iii) $\Rightarrow$ iv) Assume iii), iv) follows immediately as a special case $\underline{\boldsymbol{v_i}}.\underline{\boldsymbol{v_j}} = \boldsymbol{\delta_{ij}} \Rightarrow \boldsymbol{A}\underline{\boldsymbol{v_i}}.\boldsymbol{A}\underline{\boldsymbol{v_j}} = \boldsymbol{\delta_{ij}}$
iv) $\Rightarrow$ v) The standard basis is orthonormal, hence by the assumption so is $\boldsymbol{A}\begin{pmatrix}\boldsymbol{1}\\\boldsymbol{0}\\\vdots\\\boldsymbol{0}\end{pmatrix}, \boldsymbol{A}\begin{pmatrix}\boldsymbol{0}\\\boldsymbol{1}\\\vdots\\\boldsymbol{0}\end{pmatrix}, \ldots, \boldsymbol{A}\begin{pmatrix}\boldsymbol{0}\\\boldsymbol{0}\\\vdots\\\boldsymbol{1}\end{pmatrix}$, but these are the column vectors of $\boldsymbol{A}$, so the column vectors of $\boldsymbol{A}$ are orthonormal
v) $\Rightarrow$ i) If the columns of $\boldsymbol{A}$ are orthonormal, then $\boldsymbol{A^TA} = \boldsymbol{I}$. Since $\boldsymbol{A}$ is a square matrix, also $\boldsymbol{AA^T} = \boldsymbol{I}$. So $\boldsymbol{A}$ is orthogonal. □

Lemma 57. Let $\boldsymbol{A}$ be a real symmetric matrix, let $\underline{\boldsymbol{v}}$ be en eigenvector of $\boldsymbol{A}$. Let $\underline{\boldsymbol{v}}^\perp = \{\underline{\boldsymbol{w}} : \underline{\boldsymbol{v}}^T\underline{\boldsymbol{w}} = \boldsymbol{0}\}$
If $\underline{\boldsymbol{w}} \in \underline{\boldsymbol{v}}^\perp$ then $\boldsymbol{A}\underline{\boldsymbol{w}} \in \underline{\boldsymbol{v}}^\perp$

Proof. Let $\underline{\boldsymbol{w}} \in \underline{\boldsymbol{v}}^\perp$, so $\underline{\boldsymbol{v}}^T\underline{\boldsymbol{w}} = \boldsymbol{0}$. Then
$\underline{\boldsymbol{v}}^T\boldsymbol{A}\underline{\boldsymbol{w}} = (\boldsymbol{A}^T\underline{\boldsymbol{v}})^T\underline{\boldsymbol{w}} = (\boldsymbol{A}\underline{\boldsymbol{v}})^T\underline{\boldsymbol{w}} = (\boldsymbol{\lambda}\underline{\boldsymbol{v}})^T\underline{\boldsymbol{w}} = \boldsymbol{0}$ as $\underline{\boldsymbol{v}}^T\underline{\boldsymbol{w}} = \boldsymbol{0}$ □

This is behind the proof, not given here, of Theorem 50, p. 105.

If $\boldsymbol{V}$ has subspaces $\boldsymbol{U}$ and $\boldsymbol{W}$ such that $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{W}, \boldsymbol{U} \cap \boldsymbol{W} = \{\underline{\boldsymbol{0}}\}$ and $\boldsymbol{\alpha} : \boldsymbol{V} \to \boldsymbol{V}$ is linear with $\boldsymbol{\alpha}(\boldsymbol{U}) \subset \boldsymbol{U}, \boldsymbol{\alpha W} \subset \boldsymbol{W}$, take a basis $\boldsymbol{B}$ with $\boldsymbol{B} = \boldsymbol{B_1} \cup \boldsymbol{B_2}$ of bases $\boldsymbol{B_1}$ of $\boldsymbol{U}$ and $\boldsymbol{B_2}$ of $\boldsymbol{W}$. Then

$$[\boldsymbol{\alpha}]_{\boldsymbol{B}} = \begin{pmatrix}\boldsymbol{*} & \boldsymbol{0}\\\boldsymbol{0} & \boldsymbol{*}\end{pmatrix}$$

9.3 Quadratic Forms and Quadrics

Let $\boldsymbol{A}$ be a real symmetric $\boldsymbol{n}$x$\boldsymbol{n}$ matrix. Then the map

$$\boldsymbol{Q}(\underline{v}) = \underline{v}^{\boldsymbol{T}} \boldsymbol{A} \underline{v}$$

$\boldsymbol{Q}: \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}$ is a real quadratic form on $\mathbb{R}^{\boldsymbol{n}}$

If $\boldsymbol{A} = (\boldsymbol{a_{ij}}), \underline{v} = \begin{pmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \end{pmatrix}$ then

$$\boldsymbol{Q}(\underline{v}) = \sum_{ij} \boldsymbol{a_{ij} x_i x_j}$$

e.g. $\boldsymbol{n = 2}, \underline{v} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}$ $\boldsymbol{Q}(\underline{v}) = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}^{\boldsymbol{T}} \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \boldsymbol{ax^2 + 2bxy + cy^2}$
Conversely, given such function $\boldsymbol{ax^2 + 2bxy + cy^2}$ this can be written as $\boldsymbol{Q}(\underline{v}) = \underline{v}^{\boldsymbol{T}} \boldsymbol{A} \underline{v}$ with $\boldsymbol{A}$ as a suitable real symmetric matrix

Example: $\boldsymbol{2x^2 - 4xy + 5y^2}$ $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{2} & \boldsymbol{-2} \\ \boldsymbol{-2} & \boldsymbol{5} \end{pmatrix}, \underline{v} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}$

If we change coordinates: $\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix}$ with $\boldsymbol{P}$ being the change of basis matrix from the standard to a new one.
Then $\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}^{\boldsymbol{T}} \boldsymbol{A} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix}^{\boldsymbol{T}} \boldsymbol{P^T A P} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix}$, so the matrix of the form becomes $\boldsymbol{P^T A P}$ with respect to the new coordinate system.
Now choose $\boldsymbol{P}$ to be an orthogonal amtrix with columns being the eigenvectors of $\boldsymbol{A}$ (see Theorem 50, p. 105). Then $\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}^{\boldsymbol{T}} \boldsymbol{A} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix}^{\boldsymbol{T}} \begin{pmatrix} \boldsymbol{\lambda} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\mu} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix} = \boldsymbol{\lambda \xi^2 + \mu \eta^2}$ where $\boldsymbol{\lambda, \mu}$ are the eigenvalues of $\boldsymbol{A}$. For matrices of size $\boldsymbol{n}$ we get $\sum_{i=1}^{n} \boldsymbol{\lambda_i \xi_i^2}$

Principal Axes - The orthogonal frame with respect to which the form is diagonal i.e. where the basis is made of unit eigenvectors of $\boldsymbol{A}$

To continue the example, the eigenvalues are 1 and 6 with eigenvectors $\begin{pmatrix} \boldsymbol{2} \\ \boldsymbol{1} \end{pmatrix}, \begin{pmatrix} \boldsymbol{1} \\ \boldsymbol{-1} \end{pmatrix}$ - normalise to get $\boldsymbol{P} = \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{5}}} \begin{pmatrix} \boldsymbol{2} & \boldsymbol{1} \\ \boldsymbol{1} & \boldsymbol{-2} \end{pmatrix}$. The form becomes $\boldsymbol{\xi^2 + 6\eta^2}$
If $\boldsymbol{Q}$ is a quadratic form, the equation $\boldsymbol{Q}(\underline{v}) = \boldsymbol{1}$ is quadratic in $\mathbb{R}^{\boldsymbol{n}}$

Example: In $\boldsymbol{n = 2}$

$\boldsymbol{\lambda \xi^2 + \mu \eta^2}$	$= 1$
$\boldsymbol{\lambda > 0, \mu > 0}$	$\Rightarrow$ ellipse (with semi-axes $\frac{\boldsymbol{1}}{\sqrt{\boldsymbol{\lambda}}}, \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{\mu}}}$)
$\boldsymbol{\lambda > 0, \mu < 0}$	$\Rightarrow$ hyperbola meeting $\boldsymbol{x}$-axis at $\pm \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{2}}}$

Example: In $\boldsymbol{n = 3}$

$\boldsymbol{\lambda\xi^2 + \mu\eta^2 + \nu\zeta^2} = 1$
$\boldsymbol{\lambda > 0, \mu > 0, \nu > 0} \Rightarrow$ ellipsoid (squashed sphere)
$\boldsymbol{\lambda > 0, \mu > 0, \nu < 0} \Rightarrow$ hyperboloid of one sheet
$\boldsymbol{\lambda > 0, \mu < 0, \nu < 0} \Rightarrow$ hyperboloid of two sheets
$\boldsymbol{\lambda > 0, \mu > 0, \nu = 0} \Rightarrow$ elliptic cylinder

Another application: HESSIAN of a function $\boldsymbol{f}$ in two variables

$$\begin{pmatrix} \frac{\delta^2 f}{\delta x^2} & \frac{\delta^2 f}{\delta x \delta y} \\ \frac{\delta^2 f}{\delta y \delta x} & \frac{\delta^2 f}{\delta y^2} \end{pmatrix}$$ often symmetric

Used to investigate turning points of $\boldsymbol{f}$

9.4 More on real orthogonal matrices

Lemma 58. If λ is an eigenvalue of a real orthogonal matrix then $|\lambda| = 1$. If λ is real then $\lambda = \pm 1$

Proof. If $A\underline{v} = \lambda\underline{v}$, recalling $|A\underline{w}| = |\underline{w}|\ \forall \underline{w}$ we have $|lambda\underline{v}| = |\underline{v}|$ so $|\lambda| = 1$. If $\lambda \in \mathbb{R}, \lambda = \pm 1$ □

Lemma 59. If A is real orthogonal, then $\det A = \pm 1$

Proof. $AA^T = 1 \Rightarrow \det A . \det A^T = 1 \therefore (\det A)^2 = 1 \therefore \det A = \pm 1$ □

Notation: $O_n = O(n) = O(\mathbb{R}^n) = \{An\text{x}n \text{ real } |AA^T = I\}$
$SO_n = \{A \in O(\mathbb{R}^n) | \det A = 1\}$

Proposition 60. i) Any matrix in SO_2 is $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ for some $\theta \in [0, 2\pi]$ - rotation

ii) Any matrix in $O_2 \setminus SO_2$ is conjugate (in O_2) to $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ a reflection

Proof. i) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; A \in SO_2 \Rightarrow A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$
So $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$
Now $\det A = 1 \Rightarrow a^2 + b^2 = 1$
so (a, b) is a point on the unit circle, and therefore $(a, b) = (\cos\theta, \sin\theta)$ for a unique $\theta \in [0, 2\pi]$

ii) If $A \in O_2 \setminus SO_2$ then $\det A = -1$. Then $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} A \in SO_2$, so $A = \begin{pmatrix} a & -b \\ -b & a \end{pmatrix}$. This is symmetric so has real eigenvalues and these are ± 1. Since $\det A = -1$, have eigenvalues 1 and -1
So can conjugate in O_2 to $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ □

Now for $n = 3$

Proposition 61. i) If $A \in SO_3$ then A is conjugate (in O_3) to $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$
ii) If $A \in O_3 \setminus SO_3$, then A is conjugate to $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$

Proof. i) Claim 1 is an eigenvalue

There exists a real eigenvalue (as the characteristic polynomial is a real cubic)
This must be $\pm\mathbf{1}$
The possibilities for eigenvalues (in $\mathbb{C}$) are 1, 1, 1; 1, -1, -1 or 1, $\boldsymbol{\lambda}, \overline{\boldsymbol{\lambda}}$ ($\boldsymbol{\lambda} \in \mathbb{C} \setminus \mathbb{R}$) as $\mathbf{det} = \pm\mathbf{1}$

Let $\underline{\boldsymbol{v}} \in \boldsymbol{V}$ with $|\underline{\boldsymbol{v}}| = \mathbf{1}, \boldsymbol{A}\underline{\boldsymbol{v}} = \underline{\boldsymbol{v}}$. Find $\boldsymbol{B}\{\underline{\boldsymbol{v_1}}, \underline{\boldsymbol{v_2}}, \underline{\boldsymbol{v_3}}\}$ and orthonormal basis
Now $\underline{\boldsymbol{v}}^{\perp} - < \underline{\boldsymbol{v_2}}, \underline{\boldsymbol{v_3}} >$ is $\boldsymbol{A}$-invariant (by Theorem 56 iii), page 107)
So, for suitable $\boldsymbol{P}$ orthogonal, we have

$$\boldsymbol{P^{-1}AP} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B_{11}} & \boldsymbol{B_{12}} \\ \mathbf{0} & \boldsymbol{B_{21}} & \boldsymbol{B_{22}} \end{pmatrix}$$

with $\boldsymbol{B}$ a 2x2 square matrix $\boldsymbol{SO_2}$
Hence $\boldsymbol{P^{-1}AP} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\cos\theta} & \mathbf{\sin\theta} \\ \mathbf{0} & \mathbf{-\sin\theta} & \mathbf{\cos\theta} \end{pmatrix}$ by Proposition 60 i)
ii) If $\boldsymbol{A} \in \boldsymbol{O_3} \setminus \boldsymbol{SO_3}$, observe $-\boldsymbol{I} \in \boldsymbol{O_3} \setminus \boldsymbol{SO_3}$, so $-\boldsymbol{A} \in \boldsymbol{SO_3}$
so $\boldsymbol{A}$ is conjugate to $\begin{pmatrix} \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{-\cos\theta} & \mathbf{-\sin\theta} \\ \mathbf{0} & \mathbf{\sin\theta} & \mathbf{-\cos\theta} \end{pmatrix}$, put $\boldsymbol{\theta} = \boldsymbol{\theta} + \boldsymbol{\pi}$ □

Lemma 62. i) $\boldsymbol{I} \in \boldsymbol{(S)O_n}$
ii) If $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{(S)O_n}$ then so are $\boldsymbol{AB}$ and $\boldsymbol{A^{-1}}$

Proof. i) $\boldsymbol{II^T} = \boldsymbol{I} \quad \mathbf{det}\, \boldsymbol{I} = \mathbf{1} \quad \therefore \boldsymbol{I} \in \boldsymbol{(S)O_n}$
ii) Take $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{O_n}$
$\boldsymbol{AB(AB)^T} = \boldsymbol{ABB^TA^T} = \boldsymbol{I}$ as $\boldsymbol{BB^T} = \boldsymbol{AA^T} = \boldsymbol{I}$ as $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{O_n}$ - further
$\mathbf{det}(\boldsymbol{AB}) = \mathbf{det}(\boldsymbol{A})\,\mathbf{det}(\boldsymbol{B}) = \mathbf{1}$ if $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{SO_n}$ □

10 Groups: Axioms and Examples

Group - A group is a set with an operation $*$ such that

$$* : G \times G \to G$$

$$(a, b) \mapsto a * b$$

such that

⓪ $a, b \in G \Rightarrow a * b \in G$	CLOSURE
① $a, b, c \in G \Rightarrow (a * b) * c = a * (b * c)$	ASSOCIATIVE
② $\exists e \in G$ such that $e * g = g = g * e \ \forall g \in G$	IDENTITY
③ $\forall g \in G, \exists g^{-1} \in G$ such that $g * g^{-1} = e = g^{-1} * g$	INVERSES

Example 63. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ under addition
Look at $\mathbb{Q}$

$a, b \in \mathbb{Q} \Rightarrow a + b \in \mathbb{Q}$	✓
addition is associative	✓
$0 \in \mathbb{Q}, a + 0 = a = 0 + a$	✓
$a \in \mathbb{Q}_+ \Rightarrow -a \in \mathbb{Q}\ a + (-a) = 0$	✓

under multiplication (all $\setminus\{0\}$)

Other groups: $\mathbb{R}^n$ under $+$, $\mathbb{Q}, \mathbb{R}, \mathbb{C}$

Example: Given any field F, say $\mathbb{R}, \mathbb{C}$, $G = GL_n(F) = \{A\ n\text{x}n$ over $F \mid \det A \neq 0\}$ is a group under $\times$. (GL_n is the General Linear group)
Check: ⓪ - $A, B \in G \Rightarrow AB \in G$ since AB is nxn, all entries $\in F$ (since F is closed under $\times, +$) and $\det AB = \det A \det B \neq 0$ since $\det A, \det B \neq 0$
① - Multiplication of matrices is associative, $(AB)C = A(BC)\ \forall A, B, C \in G$
② - Identity $= I$, since $AI = A = IA\ \forall A \in G$
③ - If $A \in G$ then A is invertible since $\det A \neq 0$ and $A^{-1} \in G$ as $detA^{-1} = (\det A)^{-1} \neq 0$

Example: $SL_n(F)$ $\{A \in GL_n(F) \mid \det A = 1\}$ is also a closed group under multiplication.

On $\mathbb{R}$, $SO_n(\mathbb{R})$ are groups by Lemma 62 p. 112.

Abelian Group - A group G is abelian (or commutative) if $a * b = b * a\ \forall a, b \in G$ e.g. $GL_n(F)$ is not abelian unless $n = 1$

Subgroup - If G is a group under $*$, the subset $H \subset G$ is a subgroup if H is a group under the restriction of $*$ to $H \times H$. We sometimes also denote this $H \leq G$.

Examples: $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$
$SO_n(\mathbb{R}) < SL_n(\mathbb{R}) < GL_n(\mathbb{R})$

Example 64. Subgroups of $\mathbb{Z}$ under $+$
Notation: For $n \in \mathbb{N} \cup \{0\}$ $n\mathbb{Z} = \{m \in \mathbb{Z} | n$ divides $m\}$ i.e. $m = ng$ for

some $g \in \mathbb{Z}$
$n\mathbb{Z} \leq \mathbb{Z}$ as $n|m_1, n|m_2 \Rightarrow n|(m_1 + m_2), n|0$ and $n|m \Rightarrow n| - m$
The associative law is inherited from $\mathbb{R}$
Any subgroup of $\mathbb{Z}$ is one of these.
Let $H \leq \mathbb{Z}$ with $H \neq \{0\}$
Let n be the smallest positive integer in H - note that H necessarily contains some positive integers. Then $H = n\mathbb{Z}$, because if $m \in H$, then $m = nq + r$ for some $r, q \in \mathbb{Z}$ with r either 0 or $0 < r < n$
Then $r \in H$ since both m and n are, and $r = -nq + m \in H$, but n was minimal, so $r = 0$ and $m \in n\mathbb{Z}$, and thus $H = n\mathbb{Z}$

Lemma 65. If G is a group, then $H \subset G$ is a subgroup if $H \neq \phi$ and $a, b \in H \Rightarrow a^{-1} * b \in H$

Proof. $H \neq \phi \Rightarrow$ let $a \in H$, then $a^{-1} * a = e \in H$
So $a^{-1} = a^{-1} * e \in H$ since $a, e \in H$
If $c, d \in H$ then $c * d = (c^{-1})^{-1} * d \in H \Rightarrow c^{-1} \in H$
So H contains e, is closed under $*$ and taking inverses, and so is a subgroup (as the associative law is inhertied from G). □

Lemma 66. Let G be a group under $*$
i) The identity is unique i.e. if e' satisfies $g * e' = g \; \forall g$, then $e' = e$
ii) Each g has a unique inverse
iii) $(g^{-1})^{-1} = g$
iv) $(g * h)^{-1} = h^{-1} * g^{-1}$

Proof. i) $e = e * e' = e'$
ii) Fix $g \in G$ and assume that x, y satisfy $x * g = e, g * y = e$ i.e. x, y both inverses for g
Then $y = e * y = (x * g) * y = x * (g * y) = x * e = x$
iii) $g * g^{-1} = e = g^{-1} * g$, so g is the inverse of g^{-1}. But $(g^{-1})^{-1}$ was defined to be this, so $g = (g^{-1})^{-1}$ by uniqueness
iv) $(g * h) * (h^{-1} * g^{-1}) = g * (h * h^{-1}) * g^{-1} = g * e * g^{-1} = g * g^{-1} = e$
Similarly $(h^{-1} * g^{-1}) * (g * h) = e$ so $h^{-1} * g^{-1}$ is an inverse of $g * h$ that is unique
$\therefore h^{-1} * g^{-1} = (g * h)^{1}$ □

There is another important example - the symmetric groups

Recall: Given a set X and functions $f, g : X \to X, f : x \mapsto f(x), g : x \mapsto g(x)$ with the composite $f \circ g : X \to X, f \circ g : x \mapsto f(g(x))$

Lemma 67. The composition of functions $X \to X$ is associative

Proof. Let $f, g, h : X \to X$, for $x \in X$
$(f \circ (g \circ h))(x) = f(g \circ h(x)) = f(g(h(x)))$ and equally $((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x)))$

This holds for any $x \in X$ so the functions $f \circ (g \circ h)$ and $(f \circ g) \circ h$ on X are equal. □

Note that $f : X \to X$ is a permutation of X if it is a bijection on X

Proposition 68. The set $\mathbf{Sym}\, X = \{f : X \to X | f \text{ a permutation of } X\}$ is a group under composition.

Also note that if $|X| = n$ $|\mathbf{Sym}\, X| = n!$

Proof. If $f, g \in \mathbf{Sym}\, X$ then $f \circ g \in \mathbf{Sym}\, X$: f, g are bijections from X to X, so the composite $f \circ g : X \to X$ is bijective
If $f \circ g(x_1) = f \circ g(x_2) \underbrace{\Rightarrow}_{f \text{ is inj.}} g(x_1) = g(x_2) \underbrace{\Rightarrow}_{g \text{ is inj.}} x_1 = x_2$
If $y \in X$ let $z \in X$ be such that $f(z) = y$
Let $x \in X$ with $g(x) = z$, so $f \circ g(x) = y \Rightarrow f \circ g \in \mathbf{Sym}\, X$
Composition is associative, as shown in Lemma 67.
The map $i : X \to X, x \mapsto x$ is a permutation, so is in $\mathbf{Sym}\, X$
And for $f \in \mathbf{Sym}\, X$, have $f \circ i = f$ $i \circ f$, so i is the identity in $\mathbf{Sym}\, X$
Finally, inverses exist:
If $f \in \mathbf{Sym}\, X$ then f is bijective, define the function $f^{-1} : X \to X$ the inverse function to f: for $x \in X$, define $f^{-1}(x)$ to the unique element of X mapped to x by f. Then f^{-1} is a function from X to X, and $f \circ f^{-1} = i = f^{-1} \circ f$ □

Notation: If X is finite with n points (usually take $x = \{1, 2, \ldots, n\}$) we write $\mathbf{Sym}_n = S_n = \mathbf{Sym}_X$

Order (group) - The order $|G|$ of the group G is the number of its elements.
e.g. $|S_n| = n!$

Notation 69. For permutations, if $X = \{1, 2, \ldots, n\}$ write $g \in \mathbf{Sym}\, X$ as
$g = \begin{pmatrix} 1 & 2 & \ldots & n \\ g(1) & g(2) & \ldots g(n) & \end{pmatrix}$

Examples:
$i = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}$
Given $f = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, g = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$
$f \circ g = \begin{pmatrix} 1 & 2 & 3 \\ f(g(1)) & f(g(2)) & f(g(3)) \\ f(2) & f(1) & f(3) \\ 3 & 1 & 2 \end{pmatrix}$
$g \circ f = \begin{pmatrix} 1 & 2 & 3 \\ g(f(1)) & g(f(2)) & g(f(3)) \\ g(1) & g(3) & g(2) \\ 2 & 3 & 1 \end{pmatrix}$

As $f \circ g \neq g \circ f$ it is not of abelian form

The inverse of f is produced by inverting the rows of f
e.g. $(gf)^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$
Exercise: $|S_3| = 6$, list all elements.
Find all subgroups (there is one of order 1, 2, 3 and 3 of order 2)

Example 70. Group of symmetries of a regular n-gon under composition
Let the vertices of the n-gon be, for example, the $n\vec{}$roots of 1. Composition of symmetries is a symmetry, composition is associative (Lemma 67), we have an identity and the inverse of a symmetry is a symmetry. The group we obtain is the DIHEDRAL group D_{2n} of order $2n$
Note: $|D_{2n}| = 2n$, but $D_{2n} \leq S_n$

We have n rotations: $i, a, \underbrace{a^2}_{a*a}, \ldots, a^{n-1}$ where a is the rotation given by $\frac{2\pi}{n}$ about the origin.
We also have n reflections: $b, ab, a^2b, \ldots, a^{n-1}b$
It we take the example where $n = 3$ and label the points we can express this as a permutation group

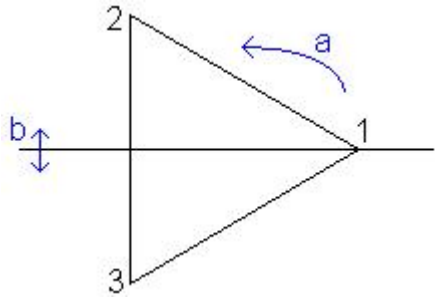

Figure 39: Permutations of the cubic roots of unity

$a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ and $ba = a^{-1}b$
(proof: $b : z \mapsto \overline{z}, a : z \mapsto zw$ with $w = e^{\frac{2pii}{n}}$
$ba(w^j) = bw^{j+1} = w^{-j-1} = a^{-1}b(w^j)$

Note also that when $n = 4$ D_8 has subgroups of orders 1, 2, 4, 8
$n = 5$ D_10 has subgroups of orders 1, 2, 5, 10

Notation: If G is a group, and $S \subset G$, write $< S >$ for the smallest subgroup of G containing S - this is the subgroup generated by S.

In fact $< S >$ consists of all elements of G which can we written in elements of S and their powers (in any order, possibly with repetitions)
The dihedral groups are generated by the elements a, b. In fact, as an abstract group, $D_{2n} =< a, b | a^n = e = b^2, bab = a^{-1} >$

Example 71. Cyclic Groups A group G is cyclic if $\exists g \in G$ such that each element of G is a power of g (or g^{-1})
e.g. $G = \mathbb{Z}$ under +, $G =< 1 >$ or $< -1 >$
e.g. The group $< a >$ in D_{2n} above is cyclic of order n, the cyclic group C_n

Abstractly C_n is the group $< g | g^n = e >$

Order (element) - Given a group G, and element $g \in G$, the order of g in G is the smallest positive integer n so that $g^n = e$ if such an n exists. If no such n exists we say the order is infinite. Write $o(g)$ for the order of g

Remark 72. If $g \in G$ then $o(g) = | < g > |$ where $< g >=< \{g\} >$
Note: $< g >= \{e, g, g^2, \ldots, g^{n-1}\}$ when $n = o(g)$ if n is infinite.

Note that if G is finite and $g \in G$ then $o(g) \leq |G|$. In fact it is true that $o(g) \Big| |G|$ (proof later)

Example: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix}$ has order 6, as it is the lowest common multiple of $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 5 & 4 \end{pmatrix}$ (order 2) and $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix}$ (order 3)

We need to be able to compare groups

Homomorphism - Given two groups $(G_1, *_1), (G_2, *_2)$ a mapping $\theta : G_1 \to G_2$ is a homomorphism if $\theta(a *_1 b) = \theta(a) *_2 \theta(b)\ \forall a, b \in G_1$

Isomorphism - It is an isomorphism if it is a bijective homomorphism.

Examples:
$\theta : (\mathbb{R}, +) \to (\mathbb{R}_{\geq 0}, \times) \quad x \mapsto e^x$ - bijective
$\theta(x + y) = e^{x+y} = e^x . e^y = \theta(x).\theta(y)$

$G_1 = \{\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} | a \in \mathbb{R}\}$ under multiplication, $G_2 = \mathbb{R}$ under addition
$\theta : \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \mapsto a$ is an isomorphism.

Let G_1 be the group of all non-singular linear maps from V to V (where V is the n-dimensional vector space F) under composition. Let G_2 be the group of non-singular nxn matrices over F. Then G_1, G_2 are isomorphic. Fix a basis B. Then the map $\alpha \mapsto [\alpha]_B$ is an isomorphism, and the map $\theta : G_1 \to G_2 = GL_n(F), \alpha \mapsto [\alpha]_B$ is a bijective homomorphism, $\theta(\alpha \circ \beta) = [\alpha \circ \beta]_B = [\alpha]_B [\beta]_B$

10.1 A Theorem of Lagrange

Let $\boldsymbol{H \subset G}$. We shall prove that, if $\boldsymbol{G}$ is finite, $|\boldsymbol{H}| \Big| |\boldsymbol{G}|$. As a consequence, we shall prove that the order of any element of a finite group divides $|\boldsymbol{G}|$

Left coset Let $\boldsymbol{G}$ be a group, let $\boldsymbol{H \subset G}$. For $\boldsymbol{g \in G}$ define $\boldsymbol{gH = \{gh | h \in H\}}$. Such sets are the left cosets of $\boldsymbol{H}$ in $\boldsymbol{G}$

We shall prove:
i) $\boldsymbol{G}$ is the disjoint union of distinct left cosets
ii) The (left) cosets all have the same size $|\boldsymbol{H}|$

To prove i) we pose:
ia) Each element of $\boldsymbol{G}$ is in some left coset
ib) Distinct left cosets are disjoint

e.g. $\boldsymbol{G = (\mathbb{Z}, +)} \quad \boldsymbol{n \in \mathbb{N}}$
$\boldsymbol{H = \{m : n | m\} = n\mathbb{Z}}$

The cosets of $\boldsymbol{H}$ are:

$$\left.\begin{array}{ll} \boldsymbol{0 + H} & \text{multiples of } \boldsymbol{n} \\ \boldsymbol{1 + H} & \text{integers } \boldsymbol{1 \mod n} \\ \boldsymbol{2 + H} & \text{integers } \boldsymbol{2 \mod n} \\ \vdots & \vdots \\ \boldsymbol{(n-1) + H} & \text{integers } \boldsymbol{(n-1) \mod n} \end{array}\right\} \boldsymbol{n} \text{ cosets, disjoint}$$

e.g. $\boldsymbol{G = S_3}$ $\boldsymbol{H = \left\langle \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\rangle}$ - order 2

$$\begin{aligned} \boldsymbol{H} &= \left\{ \boldsymbol{e}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\} \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \boldsymbol{H} &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \boldsymbol{H} \\ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \boldsymbol{H} &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \boldsymbol{H} \end{aligned}$$

Theorem 73. Lagrange's Theorem If $\boldsymbol{G}$ is a finite group, and $|\boldsymbol{H}|$ is a subgroup of $\boldsymbol{G}$, then $|\boldsymbol{H}|$ divides $|\boldsymbol{G}|$
In fact $|\boldsymbol{G}| = |\boldsymbol{H}|.|\boldsymbol{G : H}|$ where $|\boldsymbol{G : H}|$ is the index of $\boldsymbol{H}$ in $\boldsymbol{G}$, the number of distinct left cosets of $\boldsymbol{H}$ in $\boldsymbol{G}$

Proof. First, we prove that the left cosets of $\boldsymbol{H}$ in $\boldsymbol{G}$ partition $\boldsymbol{G}$, firstly, any $\boldsymbol{g \in G}$ is in some left coset, namely $\boldsymbol{g \in gH}$, secondly if $\boldsymbol{aH \cap bH \neq \phi}$ then $\boldsymbol{aH = bH}$
i) Claim that if $\boldsymbol{c \in aH}$ then $\boldsymbol{cH = aH}$

> $\boldsymbol{c \in aH \Rightarrow c = ak}$ for some $\boldsymbol{k \in H}$ so $\boldsymbol{ch = (ak)h = a(kh) \in aH}$ for $\boldsymbol{h \in H}$ so $\boldsymbol{cH \subset aH}$, but $\boldsymbol{c \in aH \Rightarrow a \in cH}$ (as $\boldsymbol{c =}$

$ak \Rightarrow a = ck^{-1} \in cH$), so it follows by the previous argument that $aH \subset cH \Rightarrow aH = cH$ as claimed.

ii) Assume now $aH \cap bH \neq \phi$, say $c \in aH \cap bH$, then $cH = aH$ and $cH = bH \therefore aH = bH$
Finally, each left coset has size $|H|$:
Fix $g \in G$, the map $\phi : H \to gH, h \mapsto gh$ is a bijection from H onto gH (with inverse $\phi^{-1} : gH \to H, x \mapsto g^{-1}x$
We have proved: The $|G : H|$ are distinct left cosets and cover G disjointly and all have the same size $|H|$

$$\therefore |H| = |G : H|\ |H|$$

□

Recall: The order $o(g)$ of $g \in G$ is the least $n \in \mathbb{N}$ such that $g^n = e$ if it exists. Otherwise $o(g) = \infty$

Theorem 74. If g is an element of a finite group G, then the order of g divides the order of group G. Hence $g^{|G|} = e$ for any $g \in G$

Proof. Consider the subgroup $H =< g >$ generated by $g \in G$. Then $H = \{e, g, g^2, \ldots, g^{n-1}\}$ where $n = o(g)$. So $| < g > | = o(g)$ and so $o(g)|G$ by Lagrange's Theorem (Theorem 73). Finally if $|G| = o(g).q$ with $q \in \mathbb{N}$, then $g^{|G|} = (g^{o(g)})^q = e^q = e$ □

Note: $g \in G, 0 \leq k < o(g)$ then $\{e, g, g^2, \ldots, g^k\}$ are all distinct. Because if not, then $g^i = g^j$ for some $0 \leq i \leq j \leq k$, then $e = g^{j-i}$ $0 \leq j - i \leq o(g)$ and so $i = j$ by the minimality of $o(g)$.
Also note that if $o(g)$ is infinite (then G infinite) and $< g >= \{e, g^{\pm 1}, g^{\pm 2}, \ldots\}$

Proposition 75. Let G be a group of prime order. Then G is cyclic and moreover, $G =< g >\ \forall g \in G \setminus \{e\}$

Proof. Let $g \in G \setminus \{e\}$. Then $\{e\} \neq < g > \leq G$ so $| < g > | \Big| |G|$ by Lagrange. But $|G|$ is prime, so $| < g > | = |G|$ and thus $< G >= G$ □

Recall: an equivalence relation R on a set$S$ is a relation which is:
a) Reflexive $aRa\ \forall a \in S$
b) Symmetric $aRb \Rightarrow bRa\ \forall a, b \in S$
c) Transitive $aRb, bRc \Rightarrow aRc\ \forall a, b, c \in S$
Examples:
i) $S = \mathbb{Z}, R = \equiv$ defined as $a \equiv b \mod n$ if $n|(a - b)$
ii) On the set of all groups the relation $\simeq$ (homomorphism) is an equivalence relation

Equivalence Classes - Take R an equivalence relation on S, and $a \in S$. $[a] = \{b \in S : aRb\}$ is the equivalence class of a. Claim the equivalence classes partition S

Proof. Each point is in some class $a \in [a]$ as aRa
If two classes have a non-empty intersection then they are equal ($[a] \cap [b] \neq \phi \Rightarrow [a] = [a]$) (as $\forall x \in [a]$ if $\exists y \in ([a] \cap [b])$ then $xRa, aRy, yRb \Rightarrow xRb$ so $x \in [b]$ and vica verse)

□

We can think of a relation R as a graph, with the vertices corresponding to the elements of S, and edges from a to b iff aRb

The equivalence classes in the graph are cliques - complete subgraphs in which all edges are present.
Equivalence relations give cosets as equivalence classes.

Lemma 76. Let $H \leq G$
i) $aH = bH$ iff $a^{-1}b \in H$
ii) The relation R on G given by aRb if $a^{-1}b \in H$ is an equivalence relation

Proof. i) $aH = bH$ iff $H = a^{-1}bH$ (multiplying by a^{-1})
so need to show $gH = H \Leftrightarrow g \in H$
$g \in H$ implies that for $h \in H, gh \in H$ so $gH \subset H$
and $h = g(g^{-1}h) \in gH$ so $H \subset gH \Rightarrow$ if $H = gH$ then $g = g.e \in gH = H \therefore g \in H$
ii) $a^{-1}a = e \in H$, so R is reflexive
if $a^{-1}b \in H$ then $b^{-1}a = (a^{-1}b)^{-1} \in H$ so R is symmetric
if $a^{-1}b \in H, b^{-1}cinH$ then $(a^{-1}b)(b^{-1}c) = (a^{-1}c) \in H$ so R is transitive
□

Some applications of Lagrange
Consider $\mathbb{Z}$ under $+$, let $H = \{m \in \mathbb{Z} | n \text{ divides } m\}$ for a fixed $n \in \mathbb{N}$. The cosets are $\{\underbrace{[0]}_{0+H}, \underbrace{[1]}_{1+H}, \ldots, \underbrace{[n-1]}_{(n-1)+H}$
Can define $[a] + [b] = [a + b]$ and $[a][b] = [ab]$
This is well-defined. If $[a] = [a']$ and $[b][b']$ then $[ab] = [a'b']$
If $a' = a + nx, b' = b + ny$ then $a'b' = ab + n(bx + ay + nxy)$

To get a group, take "units" only - elements with inverses.
$\bigcup_n = \{[a] | a \text{ prime to } n\}$
$\phi(n) = |\bigcup_n|$ - the Euler function
e.g. $\phi(p) = p - 1$ if p is prime
$\phi(4) = 2$

Lemma 77. $\bigcup_n$ is a group under multiplication **mod** n (as above)

Proof. The operation is well-defined.
Closure: a, b coprime to $n \Rightarrow ab$ coprime to n
Identity: $[1]$
Inverses: Let $[a] \in \bigcup_n$ then map $[x] \mapsto [ax]$ is one to one

(if $[ax_1] = [ax_2]$ then n divides $ax_1 - ax_2 = a(x_1 - x_2)$ and so n divides $(x_1 - x_2)$ since a coprime to n, so $[x_1] = [x_2]$)

Since $\bigcup_n$ is finite, any injection $\bigcup_n \to \bigcup_n$ is onto, so for some x, have $[a][x] = [ax] = 1$
The multiplication is commutative □

Theorem 78. Fermat-Euler Theorem Let $n \in \mathbb{N}$. If $a \in \mathbb{N}$ is coprime to n then $a^{\phi(n)} \equiv 1 \mod n$. If p is prime, then $a^{p-1} \equiv 1 \mod p$ for any a not divisible by p

Proof. By Theorem 74 (p. 119) $[a]^{\phi(n)} = [1]$, so $a^{\phi(n)} \equiv 1 \mod n$
And $\phi(p) = p - 1$ for p a rpime □

10.2 Orders of Subgroups

Example: Subgroups of $\boldsymbol{d_10}$, the dihedral group of order 10:
By Lagrange, any such subgroup will have order 1, 2, 5, 10.
Any subgroup of order 5 consists of $\boldsymbol{e}$ and four elements of order $\mathbf{5}$, so there is only one such, namely $< \boldsymbol{a} >$, where $\boldsymbol{a}$ is a rotation - this is the group of all rotations.
Any subgroup of order 2 consists of $\boldsymbol{e}$ and an element of order 2. There are five such subgroups, all reflections, so five such subgroups exist.

Exanple: Subgroups of $\boldsymbol{D_8}$, 5 subgroups of order 2, 3 subgroups of order 4 (one of which is cyclic)

Note: The converse of Lagrange's Theorem is false e.g. $|\boldsymbol{A_4}| = \mathbf{12}$ but $\boldsymbol{A_4}$ has no subgroup of order 6

10.3 Right cosets

Given $\boldsymbol{H \leq G, g \in G} \quad \boldsymbol{Hg = \{hg | h \in H\}}$, the right coset of $\boldsymbol{H}$ corresponding to $\boldsymbol{g \in G}$
This has similar properties to left cosets e.g. $\boldsymbol{Ha = Hb}$ iff $\boldsymbol{ab^{-1} \in H}$

e.g. $\boldsymbol{G = S_3} \quad \boldsymbol{|G| = 6}$
i)$\boldsymbol{K} = \left\langle \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{3} & \mathbf{1} & \mathbf{2} \end{pmatrix} \right\rangle$ has order 3
$\boldsymbol{K, G \setminus K}$ are the two left cosets and the two right cosets (as we only have 2 by cosets being distinct and order dividing the order of the set, so these 2 must represent both)
so $\boldsymbol{gK = Kg \; \forall g \in G}$ ii) $\boldsymbol{H} = \left\langle \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{2} & \mathbf{1} & \mathbf{3} \end{pmatrix} \right\rangle \; \boldsymbol{|H| = 2}$
$\begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{2} & \mathbf{3} & \mathbf{1} \end{pmatrix} \boldsymbol{H} = \left\{ \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{2} & \mathbf{3} & \mathbf{1} \end{pmatrix}, \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{3} & \mathbf{2} & \mathbf{1} \end{pmatrix} \right\}$
by multiplying each element of $\boldsymbol{H}$ in turn
$\boldsymbol{H} \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{2} & \mathbf{3} & \mathbf{1} \end{pmatrix} = \left\{ \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{2} & \mathbf{3} & \mathbf{1} \end{pmatrix}, \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{1} & \mathbf{3} & \mathbf{2} \end{pmatrix} \right\}$

We say $\boldsymbol{K}$ is a normal subgroup and $\boldsymbol{H}$ is not normal

11 Normal Subgroups and Homomorphisms

Normal subgroup - A subgroup K of G is a normal subgroup if $\forall g \in G, \forall k \in K, gkg^{-1} \in K$
Notation: $K \lhd G$

Note: This is equivalent to $Kg = gK \ \forall g \in G$ and to $gKg^{-1} = K \ \forall g \in G$
Note: Can switch g and g^{-1} (as $g^{-1} \in G$ so could just call it g instead)

Lemma 79. i) If $K < G$ of index 2, then $K \lhd G$
ii) If G is abelian then any subgroup is normal

Proof. i) If $g \in K$ then $gK = K = Kg$
If $g \in G \setminus K$ then $gK = G \setminus K$ (the left coset, not K) $= Kg$ (the right coset, also not K)
ii) If G is abelian, then $gkg^{-1} = k \ \forall k \in K, g \in G$ so certainly $gkg^{-1} \in K \ \forall k \in K, g \in G$ □

Theorem 80. Let K be a normal subgroup of G. The set of (left) cosets of K in G is a group under multiplication

$$aK.bK = abK$$

called the quotient group or factor group - $\frac{G}{K}$

For example, take $G = \mathbb{Z}$ under $+$, $K =< n >= \{m \in \mathbb{Z} : n \text{ divides } m\}$
G is abelian, and $K \lhd G$

$\frac{G}{K} = \{K, 1+K, 2+k, \ldots, (n-1)+K\}$ is a group, as $(a+K)+(b+K) = (a+b)+K$. This is the integers under addition, and can be denoted $\frac{\mathbb{Z}}{<n>}$, $\frac{\mathbb{Z}}{n\mathbb{Z}}$ or $\mathbb{Z}_n$

Proof. The multiplication is well-defined:
If $akK = a'K$ and $bK = b'K$ then $abK = a'b'K$

$$\begin{aligned} a'b' &= ak_1bk_2 \text{ for some } k_1, k_2 \in K \\ &= abk_1'k_2 \text{ for some } k_1' in K \text{ as } K \lhd G \\ &= abk_3 \text{ where } k_3 = k_1'k_2 \in K \end{aligned}$$

so $a'b'K = abK$
Closed under multiplication - clear as abK is a coset
Associative: $(aKbK)cK = abKcK = (ab)cK \underbrace{=}_{\times \text{ is associative in } G} a(bc)K = aKbcK = aK(bKcK)$
Identity: K is the identity $K = eK$ and $eK.gK = gK = gKeK \ \forall gK$
Inverses: Fix gK such that $(gK)^{-1} = g^{-1}K$ for $gKg^{-1}K = gg^{-1}K = g^{-1}KGK$

□

This does not work unless $\boldsymbol{K}$ is normal, the proposed is not well-defined unless $\boldsymbol{K}$ is normal.

For example, take $\boldsymbol{V}$ a vector space (which is by definition an abelian group), $\boldsymbol{U} \subset \boldsymbol{V}$ also a subspace (and by defintion normal). We can form the quotient space $\frac{\boldsymbol{V}}{\boldsymbol{U}}$. This is not a subspace of $\boldsymbol{V}$. On the other hand, there are subspaces $\boldsymbol{W}$ of $\boldsymbol{V}$ such that $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{W}$ and $\boldsymbol{U} \cap \boldsymbol{W} = \boldsymbol{\phi}$, the complement of $\boldsymbol{U}$ in $\boldsymbol{V}$, and each $\boldsymbol{W}$ is isomorphic to $\frac{\boldsymbol{V}}{\boldsymbol{U}}$ but $\frac{\boldsymbol{V}}{\boldsymbol{U}}$ is not one of them.

11.1 Homomorphisms

Recall: Given G_1, G_2 groups, $\theta : G_1 \to G_2$ is a homomorphism if $\theta(ab) = \theta(a)\theta(b)\ \forall a, b \in G$

Kernel - The kernel of a map is the set of all elements which map to the identity
$\ker(\theta) = \{g \in G_1 : \theta(g) = e\}$

Image - The image of a map is the set of all elements in G_2 which are mapped to by an element in G_1
$\operatorname{Im}\theta = \theta(G_1) = \{\theta(g) : g \in G_1\}$

Remark
i) θ is injective iff $\ker(\theta) = \{e\}$
ii) θ is surjective iff $\operatorname{Im}(\theta) = G_2$
$\ker\theta \lhd G_1$
Note: $\frac{G}{K}$ is a new group, there may or may not even be a subgroup of G isomorphic to $\frac{G}{K} : G = \mathbb{Z}$ under $+$, any proper quotient is a finite group $\frac{\mathbb{Z}}{<n>}$ (we have seen the subgroups of $\mathbb{Z}$ are precisely $< n >$ for various n, all such infinite)

Note 81. i) $\theta(e) = e$
ii) $\theta(g^{-1}) = [\theta(g)]^{-1}$
iii) $o(\theta(g))|o(g)$

Proof. i) $\theta(e) = \theta(ee) = \theta(e)\theta(e)$ so cancel to get $\theta(e) = e$
ii) $\theta(g)\theta(g^{-1}) = \theta(gg^{-1}) = \theta(e) = e[\theta(g)]^{-1}\theta(g)$ so $\theta(g^{-1}) = [\theta(g)]^{-1}$
iii) Covered in example sheets □

Lemma 82. If $\theta : G_1 \to G_2$ is a homomorphism, then
i) $\ker(\theta) \lhd G$
ii) $\operatorname{Im}(\theta) \leq G$

Proof. i) Firstly, $\ker\theta \leq G_1$
$e \in \ker\theta$, let $a, b \in \ker\theta$ and $a^{-1}b \in \ker\theta$
$\theta(a^{-1}b) = [\theta(a)]^{-1}\theta(b) = e^{-1}e = e$
Secondly it is normal: let $g \in G, k \in \ker(\theta)$
Then $gkg^{-1} \in \ker\theta : \theta(gkg^{-1}) = \theta(g)\theta(k)[\theta(g)]^{-1} = e$
ii) $e \in \operatorname{Im}\theta$ as $e = \theta(e)$
if $x, y \in \operatorname{Im}(\theta)$ then $x^{-1}y \in \operatorname{Im}(\theta)$ for if
$x = \theta(a), y = \theta(b)$, then $x^{-1}y = \theta(a^{-1}b)$
(in fact, if $H \leq G_1$ then $\theta(H) \leq G_2$) □

Remark 83. i) θ is injective iff $\ker\theta = \{e\}$
ii) θ is surjective iff $\operatorname{Im}\theta = G_2$

Proof. i) If $\boldsymbol{\theta}$ is injective, and $\boldsymbol{a} \in \ker \boldsymbol{\theta}$ then $\boldsymbol{\theta(a) = e = \theta(e)}$, so $\boldsymbol{a = e}$ and so $\boldsymbol{\ker(\theta) = \{e\}}$.
Conversely, assume $\boldsymbol{\ker(\theta) = \{e\}}$, if $\boldsymbol{\theta(a) = \theta(b)}$ then $\boldsymbol{\theta(a^{-1}b) = e}$, so $\boldsymbol{a^{-1}b \in \ker(\theta) = \{e\}}$ so $\boldsymbol{a = b}$ ii) Is clear from definitions □

e..g $\boldsymbol{G_1 = \mathbb{Z}}$ under $+$, $\boldsymbol{G_2}$ the group of rotations of a regular $\boldsymbol{n}$-gon, then $\boldsymbol{G_2 =< a >}$, a rotation by $\boldsymbol{\frac{2\pi}{n}}$

The map $\boldsymbol{\theta : m \mapsto a^m}$ is a homomorphism ($\boldsymbol{\theta(m_1 + m_2) = a^{m_1}a^{m_2}}$) and surjective as $\boldsymbol{\ker(\theta) =< n >}$
$\boldsymbol{G_1 = GL_n(\mathbb{R}), G_2 = \mathbb{R} \setminus \{0\}, \times}$ under the multiplication of matrices
Then $\boldsymbol{\theta : G_1 \to G_2, A \mapsto \det A}$ is a homomorphism: $\boldsymbol{\det(AB) = \det(A)\det(B)}$, so $\boldsymbol{\theta(AB) = \theta(A)\theta(B)}$
The kernel is $\boldsymbol{SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) : \det A = 1\}}$

Let $\boldsymbol{G}$ be a group, $\boldsymbol{K \lhd G}$
A natural homomorphism from $\boldsymbol{G}$ to $\boldsymbol{\frac{G}{K}}$ is $\boldsymbol{q : G \to \frac{G}{K}, g \mapsto gK}$
This is a homomorphism, as $\boldsymbol{q(ab) = abK = aKbK = q(a)q(b)}$
Clearly $\boldsymbol{q}$ is surjective, and its kernel is $\boldsymbol{K}$, for $\boldsymbol{g \in \ker(q)}$ iff, $\boldsymbol{q(g) = K}$ iff $\boldsymbol{gK = K}$ iff $\boldsymbol{g \in K}$, so

Lemma 84. Any normal subgroup $\boldsymbol{K}$ of $\boldsymbol{G}$ is the kernel of some homomorphism $\boldsymbol{G \to G_2}$ (in fact $\boldsymbol{G_2 = \frac{G}{K}}$)

Theorem 85. The Isomorphism Theorem Let $\boldsymbol{\theta : G \to G_2}$ be a homomorphism, let $\boldsymbol{K = \ker(\theta)}$. Then $\boldsymbol{K \lhd G}$ and $\boldsymbol{\frac{G}{K} \simeq \operatorname{Im}\theta}$

Proof. $\boldsymbol{K \lhd G}$ is done above
Define $\boldsymbol{\overline{\theta} : \frac{G}{K} \to G_2, gK \mapsto \theta(g)}$
So that $\boldsymbol{\theta = \overline{\theta}q}$
Then $\boldsymbol{\overline{\theta}}$ is well-defined, as $\boldsymbol{G}$ contains subsets, each of which associates with a unique point in $\boldsymbol{G_2}$ and $\boldsymbol{\frac{G}{K}}$, so $\boldsymbol{\overline{\theta}}$ takes 1 point to 1 point.
If $\boldsymbol{aK = bK \Rightarrow a^{-1}b \in K (= \ker(\theta))}$ so $\boldsymbol{\theta(a^{-1}b) = e}$
So $\boldsymbol{\theta(a) = \theta(b)}$, so $\boldsymbol{\overline{\theta}(aK) = \overline{\theta}(bK)}$
Injective: reverse the above argument
If $\boldsymbol{\overline{\theta}(aK) = \overline{\theta}(bK)}$ then $\boldsymbol{\theta(a) = \theta(b)}$, so $\boldsymbol{a^{-1}b \in \ker(\theta) = K}$, so $\boldsymbol{aK = bK}$
$\boldsymbol{\therefore \overline{\theta}}$ is onto $\boldsymbol{\operatorname{Im}(\theta)}$ - clear.
Finally, $\boldsymbol{\overline{\theta}}$ is a homomorphism.
$\boldsymbol{\overline{\theta}(aKbK) = \overline{\theta}abK = \theta(ab) = \theta(a)\theta(b) = \overline{\theta}(aK)\overline{\theta}(bK)}$ □

Remark
If $\boldsymbol{G}$ is finite, we get $\boldsymbol{\frac{|G|}{|\ker\theta|} = |\operatorname{Im}\theta|}$
compare with the rank-nullity theorem for linear maps

$$\boldsymbol{\alpha : V \to W \quad \dim V = k(\alpha) + n(\alpha)}$$

e.g. $\boldsymbol{= \mathbb{Z}_p^*}$, $\boldsymbol{\frac{\mathbb{Z}}{<p>} \setminus \{0\}}$ under multiplication
$\boldsymbol{\theta : G \to G, x \mapsto x^2}$ is a homomorphism as $\boldsymbol{G}$ is abelian

$\boldsymbol{(xy)^2 = (xy)(xy) = x^2y^2}$ as it is commutative
$\boldsymbol{\ker\theta = \{\pm 1\}}$
$\boldsymbol{|\operatorname{Im}\theta| = \frac{p-1}{2}}$ $\boldsymbol{\operatorname{Im}(\theta)\{}$ quadratic residues$\boldsymbol{\}}$

Lemma 86. Any cyclic group is isomorphic to one of $\mathbb{Z}$ under + or $\frac{\mathbb{Z}}{<n>} = \mathbb{Z}_n$ under + (for various $n \in \mathbb{N}$)
So a unique cyclic group of each order, up to isomorphism

Proof. Let H be a cyclic group, say $H =< g >$
Let $\theta : \mathbb{Z} \to H, m \mapsto g^m$. Then θ is a homomorphism. ($g^{m_1+m_2} = g^{m_1}g^{m_2}$ so $\theta(m_1 + m_2) = \theta(m_1)\theta(m_2)$)
Also θ is clearly surjective (as H is generated by g).
Now $\ker\theta = \{k \in \mathbb{Z} : g^k = e\}$
If H is infinite then $\ker\theta = \{0\}$ so $H \simeq \mathbb{Z}$ (as θ is an isomorphism)
If H is finite, or order n, then $o(g) = n$ and $g^k = e$ iff $n|k$. So $H \simeq \frac{\mathbb{Z}}{<n>}$ by the isomorphism theorem □

Example: Subgroups of cyclic groups are cyclic
Quotients (i.e homomorphic images) of cyclic groups are cyclic

Simple Groups - A group G is simple if it has only trivial subgroups (the identity, or G). These groups are the building blocks of all finite groups. Finite simple groups have only just been classified

12 Actions of Groups

Recall: Permutations of a set $\boldsymbol{X}$, composition of permutations.

Acting on a group - Given a group $\boldsymbol{G}$ and a set $\boldsymbol{X}$, we say that $\boldsymbol{G}$ acts on $\boldsymbol{X}$ if the elements of $\boldsymbol{G}$ are permutations of $\boldsymbol{X}$ and the group multiplications yields the composition of permutations

Example: Take a regular $\boldsymbol{n}$-gon in a plane as our set $\boldsymbol{X}$
$\boldsymbol{D_{2n}} = \boldsymbol{G}$ =the group of all symmetries on this $\boldsymbol{n}$-gon (isometries on the Plane which don't damage the $\boldsymbol{n}$-gon)
Then $\boldsymbol{G}$ acts on the set of $\boldsymbol{n}$ vertices.
This makes $\boldsymbol{D_{2n}}$ a subgroup of $\boldsymbol{S_n}$
The action is "faithful"

Lemma 87. If $\boldsymbol{G}$ acts on $\boldsymbol{X}$, and if $\boldsymbol{x} \in \boldsymbol{X}$, then $\boldsymbol{G_x} = \{\boldsymbol{g} \in \boldsymbol{G} : \boldsymbol{g(x)} = \boldsymbol{x}\}$ is a subgroup, called the stabliser of $\boldsymbol{x}$ in $\boldsymbol{G}$

Proof. $\boldsymbol{e} \in \boldsymbol{G_x}$
If $\boldsymbol{g, h} \in \boldsymbol{G}$ then $\boldsymbol{g^{-1}h(x)} = \boldsymbol{g^{-1}(x)} = \boldsymbol{x}$ so $\boldsymbol{g^{-1}h(x)} \in \boldsymbol{G_x}$ $\therefore \boldsymbol{G_x}$ is a subgroup
If $\boldsymbol{G}$ acts on $\boldsymbol{X}, \boldsymbol{x} \in \boldsymbol{X}$, then write
$\boldsymbol{G(x)} = \{\boldsymbol{g(x)} : \boldsymbol{g} \in \boldsymbol{G}\}$ - called the $\boldsymbol{G}$-orbit of $\boldsymbol{x}$ in $\boldsymbol{X}$
$\boldsymbol{G}$ is transitive on $\boldsymbol{X}$ if $\boldsymbol{X}$ itself is an orbit $\forall \boldsymbol{x_1, x_2} \in \boldsymbol{X}$ $\exists \boldsymbol{g} \in \boldsymbol{G}$ such that $\boldsymbol{g(x_1)} = \boldsymbol{x_2}$ □

Lemma 88. The distinct orbits of $\boldsymbol{G}$ on $\boldsymbol{X}$ form a partition of $\boldsymbol{X}$

Proof. ① Define $\sim$ on $\boldsymbol{X}$ by $\boldsymbol{x_1} \sim \boldsymbol{x_2}$ if $\boldsymbol{g(x_1)} = \boldsymbol{x_2}$ and show $\sim$ is an equivalence relation, and thus the equivalence classes the set.
② Each point $\boldsymbol{x} \in \boldsymbol{X}$ lies in some orbit
If $\boldsymbol{a} \in (\boldsymbol{G_{x_1}} \cap \boldsymbol{G_{x_2}})$ then $\boldsymbol{G_{x_1}} = \boldsymbol{G_{x_2}}$

$\boldsymbol{a} = \boldsymbol{g_1(x_1)} = \boldsymbol{g_2(x_2)}$ for some $\boldsymbol{g_i} \in \boldsymbol{G}$ so for $\boldsymbol{g} \in \boldsymbol{G_1}$
$\boldsymbol{g(x_1)} = \boldsymbol{g(g_1^{-1}(a))} = \boldsymbol{gg_1^{-1}g_2(x_2)} \in \boldsymbol{G(x_2)}$

So $\boldsymbol{G(x_1)} \subset \boldsymbol{G(x_2)}$ so equal by symmetry □

Theorem 89. The Orbit-Stabiliser Theorem If $\boldsymbol{G}$ is a finite group acting on the set $\boldsymbol{X}$ then $|\boldsymbol{G}| = |\boldsymbol{G(x)}|.|\boldsymbol{G_x}|$ $\forall \boldsymbol{x} \in \boldsymbol{X}$

Proof. Claim: if $\boldsymbol{y} \in \boldsymbol{G(X)}$, say $\boldsymbol{y} = \boldsymbol{g(x)}$, then $\{\boldsymbol{f} \in \boldsymbol{G} : \boldsymbol{y} = \boldsymbol{f(x)}\} = \boldsymbol{gG_x}$ - the left coset of the stabiliser
For, $\boldsymbol{f}$ is in the LHS iff $\boldsymbol{f(x)} = \boldsymbol{g(x)}$ iff $\boldsymbol{g^{-1}f} \in \boldsymbol{G_x}$ iff $\boldsymbol{f} \in \boldsymbol{gG_x}$
Now, each stabiliser of $\boldsymbol{G_x}$ has size $|\boldsymbol{G_x}|$, and there are $|\boldsymbol{G(x)}|$ points in the orbits $\boldsymbol{G(x)}$, so $|\boldsymbol{G}| = |\boldsymbol{G(x)}|\ |\boldsymbol{G_x}|$ □

Example: Let $\boldsymbol{G}$ be the group of all symmetries of a regular $\boldsymbol{n}$-gon in the plane. Let $\boldsymbol{X}$ be the set of vertices of this $\boldsymbol{n}$-gon.

Then G acts on X, transitively (given x_1, x_2 $\exists$ a rotation taking one to the other)
So $|G| = n.|G_x|$ where $x \in X$
But $G_x = \{e, t\}$ where t is the reflection in the axis through x
So $|G_x| = 2$ and thus $|G| = 2n$

Example 90. Take G, the group of symmetries of a regular tetrahedron.
G acts on the set X of vertices of our tetrahedron, it is transitive [rotate suitably]
$x \in X, |G| = 4|G_x|$ (G_x is transitive on the other vertices
Take $y \in X \setminus \{x\}$ then $|G_x| = 3|G_{xy}|$
$G_{xy} = \{e,$ reflection through a plane on the edge $xy\}$
So $|G_{xy}| = 2$ and thus $|G| = 4.3.2 = 24$
It follows that $G = S_4$
Also, if G^+ is the subgroup consisting of rigid motions (i.e. rotationaly symmetries) then the above applies to give $|G^+| = 12$. In fact, G^+ is the alternating group A_4

Exercise: A_4 has no subgroup of order 6 (but $6 \Big| |A_4|$)

Explicit example of left cosets of the stabiliser group:

Take $G = D_8$ the set of symmetries of a square, with b representing reflection in the line 13 and a representing rotation by $\frac{\pi}{2}$ anti-clockwise

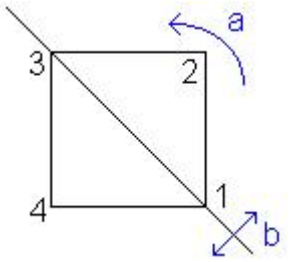

Figure 40: Dihedral group of order 8

$$G_1 = \{e, b\}1 \mapsto 1$$
$$aG_1 = \{a, ab\}1 \mapsto 2$$
$$a^2G_1 = \{a^2, a^2b\}1 \mapsto 3$$
$$a^3G_1 = \{a^3, a^3b\}1 \mapsto 4$$

12.1 General definition of action

Group acting on a set - Let G be a group. Let X be a set. Then G acts on X if there is a mapping $G \times X \to X, (g, \alpha) \mapsto g(\alpha)$ such that

0) $g(\alpha) \in X \quad \forall g \in G, \forall \alpha \in X$
1) $e(\alpha) = \alpha$
2) $(g_1, g_2)(\alpha) = g_1(g_2(\alpha))$

Lemma 91. Assume G as on X. Then each g induces a permutation θ_g. The mapping $\theta : G \to \mathbf{Sym}\, X$ is a homomorphism $g \mapsto \theta_g$

Proof. If $g_1, g_2 \in G$ then $\theta_{g_1}, \theta_{g_2} = \theta_{g_1 g_2}$, as $\theta_{g_1}\theta_{g_2}(\alpha) = g_1(g_2(\alpha)) = (g_1 g_2)(\alpha) = \theta_{g_1 g_2}(\alpha)\ \forall \alpha \in X$

θ is a permutation of X_1 for each $g \in G$:
Certainly, it is a map $X \to X$, it is bijective, since the inverse of θ_g is $\theta_{g^{-1}}$, $\theta_g \theta_{g^{-1}} = \theta_e$ and θ_e is the identity permutation, hence the claim follows □

If G acts on the set X, the kernel of the action is $K = \{g \in G : g(\alpha) = \alpha\ \forall \alpha \in X\}$
Then K is the kernel of the homomorphism θ in the above. Hence:

Lemma 92. If G acts on X, with kernel K, then $K \lhd G$ and $\frac{G}{K} \leq \mathbf{Sym}\, X$

Proof. The above and the isomorphism theorem □

Example 93. Let G be the group of symmetries of a cube, let G^+ be the subgroup of rigid motions (i.e. rotational symmetries). Consider the action of G on the set X of vertices - so $|X| = 8$. Let $\alpha \in X$. The action of G (and of G^+ on X is transitive.
So $|G| = 8|G_\alpha|$. The orbits of G_α on X have sizes
1, 3 (neighbours of α, connected to it by an edge), 3, 1 (antipodal)
G_X is clear transitive on the set of neighbours of α
Let β be a "neighbour" of α
$|G_\alpha| = 3|G_{\alpha\beta}|.|G^+| = 3|G^+_{\alpha\beta}|$
Now $G^+_{\alpha\beta} = 1(\{e\})$ so $|G^+| = 8|G^+_\alpha| = 8.3 = 24$
And $G_{\alpha\beta} = \{e,$ reflection in a mirror in a plane on the edge $\alpha\beta\}$
sp $|G_{\alpha\beta}| = 2$, so $|G| = 8.3.2 = 48$
Now consider G acying on the set X' of the four diagonals (joining antipodal pairs). Then G does act - diagonals are taken to diagonal by symmetries. G is transitive on X' - in fact G^+ is. Let $d_1 \in X'$. Then $G^+_{d_1}$ is transitive on the other 3 diagonals, and in fact $G_{d_1 d_2}$ swaps d_3, d_4
But $G_{d_1 d_2 d_3 d_4}$ is the kernel K of the action of G on X', and has order 2: if the cube has centre in O, then kernel is $\{\pm I\}$, $\frac{G}{k} \simeq S_4, G^+ \simeq S_4$
Later, we shall see that G is in fact isomorphic to the direct product $S_4 \times c_2$ (or $G^+ \times K$)

Example 94. Left coset action
Take G a group, $H \leq G$, let X be the set of all left cosets.
Left coset action: $a(bH) = abH$
This is an action of $G, a \in G, bH \in X$
It is transitive: $cb^{-1}bH = cH$ for $bH, cH \in X$
Stabiliser of H is precisely H - the stabiliser of bH is bHb^{-1}, $g(bH_{=}bH$ iff $b^{-1}gb \in H$ iff $g \in bHb^{-1}$
The kernel of this action is the intersection of all stabiliser groups $= \bigcap_{b \in G} bGb^{-1}$
This is the largest normal subgroup of G contained in H

Left regular action:
Take G a group, and let $X = G$
G acts on X by left multiplication, transitively, and $G_X = \{e\}$ ($g(x) = x \Leftrightarrow g = e$)
Hence a theorem of Cayley:
Any group is isomorphic to a subgroup of a symmetrix group
In fact, here, $G \leq \mathbf{Sym}\, X$

Conjugation - Let G be a group

> i) The subgroups H_1, H_2 of G are G-conjugate if $gHg^{-1} = H$ for some $g \in G$
> ii) The elements $h_1, h_2 \in G$ are conjugate if $gh_1g^{-1} = h_2$ for some $g \in G$

Lemma 95. Assume G acts on X, let $\alpha \in X_1, g \in G$
Then $G_{g(\alpha)} = gG_\alpha g^{-1}$

Proof. $\supset$: if $h \in G_\alpha$ then $ghg^{-1}(g(\alpha)) = gh(\alpha) = g(\alpha)$
$\subset$: If $h \in G_{g(\alpha)}$ then $h(g(\alpha)) = g(\alpha)$ so $g^{-1}hg(\alpha) = \alpha$ so $h \in gG_\alpha g^{-1}$ □

Remark 96. If we know G_α, to study G_β with $\beta \in G(\alpha)$, just onjugate by an element taking α to β

Conjugate Actions:
Let X be the set of all elements of G
Then G acts on X by conjugation, $g : x \mapsto gxg^{-1}$ for $x \in X, g \in G$
It is an action: as it is closed, has the identity fixing x, $(g_1g_2)(x) = g_1g_2xg_2^{-1}g_1^{-1} = g_1(g_2(x))$
Orbits are g-conjugacy classes:
$G(x) = ccl_G x$ (read as "conjugacy class of x in G") $= \{gxg^{-1} : g \in G\}$
$G_x = C_G(x)$ - the centraliser of x in $G = \{g \in G : gx = xg\}$
The kernel of the conjugation action: $C(G)$ (or $Z(G)$) $= \{g \in G : gx = xg \forall x\}$ - the centre of G

Take another action G by conjugation on the set of all subgroups of G. The stabiliser of H here is $N_G(H) = \{g \in G : gHg^{-1} = h\}$, the normaliser of H in G

Lemma 97.
i) $|G| = |C_G(x)|\ |ccl_G x|$ for $x \in G$
ii) $|G| = |N_G(H)|$**.** the size of the ***ccl*** of H in G

13 More examples of groups

Isometries of the plane - maps (both linear and others) preserving distance

$\boldsymbol{G = \{z \mapsto az + b}$ or $\boldsymbol{a\overline{z} + b : a, b \in \mathbb{C}, |A| = 1\}}$
$\boldsymbol{G^+ = \{z \mapsto az + b : a, b \in \mathbb{C}, |a| = 1\}}$ - called direct isometries

Both $\boldsymbol{G, G^+}$ are groups under composition - check

> e.g. Take $\boldsymbol{g_1 : z \mapsto a_1 z + b_1}$, $\boldsymbol{g_2 : z \mapsto a_2 z + b_2}$
> $\boldsymbol{g_2 g_1(z) = a_2(a_1 z + b_1) + b_2 = a_1 a_1 + a_2 b_1 + b_2}$ and as $\boldsymbol{|a_1 a_2| = 1}$ this is still in $\boldsymbol{G}$ or $\boldsymbol{G^+}$ respectively
> Equally the inverse of $\boldsymbol{z \mapsto az + b}$ is $\boldsymbol{z \mapsto a^{-1}z - a^{-1}b}$

In fact, $\boldsymbol{G^+}$ is a subgroup of $\boldsymbol{G}$ of index 2. If $\boldsymbol{g_1, g_2 \in G \setminus G^+}$ then

$$\begin{aligned} g_2^{-1} g_1(z) &= g_2^{-1}(a_1\overline{z} + b_1) \\ &= a_2^{-1}\overline{(a_z\overline{z} + b_1)} - a_2^{-1}b_2 \\ &= \overline{a_1}a_2^{-1}z + (\overline{b_1}a_2^{-1} - a_2^{-1}b_2) \in G^+ \\ \therefore g_1 G^+ &= g_2 G^+ \end{aligned}$$

Recall: The group of isometries of $\mathbb{C}$ under composition.

Proposition 98. Every isometry of $\mathbb{C}$ is the product of a translation, a rotation about O and possibly a reflection in a line through the origin. Hence:

$$\boldsymbol{G = \{z \mapsto az + b \text{ or } z \mapsto a\overline{z} + b : a, b \in \mathbb{C}, |a| = 1\}}$$

is the full group of isometries of $\mathbb{C}$

Proof. Let $\boldsymbol{g}$ be an isometry of $\mathbb{C}$. If $\boldsymbol{g(0) = b}$ then $\boldsymbol{tg(0) = 0}$ where $\boldsymbol{t}$ is the translation $\boldsymbol{z \mapsto z - b}$. Note that $\boldsymbol{t \in G}$
So assume $\boldsymbol{g(0) = 0}$.
If $\boldsymbol{g(1) = a}$, then $\boldsymbol{|a| = 1}$ (as distance from O is fixed)
Then $\boldsymbol{\rho g(1) = 1}$ where $\boldsymbol{\rho}$ is the rotation $\boldsymbol{z \mapsto a^{-1}z}$. Again $\boldsymbol{\rho \in G}$
So can assume $\boldsymbol{g = \rho g}$ fixes 0 and 1
Then $\boldsymbol{g(i)}$ is $\boldsymbol{i}$ or $\boldsymbol{-i}$. If $\boldsymbol{g(i) = -i}$ then $\boldsymbol{rg}$ fixes $\boldsymbol{i}$ (where $\boldsymbol{r : z \mapsto \overline{z}}$) - again $\boldsymbol{r \in G}$
So can assume $\boldsymbol{g = rg}$ fixes $\boldsymbol{0, 1, i}$. Generally, an isometry that fixed 3 non-linear points is the identity (as 3 circles intersect at at most one point, unless their centres are on a line). Thus $\boldsymbol{g}$ is the identity, and thus every isometry in $\boldsymbol{G}$ can be produced from the identity combined with suitable inverses for $\boldsymbol{r, \rho, t}$. □

Note 99. $\boldsymbol{G_O = O_2 = \{z \mapsto az \text{ or } z\overline{z} : |a| = 1\}}$, $\boldsymbol{G_O^+ = SO_2}$
Let $\boldsymbol{v \in \mathbb{C}}$, $\boldsymbol{G_v}$ is the just the conjugate $\boldsymbol{gG_O g^{-1}}$ where $\boldsymbol{g : z \mapsto z + v}$ and $\boldsymbol{g^{-1} : z \mapsto z - v}$
$\boldsymbol{G_v^+ = \{z \mapsto az + v(1 - a)\}}$ - rotations about v

Remark 100. More generally, any isometry of $\mathbb{R}^n$ can be written as a composite of a translation and a linear orthogonal map in O_n.
Let g be an isometry. Composing with a translation, we may assume $g(0) = 0$. Now claim that such a g is linear, and hence orthogonal. Now g preserves distance, so it also preserves inner products (as $x.y = \frac{1}{4}(|x+y| - |x-y|)$) Hence g takes the standard basis of $\mathbb{R}^n$ to an orthonormal set of vectors, so to an orthonormal basis. Composing g with a suitable element of O_n, we may assume g fixes the standard basis. But then g fixes $\mathbb{R}^n$ elementwise, so $g = e$, because given a general member of $\mathbb{R}^n$, say $\underline{v} = \sum a_i \underline{e_i}$

$$\begin{aligned} g(\underline{v}).g(\underline{e_i}) &= g(\underline{v}).\underline{e_i} \text{ as } g \text{ fixes } \underline{e_i} \\ &= \underline{v}.\underline{e_i} \text{ as } g \text{ preserves inner products} \\ &= a_i \end{aligned}$$

So $g(\underline{v}) = \sum a_i \underline{e_i} = \underline{v}$

There is a related group on $\mathbb{C}$ - the group of similarities on $\mathbb{C}$

$$\{z \mapsto az + b \text{ or } a\overline{z} + b : a, b \in \mathbb{C}, a \neq 0\}$$

which we will use later.

13.1 Möbius Groups

We will study maps $\boldsymbol{T}:\mathbb{C}\to\mathbb{C}$ with the following properties:

$$z\mapsto\frac{az+b}{cz+d}\quad a,b,c,d\in\mathbb{C}\quad \left|\begin{array}{cc} a & b\\ c & d\end{array}\right|\neq 0$$

If $c\neq 0$, we get a problem at $z=-\frac{d}{c}$
Add a new point, ∞ to $\mathbb{C}$ to form the extended complex plane $\boldsymbol{C}\cup\{\infty\}$. Now $\boldsymbol{T}:\mathbb{C}\cup\{\infty\}\to\boldsymbol{C}\cup\{\infty\}$

$$c=0\begin{cases} T(z) &= \frac{az+b}{d}\ z\in\mathbb{C}\\ T(\infty) &= \infty\end{cases}$$

$$c\neq 0\begin{cases} T(z) &= \frac{az+b}{cz+d}\ z\neq -\frac{d}{c}\\ T(-\frac{d}{c}) &= \infty\\ T(\infty) &= \frac{a}{c}\end{cases}$$

At present, ∞ is just a new point, useful here. Do not use it elsewhere (e.g. in Analysis I).

Under composition: Let $\boldsymbol{T_i}(z)=\frac{a_iz+b_i}{c_iz+d_i}\quad \left|\begin{array}{cc} a_i & b_i\\ c_i & d_i\end{array}\right|\neq 0$

$$\begin{aligned} T_1\circ T_2(z) &= \frac{a_1\frac{a_2z+b_2}{c_2z+d_2}+b_1}{c_1\frac{a_2z+b_2}{c_2z+d_2}+d_1}\\ &= \frac{(a_1a_2+b_1c_2)z+(a_1b_2+b_1d_2)}{(c_1a_2+d_1c_2)z+(c_1b_2+d_1d_2)}\end{aligned}$$

Now $(a_1a_2+b_1c_2)(c_1b_2+d_1d_2)-(a_1b_2+d_2b_1)(c_1a_2+d_1c_2)=(a_1d_1-b_1c_1)(a_2d_2-b_2c_2)\neq 0$ If we write $\boldsymbol{M}$ for the set of all Möbius transformations, then $\boldsymbol{M}$ is closed under composition. Composition is associative, and we have:
Identity: $z\mapsto z\quad a,d=1,b,c=0\quad \left|\begin{array}{cc} a & b\\ c & d\end{array}\right|=1\neq 0$
Inverse: Given a map $z\mapsto\frac{az+b}{cz+d}$ the inverse is $z\mapsto\frac{dz-b}{-cz+a}$ and $\left|\begin{array}{cc} d & -b\\ -c & a\end{array}\right|\neq 0$ if $\left|\begin{array}{cc} a & b\\ c & d\end{array}\right|\neq 0$.
This leads us to:

Theorem 101. The set of all Möbius transformations on $\mathbb{C}$ is a group, $\boldsymbol{M}$, under composition

$$\begin{aligned} M:\mathbb{C}\cup\{\infty\}&\to\mathbb{C}\cup\{\infty\}\\ z:&\mapsto\frac{az+b}{cz+d}\quad z\in\mathbb{C}\setminus\{-\frac{d}{c}\}\\ \infty:&\mapsto\frac{a}{c}\\ -\frac{d}{c}:&\mapsto\infty\end{aligned}$$

Proof. All the necessary properties have been checked above □

Now we have a connection to $\boldsymbol{GL_2(\mathbb{C})}$, the group of all 2x2 non-symmetrix complex matrices

Given $\boldsymbol{A \in GL_2(\mathbb{C}), A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}}$ with $\boldsymbol{\det A \neq 0}$ consider the Möbius transformation $\boldsymbol{T_A : z \mapsto \frac{az+b}{cz+d}}$

The map $\boldsymbol{A \mapsto T_A}$ is a homomorphism $\boldsymbol{GL_2(\mathbb{C}) \to M}$

$\boldsymbol{\theta(A_1A_2) = \theta(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}) = \theta \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix}}$

$\boldsymbol{\therefore T_{A_1}T_{A_2} = \theta(A_1)\theta(A_2)}$

Also, $\boldsymbol{\theta}$ is surjective - clear

The kernel: $\boldsymbol{\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} : z = \frac{az+b}{cz+d}\}, \left| \begin{matrix} a & b \\ c & d \end{matrix} \right| \neq 0, \forall z in \mathbb{C} \cup \infty\}}$

if $\boldsymbol{z = \infty \frac{az+b}{cz+d} = \infty}$ iff $\boldsymbol{c = 0}$

if $\boldsymbol{z = 0}$ $\boldsymbol{\frac{a(0)+b}{d} = 0}$ iff $\boldsymbol{b = 0}$

if $\boldsymbol{z = 1}$ $\boldsymbol{\frac{a(1)}{d(1)} = 1}$ iff $\boldsymbol{a = d}$

$\boldsymbol{\therefore \ker(\theta) = \{\lambda I : \lambda \in \mathbb{C}, \lambda \neq 0\} = Z}$

$$\therefore \boldsymbol{M \simeq \frac{GL_2(\mathbb{C})}{Z}} \text{ by the isomorphism theorem}$$
$$= \boldsymbol{PGL_2(\mathbb{C})} \text{ the projective general linear group}$$

Note: $\boldsymbol{T_A = T_B}$ iff $\boldsymbol{A = \lambda B}$ for some $\boldsymbol{\lambda \in \mathbb{C} \setminus \{0\}}$

In fact, also $\boldsymbol{M \simeq \frac{SL_2(\mathbb{C})}{\{\pm I\}} = PSL_2(\mathbb{C})}$

Theorem 102. The group $\boldsymbol{M}$ is isomorphic to $\boldsymbol{\frac{SL_2(\mathbb{C})}{\{\pm I\}} = PSL_2(\mathbb{C})}$

Proof. Consider the restriction of $\boldsymbol{\theta}$ above to $\boldsymbol{SL_2(\mathbb{C})}$ - this is still surjective, since it determines a matrix only modulo scalars. The kernel becomes $\boldsymbol{\{\pm I\}}$ (scalars $\boldsymbol{\det 1}$) □

13.2 Permutation properties of M

The elements of $\boldsymbol{M}$ are permutations of $\mathbb{C} \cup \{\infty\}$

Theorem 103. If z_1, z_2, z_3 and w_1, w_2, w_3 are two triples of distinct points in $\mathbb{C} \cup \{\infty\}$ there is a unique Möbius transformation $\boldsymbol{T}$ with $\boldsymbol{T}(z_i) = w_i \ \forall i$ (so $\boldsymbol{M}$ is sharply transitive).

Proof. Existence - First, the special case $w_1 = \infty, w_2 = 0, w_3 = 1$. Take:

$$\boldsymbol{T}(z) = \frac{z - z_2}{z - z_1} . \frac{z_3 - z_1}{z_3 - z_2}$$

(assuming here that all the $z_i \in \mathbb{C}$, otherwise we have to use one of the following:

$$z_1 = \infty \Rightarrow \boldsymbol{T}(z) = \frac{z - z_2}{z_3 - z_2}$$
$$z_2 = \infty \Rightarrow \boldsymbol{T}(z) = \frac{z_3 - z_1}{z - z_1}$$
$$z_3 = \infty \Rightarrow \boldsymbol{T}(z) = \frac{z - z_2}{z - z_1}$$

which then satisfies the requirements)
Next - the general case:
Let $\boldsymbol{T_1}$ take z_1, z_2, z_3 to $\infty, 0, 1$
Let $\boldsymbol{T_2}$ take w_1, w_2, w_3 to $\infty, 0, 1$
$\boldsymbol{T_2^{-1}} \circ \boldsymbol{T_1}$ is a Möbius map taking z_i to w_i as required.
Uniqueness - This comes from the next 2 lemmas
First the stabiliser $\boldsymbol{M_{\infty 01}} = \{e\}$

Lemma 104. $\boldsymbol{M_{\infty 01}} = \{e\}$

Proof. $\boldsymbol{M_\infty} = \{z \mapsto \frac{az+b}{d} : ad \neq 0\} = \{z \mapsto a'z + b' : z' \neq 0\}$ where $a' = \frac{a}{d}, b' = \frac{b}{d}$ - the group of direct similarities
$\boldsymbol{M_{\infty 0}} = \{z \mapsto a'z : a' \neq 0\}$
$\boldsymbol{M_{\infty 01}} - \{z \mapsto z\} = \{e\}$ □

Next, we prove this in general:

Lemma 105. $\boldsymbol{M_{z_1 z_2 z_3}} = \{e\}$

Proof. Let $\boldsymbol{T}$ take $z_1 \mapsto \infty, z_2 \mapsto 0, z_3 \mapsto 1$
as above $\boldsymbol{M_{\infty 01}} = \boldsymbol{T} \therefore \boldsymbol{M_{\infty 01}} = e$ □

Finally, if $\boldsymbol{T_1}, \boldsymbol{T_2}$ both send z_1 to w_1, z_2 to w_2, z_3 to w_3 then $\boldsymbol{T_2^{-1} T_1} \in \boldsymbol{M_{\infty 01}} = \{e\} \Rightarrow \boldsymbol{T_1} = \boldsymbol{T_2}$ and thus unique □

Remark: A Möbius maps fixing three points of $\mathbb{C} \cup \{\infty\}$ must be trivial, because if $z \mapsto \frac{az+b}{cz+d} = z$ has at least three roots, given that it has at most degree 2 it must be trivial.

Lemma 106. If $\boldsymbol{T(z) = \frac{az+b}{cz+d}}$ with $\boldsymbol{ad \neq bc \Rightarrow ad - bc \neq 0}$ then $\boldsymbol{T}$ is either a direct similarity (so $\boldsymbol{c = 0}$) or can be written as a composite of an inversion and direct similarities. More precisely, if $\boldsymbol{c = 0}$ then $\boldsymbol{T = T_1 T_2}$ with

$$\boldsymbol{T_2(z) = \frac{a}{d} z} \text{ rotation and dilation}$$
$$\boldsymbol{T_1(z) = z + \frac{b}{d}} \text{ translation}$$

If $\boldsymbol{c \neq 0}$ then $\boldsymbol{T = T_1 T_2 T_3 T_4}$ with

$$\boldsymbol{T_4(z) = z + \frac{d}{c}} \text{ translation}$$
$$\boldsymbol{T_3(z) = \frac{1}{z}} \text{ inversion}$$
$$\boldsymbol{T_2(z) = -\frac{ad - bc}{c} z} \text{ rotation and dilation}$$
$$\boldsymbol{T_1(z) = z + \frac{a}{c}} \text{ translation}$$

Proof. This is immediate □

Remark: $\boldsymbol{\frac{az+b}{cz+d} = \frac{a}{c} - \frac{ad-bc}{c(cz+d)}}$ (replacing $\boldsymbol{\infty \mapsto \frac{a}{c}}$ by $\boldsymbol{\infty \mapsto 0}$)
$\therefore \boldsymbol{-\frac{ad-bc}{c(cz+d)} = -\frac{ad-bc}{c^2} \cdot \frac{1}{z+\frac{a}{c}}}$

Lemma 107. The general equation of a circle on a straight line in $\mathbb{C}$ is:

$$\boldsymbol{Az\overline{z} + Bz + \overline{B}\overline{z} + C = 0}$$

where $\boldsymbol{A, C \in \mathbb{R}, |B|^2 > AC}$

Proof. Substitute in $\boldsymbol{z = x + iy}$ □

Proposition 108. A Möbius transformation takes $\boldsymbol{\{}$circles, straight lines$\boldsymbol{\}} \to \boldsymbol{\{}$circles, straight lines$\boldsymbol{\}}$ - where we have "circles" $\mathbb{C} \cup \boldsymbol{\{\infty\}}$ Note: Not respectively

(!!MISSING PAGE!!)
(!!MISSING PAGE!!)
(!!MISSING PAGE!!)
(!!MISSING PAGE!!)
(!!MISSING PAGE!!)
(!!MISSING PAGE!!)
(!!MISSING PAGE!!)
(!!MISSING PAGE!!)
(!!MISSING PAGE!!)

Say $\boldsymbol{A}$ is the kernel. Then

$$\begin{aligned} \boldsymbol{A}\underline{\boldsymbol{e_1}} &= \boldsymbol{\lambda}\underline{\boldsymbol{e_1}} \\ \boldsymbol{A}\underline{\boldsymbol{e_2}} &= \boldsymbol{\mu}\underline{\boldsymbol{e_2}} \\ \Rightarrow \boldsymbol{A}(\underline{\boldsymbol{e_1}} + \underline{\boldsymbol{e_2}}) &= \boldsymbol{\lambda}\underline{\boldsymbol{e_1}} + \boldsymbol{\mu}\underline{\boldsymbol{e_2}} \end{aligned}$$

Which implies that $\boldsymbol{\lambda, \mu}$ are scalar multiples.

Now 1-dimensional subspaces of $\mathbb{C}^{\boldsymbol{2}}$ are in correspondance with "slopes"

$\left\langle \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle \leftrightarrow z_1, z_2 \in \mathbb{C} \cup \{\infty\}$

The action:

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \left\langle \begin{pmatrix} z \\ 1 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} az+b \\ cz+d \end{pmatrix} \right\rangle \quad \text{so} \quad z \mapsto \tfrac{az+b}{cz+d} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} a \\ c \end{pmatrix} \right\rangle \quad \text{so} \quad \infty \mapsto \tfrac{a}{c} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \left\langle \begin{pmatrix} d \\ -c \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} ad-bc \\ 0 \end{pmatrix} \right\rangle \quad \text{so} \quad -\tfrac{d}{c} \mapsto \infty \end{aligned}$$

Remark 109. By the theorem on normal forms of complex 2x2 matrices (Theorem 48, p. 102) each matrix in $\boldsymbol{GL_2}(\mathbb{C})$ is conjugate to one of

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \Rightarrow$ 2 1-dimensional eigenspaces, so $\boldsymbol{T_A}$ fixed 2 points

$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \Rightarrow$ 1 1-dimensional eigenspace, so $\boldsymbol{T_A}$ fixed 1 point only

$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \Rightarrow \boldsymbol{T_A}$ trivial

Hence any möbius map fixes 1, 2 or all the points of $\mathbb{C} \cup \{\infty\}$

The point ∞ - geometric view:

Let $\boldsymbol{S}$ be the unit sphere in $\mathbb{R}^{\boldsymbol{3}}$, with centre O. Let $\mathbb{C}$ be the equatorial horizontal plane $\{x, y, 0\}$
Let ζ be the north pole of $\boldsymbol{S} \therefore \zeta = (0, 0, 1)$
Consider the function $\phi : \mathbb{C} \to \boldsymbol{S} \setminus \{\zeta\}, z \mapsto z'$ stereographical projection
The line ζz meets $\boldsymbol{S}$ at the unique point z' so $\phi(z) = (\frac{2x}{|z|^2+1}, \frac{2y}{|z|^2+1}, \frac{|z|^2-1}{|z|^2+1})$
where $z = x + iy = (x, y, 0)$
(It is $(0, 0, 1) + t[(x, y, 0) - (0, 0, 1)] = (tx, ty, 1-t)$ with $(tx)^2 + (ty)^2 + (1-t)^2 = 1$)

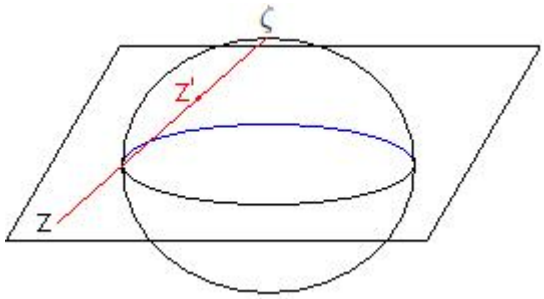

Figure 41: Stereographical Projection onto the plane

As z gets large, $\phi(z)$ gets closer to ζ. Adjoin a point ∞ to $\mathbb{C}$ and put $\phi(\infty) = \zeta$ so now we have a bijection $\phi : \mathbb{C} \cup \{\infty\} \to S$
Now consider $T = T_A$ on $\mathbb{C} \cup \{\infty\}$ in terms of S. $\phi T \phi^{-1} : S \to S$. One can check that points near to $\phi(-\frac{d}{c})$ map to points near to ζ, to get continuity of S we define $(\phi T \phi^{-1})(\phi(-\frac{d}{c})) = \zeta$ so $T(-\frac{d}{c}) = \infty$. Similarly, $T(\infty) = \frac{a}{c}$

Remark on circles and lines
Let L be a straight line in the equatorial plane of $\mathbb{C}$. Let Π be the plane cutting L and ζ, then $\mathcal{C} = \Pi \cap S$ is a circle on S containing ζ, and ϕ maps L to $\mathcal{C} \setminus \{\zeta\}$.
It is reasonable to adjoin ∞ to L and think of it as a "circle" in $\mathbb{C} \cup \{\infty\}$ - its image under ϕ is $\mathcal{C}$.
Conversely, starting from a circle $\mathcal{C}$ on S through ζ, then ϕ^{-1} takes $\mathcal{C} \setminus \{\zeta\}$ to a straight line in the equatorial plane $\mathbb{C}$, so $\phi^{-1}(\mathcal{C})$ is a circle in $\mathbb{C} \cup \{\infty\}$
Also we can check:

$$\text{circles in } \mathbb{C} \overset{\phi}{\longleftrightarrow} \text{circles on } S \setminus \{\zeta\}$$

14 Symmetric and Alternating Groups

Recall: A permutation is a bijection $X \to X$
$\mathbf{Sym}\, X$ is the group of permutations of X under composition
If $|X| = n$ write $\mathbf{Sym}_n$ or S_n for $\mathbf{Sym}\, X$

Another notation: Disjoint cycle notation:

Start at $i \in X$: $(\; i \quad f(i) \quad f^2(i) \quad \dots \quad f^{k-1}(i) \;)(\; j \quad f(j) \quad f^2(j) \quad \dots \quad f^{l-1}(j) \;)(\dots)$ with the entries all distinct, i.e. $k, l, \dots$ are minimal such that $f^k(i) = i, f^l(j) = j, \dots$

Example: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 4 & 6 \end{pmatrix} = (\; 1 \quad 2 \quad 3 \;)(\; 4 \quad 5 \;)(\; \cancel{6} \;) = (\; 1 \quad 2 \quad 3 \;)(\; 4 \quad 5 \;)$

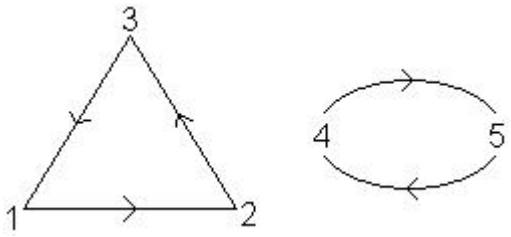

Figure 42: Disjoint cycle on 6 elements

$f = (\; i_1 \quad i_2 \quad \dots \quad i_k \;)$ is a k-cycle

Note 110. i) $(\; 1 \quad 2 \quad 3 \;) = (\; 2 \quad 3 \quad 1 \;) = (\; 3 \quad 1 \quad 2 \;)$
ii) $(\; 1 \quad 2 \quad 3 \;)(\; 4 \quad 5 \;) = (\; 4 \quad 5 \;)(\; 1 \quad 2 \quad 3 \;)$ - cycles commute
iii) But when composing cycles which are not disjoint (not here) they do not commute in general e.g.:

$$(\; 1 \quad 2 \;)(\; 2 \quad 3 \;) = (\; 1 \quad 2 \quad 3 \;)$$
$$(\; 2 \quad 3 \;)(\; 1 \quad 2 \;) = (\; 1 \quad 3 \quad 2 \;)$$

Lemma 111. Any permutation can be written as the product of disjoint cycles in an essentially unique way

Proof. Let $f \in \mathbf{Sym}\, X$. Put $i \in X$, follow it through until $f^k(i) = i$, so we have $(\; i \quad f(i) \quad f^2(i) \quad \dots \quad f^{k-1}(i) \;)$. This limits the $< f >$ orbit X_1 of i in X.
Now take $j \in X \backslash X_1$ (if such exists), repeat so we have $(\; j \quad f(j) \quad f^2(j) \quad \dots \quad f^{l-1}(j) \;)$ is the $< f >$ orbit X_2 of j in X
Now take $r \in X \setminus (X_1 \cup X_2)$, continue
To show that it is essentially unique: cycles list elements of $< f >$ orbits in cyclic order. □

Lemma 112. If $f \in \mathbf{Sym}\, X$ then the order of f is the least common multiple of all cycle lengths of f in the disjoint cycle decomposition

Proof. This comes from any k-cycle having order k, and properties like

$$[(\ 1\quad 2\quad 3\ )\,(\ 4\quad 5\ )]^l = (\ 1\quad 2\quad 3\ )^l\,(\ 4\quad 5\ )^l$$

as they commute □

Cycle Tuple - The cycle tuple of f is an ordered v-tuple $(n_1, n_2, \ldots, n_v)$ where f has cycles of length $n_1, n_2, \ldots, n_v$ in a disjoint cycle representation, and $n_1 \geq n_2 \geq \cdots \geq 1, \sum_i n_i = n$

Theorem 113. Two permutations in S_n are conjugate iff they have the same cycle type i.e. they have the same number m_k of k-cycles $\forall k$

Example: Conjugacy classes of S_4

	cycle type	size
$(\ 1\quad 2\quad 3\quad 4\ )$	4	6
$(\ 1\quad 2\quad 3\ )$	$3\quad 1$	$4 \times 2 = 8$
$(\ 1\quad 2\ )(\ 3\quad 4\ )$	$2\quad 2$	3
$(\ 1\quad 2\ )$	$2\quad 1\quad 1$	6
e	$1\quad 1\quad 1\quad 1$	1

Proof. If $g \in S_n$ and $(\ i_1\quad i_2\quad i_3\quad \ldots\quad i_k\ )$ is a k-cycle, then $g\,(\ i_1\quad i_2\quad i_3\quad \ldots\quad i_k\ )\,g^{-1}$ is the k-cycle $(\ g(i_1)\quad g(i_2)\quad g(i_3)\quad \ldots\quad g(i_k)\ )$
$g\,(\ i_1\quad i_2\quad i_3\quad \ldots\quad i_k\ )\,g^{-1}$ sends $g(i_1)$ to $g(i_2)$ and $g(i_2)$ to $g(i_3)$ and $\ldots$ and $g(i_k)$ to $g(i_1)$
If $h = h_1\ h_2\ \ldots\ h_v$ with the h_j disjoint cycles then claim

$$\begin{aligned} ghg^{-1} &= g(h_1\ h_2\ \ldots\ h_v)g^{-1} = gh_1g^{-1}\ gh_2g^{-1}\ \ldots\ gh_vg^{-1} \\ &= (\ g(i_1)\quad g(i_2)\quad \ldots\quad g(i_k)\ )\,(\ g(i_{k+1})\quad \ldots\ ) \end{aligned}$$

if $h_1 = (\ i_1\quad i_2\quad \ldots\quad i_k\ ), h_2 = (\ i_{k+1}\quad \ldots\ )$
$\Rightarrow h = h_1\ h_2\ \ldots\ h_v$ disjoint cycles

$$\begin{aligned} ghg^{-1} &= gh_1g^{-1}\ gh_2g^{-1}\ \ldots\ gh_vg^{-1} \\ &= g\,(\ i_1\quad i_2\quad \ldots\quad i_k\ )\,(\ i_{k+1}\quad \ldots\ )\ldots(\ \ldots\ )\,g^{-1} \\ &= (\ g(i_1)\quad g(i_2)\quad \ldots\quad g(i_k)\ )\,(\ g(i_{k+1})\quad \ldots\ )\ldots \end{aligned}$$

$\Leftarrow$ if $x_1, x_2 \in S_n$ of the same cycle type
$x_1 = (\ i_1\quad i_2\quad \ldots\quad i_k\ )\,(\ i_{k+1}\quad i_{k+2}\quad \ldots\quad i_l\ )\ldots$
$x_2 = (\ j_1\quad j_2\quad \ldots\quad j_k\ )\,(\ j_{k+1}\quad j_{k+2}\quad \ldots\quad j_l\ )\ldots$
Take $g = \begin{pmatrix} i_1 & i_2 & \ldots & i_k & i_{k+1} & \ldots \\ j_1 & j_2 & \ldots & j_k & j_{k+1} & \ldots \end{pmatrix}$ then $gx_1g^{-1} = x_2$ □

Back to our example, $\boldsymbol{S_4}$ = conjugacy classes in $\boldsymbol{S_4}$

	cycle type	size	centraliser size	sign
$(\mathbf{1\ 2\ 3\ 4})$	$\mathbf{4}$	$\mathbf{6}$	$\mathbf{4}$	$-$
$(\mathbf{1\ 2\ 3})$	$\mathbf{3\ 1}$	$\mathbf{8}$	$\mathbf{3}$	$+$
$(\mathbf{1\ 2})(\mathbf{3\ 4})$	$\mathbf{2\ 2}$	$\mathbf{3}$	$\mathbf{8}$	$+$
$(\mathbf{1\ 2})$	$\mathbf{2\ 1\ 1}$	$\mathbf{6}$	$\mathbf{4}$	$-$
$\boldsymbol{e}$	$\mathbf{1\ 1\ 1\ 1}$	$\mathbf{1}$	$\mathbf{24}$	$+$

Note: We have the centraliser sizes from Lemma 97, p. 133.

14.1 Digression on Normal Subgroups

Lemma 114. A subgroup K of G is normal in G iff K is the union of G's conjugacy classes

Proof. If $K \subset G$ is a union of G-conjugacy classes then for $k \in K, g \in G, gkg^{-1} \in K$ so $K \lhd G$
Conversely, if $K \lhd G$ with each $k \in K$, have all $ghg^{-1} \in K \ \forall g \in G$ so $ccl_g(k) \subset K$ □

Example: Normal subgroups of S_4 must contain e, and the order must divide 24, and must be a union of S_4 - conjugacy classes (see table). So:
order 1 = $\{e\}$
order 4 = $\{e, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\} = V_4$
order 12 = A_4 - group of rotations on regular tetrahedrons
order 24 = S_4

Quotients of S_4

$\frac{S_4}{\{e\}}$	$\frac{S_4}{V_4}$	$\frac{S_4}{A_4}$	$\frac{S_4}{S_4}$
	6	12	
S_4	S_3	C_2	$\{e\}$

Now alternating groups
A transposition is a permutation transposing 2 points and fixing the rest - a 2-cycle

Proposition 115. S_4 is generated by transpositions, that is, every permutation can be written as a product of transpositions

Proof. Any permutation is a disjoint product of cycles, so it is enough to show that each cycle is a product of transpositions, but

$$(i_1\ i_2\ \dots\ i_k) = \underbrace{(i_1\ i_2)(i_2\ i_3)\dots(i_{k-1}\ i_k)}_{k-1 \text{ transpositions}}$$

□

Note that this is not unique:

$$\begin{aligned}(1\ 2\ 3) &= (1\ 2)(2\ 3) \\ &= (1\ 2)(3\ 4)(2\ 4)(3\ 4) \\ &= \dots\end{aligned}$$

Theorem 116. For $g \in S_n$, define sign g (denoted $\varepsilon(g)$) $= (-1)^k$ where k is the number of transpositions in some expression of g as a product of transpositions. Then sign is well defined and sign: $S_n \to \{\pm 1\}$ under multiplication is a homomorphism

Even permutation - $g \in S_n$ is an even permutation if $\text{sign}(g) = 1$

Odd permutation - $g \in S_n$ is an odd permutation if $\text{sign}(g) = -1$

The set of even permutations $\in S_n$ forms a normal subgroup A_n, the alternating group of degree n

Note: $|A_n| = \frac{1}{2}n!$. To see this directly, observe that $S_n \setminus A_n$ is one coset of A_n, since the product of any 2 odd permutations is even, and hence is in A_n

Proof. Assume sign is well defined, then it is a homomorphism $S_n \to \{\pm 1\}$:
On $g = t_1 t_2 \dots t_a, h = u_1 u_2 \dots u_b$ with t_i, u_j transpositions. Then sign $g = (-1)^a$, sign $h = (-a)^b$ and $gh = t_1 t_2 \dots t_a u_1 u_2 \dots u_b$ so sign $gh = (-1)^{a+b} = (-1)^a (-1)^b =$sign $g$sign h
$\therefore$ homomorphism.
And define it on $\begin{pmatrix} 1 & 2 \end{pmatrix}$ as -1 and on e as 1.
Now just need to show that sign g is well defined.
For any $g \in S_n$ write $c(g)$ to be the number of cycles in a disjoint cycle representations of g (including fixed points). Claim that if t is a transposition then $c(gt) = c(g) \pm 1 \equiv c(g) + 1 \mod 2$. Note that this claim will be enough, as if $g = t_1 t_2 \dots t_a = u_1 u_2 \dots u_b$ with t_i, u_j transpositions then $c(g) \equiv n + a \equiv n + b \mod 2$ with $c(e) = n$, so $a \equiv b \mod 2$ and so $(-1)^a = (-1)^b$
To prove the claim take $t = \begin{pmatrix} k & l \end{pmatrix}$

Proof. Clearly, all g-cycles not involving k or l remain unaffected by multiplication by t. So consider only g-cycles involving k and l
Case 1: k, l in the same g-cycle
Changing notation, we consider $\begin{pmatrix} 1 & \dots & k-1 & k & k+1 & \dots & l-1 & l & l+1 & \dots & t \end{pmatrix} \circ \begin{pmatrix} k & l \end{pmatrix} = \begin{pmatrix} 1 & \dots & k-1 & k & l+1 & \dots & t \end{pmatrix} \begin{pmatrix} k+1 & k+2 & \dots & l-1 & l \end{pmatrix}$.
So in this case $c(gt) = c(g) + 1$
Case 2: k, l in different cycles of g (want to show here that $c(gt) = c(g) - 1$)
Consider: $\begin{pmatrix} 1 & \dots & k & \dots & s \end{pmatrix} \begin{pmatrix} r & \dots & l & \dots & v \end{pmatrix} \circ \begin{pmatrix} k & l \end{pmatrix} = \begin{pmatrix} 1 & \dots & k-1 & k & l+1 & \dots & v$
□

□

Lemma 117. $g \in S_n$ is an even permutation if the number of cycles of even length is even, and odd otherwise (if the number of cycles of even length is odd)

Examples: $\begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix}$ odd
$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix}$ even
A k-cycle is the product of $k-1$ transpositions, so cycle type tells us whether a permutation is even or odd. This leads to a second proof of Theorem 116

Proof. Consider the polynomials in n variables $x_1, \dots, x_n$ such as $\Delta(x_1, \dots, x_n) = \prod_{1 \le i \le j \le n} (x_i - x_j)$
Define an action of $S_n : g \in S_n$, let g permute the variables

e.g. $g\Delta(x_1, \ldots, x_n) = \prod_{i<j}(x_{g(i)} - x_{g(j)})$
Now, for our Δ, we have $g\Delta = \Delta$ or $-\Delta$ (since g is a permutation of $\{1, \ldots, n\}$)

So write $\varepsilon(g) = +1$ or -1 (recall that $\varepsilon(g) =$sign g)
So $g(\Delta) = \varepsilon(g)\Delta$. Then

$$\begin{aligned} g_1 g_2(\Delta) &= g_1(\varepsilon(g_2(\Delta))) \\ &= \varepsilon(g_2) g_1(\Delta) \\ &= \varepsilon(g_2)\varepsilon(g_1)\Delta \\ &= \varepsilon(g_1 g_2)\Delta \end{aligned}$$

So $\varepsilon(g_1 g_2) = \varepsilon(g_1)\varepsilon(g_2)$ so ε is a homomorphism from S_n to $\{\pm\}$, onto since $\varepsilon \begin{pmatrix} 1 & 2 \end{pmatrix} = -1$ □

Example:

$$\begin{aligned} \begin{pmatrix} 1 & 2 \end{pmatrix}(\Delta) = (x_2 - x_1)(x_2 - x_3)(x_2 - x_4)\ldots(x_2 - x_n) \\ (x_1 - x_3)(x_1 - x_4)\ldots(x_1 - x_n) \\ (x_3 - x_4)\ldots \end{aligned}$$

which contains all the same terms as Δ except that the first one is reversed.

We can apply this to our stated definition of determinant:

Given A an nxn matrix, with $A = (a_{ij})$ then

$$\det A = \sum_{g \in S_n} \varepsilon(g) a_{1g(1)} a_{2g(2)} \ldots a_{ng(n)}$$

15 Small Groups

We have already seen that groups of prime order are cyclic. In fact, if $|G| = p$, then $G \simeq C_p \simeq \frac{\mathbb{Z}}{<p>}$ (by Lagrange)
e.g. if $G =< p >$ then $\theta : G \to \frac{\mathbb{Z}}{<p>}, g^i \mapsto i$

Lemma 118. i) If G has even order, then G contains an element of order 2
ii) If all elements of $G \setminus \{e\}$ have order 2, then G is abelian: if $|G| > 2$ then $4 \Big| |G|$ (in fact, $|G| = 2^n$ for some n)

Proof. i) Elements of G of order > 2 come in inverse pairs; e and the elements of order 2 are self inverse. Hence G has an odd number of elements of order 2
ii) Assume all elements in $G \setminus \{e\}$ have order 2
Let $a, b \in G$. Then $ab = (ab)^{-1} = b^{-1}a^{-1} = ba$ so G is abelian. And $\{e, a, b, ab\}$ is a subgroup of G of order 4. So $4 \Big| |G|$ if $|G| > 2$ i.e. $a \neq b$ □

Proposition 119. Any group of order $2p$ with p an odd prime is either cyclic or dihedral

$$< a : a^{2p} = e > \qquad < a, b : a^p = b^2 = e, bab^{-1} = a^{-1} >$$

Proof. Let G be a group of order $2p$
Let $a, b \in G$ with $o(a) = p, o(b) = 2$ - these exist by Lagrange and the preceding Lemma.
Then $< a > \triangleleft G$ (since it has index 2)
So $bab^{-1} \in < a >$, so $bab^{-1} = a^i$ for some i
Now $a = b^2ab^{-2} = b(bab^{-1}b^{-1} = ba^ib^{-1} = (bab^{-1})^i = a^{i^2}$ so $p | i^2 - 1$, hence $i = \pm 1$.
If $bab^{-1} = a$ then a, b commute, so $o(ab) = 2p$ - so G is cyclic, and hence isomorphic to C_{2p}
Otherwise G is generated by a, b with $a^p = e = b^2$ and $bab^{-1} = a^{-1}$
It follows that G is dihedral, and $G \simeq D_{2p}$

G	D_{2p}
a, b	α - rotation, β - reflection
$a^p = e$	$\alpha^p = e$
$b^2 = e$	$\beta^2 = e$
$ba^i = a^{-1}b$	$\beta\alpha^i = \alpha^{-1}\beta$
$bab^{-1} = a$	$\beta\alpha\beta^{-1} = \alpha^{-1}$

Let $\theta : a^ib^j \mapsto \alpha^i\beta^j$, this is a homomorphism.
Check: $\theta(a^ib^ja^kb^l) = \theta(a^ib^j)\theta(a^kb^l)$
$\theta(a^ib^ka^kb^l) = \theta(a^{i+(-1)^jk}b^{k+1}) \mapsto \alpha^{i+(-1)^jk}\beta^{k+1}$
$\theta(a^ib^j)\theta(a^kb^l) \mapsto \alpha^i\beta^j\alpha^k\beta^l = \alpha^{i+(-1)^jk}\beta^{k+1}$ □

Remark on $\frac{S_4}{V_4} \simeq S_3$
$\frac{S_4}{V_4}$ is a group of order $6 = 2.3$, and 3 is prime
So is isomorphic to C_6 or D_6
Not isomorphic to $C)6$ as S_4 has no elements of order 6
So $\frac{S_4}{V_4}$ is isomorphic to $D_6 = S_3$

Direct product of groups If A and B are groups, then $A \times B = \{(a,b) : a \in A, b \in B\}$ is a group under $(a_1,b_1)(a_2,b_2) = (a_1a_2, b_1b_2)$. This is called the direct product of groups

Lemma 120. If $H, K \lhd G$ with $H \cap K = \{e\}$, then $HK = \{hk : h \in H, k \in K\}$ is a subgroup of G isomorphic to $H \times K$

Proof. Note first that $hk = kh \; \forall h \in H, k \in K$, since $h^{-1}k^{-1}hk \in H \cap K$ as both $H, K \lhd G$, and so this is equal to $\{e\}$
$HK \leq G : h_1k_1h_2k_2 = (h_1h_2)(k_1k_2) \in HK$ and $(hk)^{-1} = k^{-1}h^{-1} = h^{-1}k^{-1} \in HK$
Further, if $h_1k_1 = h_2k_2$ then $h_2^{-1}h_1 = k_2k_1^{-1} \in H \cap K = \{e\}$, so $h_1 = h_2, k_1 = k_2$
Define $\theta : HK \to H \times K, hk \mapsto (h,k)$ This is an isomorphism □

We can compare this to Example 93 (p. 131), where we took G as the group of symmetries of a cube, and G^+ as the group of rotations of a cube. $|G| = 48, |G^+| = 24, G^+ \simeq S_4, G^+ \lhd G, K \lhd G, |K| = 2, G^+ \cap K = \{e\}$
So $G \simeq G^+ \times K \simeq S_4 \times C_2$

Lemma 121. If G has order 4 then $G \simeq C_4$ or $G \simeq C_2 \times C_2$

Proof. If G has an element of order 4, then it is cyclic, so $G \simeq C_4$
Otherwise, all elements of $G \setminus \{e\}$ have order 2, take $a \neq b \in G \setminus \{e\}$ then $G \simeq < a > \times < b > \simeq C_2 \times C_2$ □

Lemma 122. If G has order 8, then either G is abelian and isomorphic to one of $C_8, C_4 \times C_2, (C_2 \times C_2) \times C_2$ or G is non-abelian and isomorphic to one of D_8 or $\underbrace{Q_8}_{\text{quaternion}}$

Proof. First assume G is abelian. If it contains an element of order 8 then $G \simeq C_8$
If all elements have order 2, let $a \neq b \in G \setminus \{e\}$
Then $< a, b > \simeq C_2 \times C_2$ (by preceeding lemma), take $c \in G \setminus < a, b >$, then $G \simeq (< a > \times < b >) \times < c >$
So finally assume G has no elements of order 8, but $\exists a \in G$ of order 4. Let $b \in G \setminus < a >$
Then $b^2 \in < a >$: If $b^2 = e$ then $G \simeq < a > \times < b > \simeq C_4 \times C_2$. Otherwise $b^2 = a^2$ (as $o(b) \neq 8$), replace b by ab, note $(ab)^2 = e$ as G is abelian.

$\therefore \boldsymbol{G} \simeq < \boldsymbol{a} > \times < \boldsymbol{ab} > \simeq \boldsymbol{C_4} \times \boldsymbol{C_2}$

Now assume $\boldsymbol{G}$ is non-abelian. Then there exists $\boldsymbol{a} \in \boldsymbol{G}$ of order 4 (if all elements were of order 2, $\boldsymbol{G}$ would be abelian. Let $\boldsymbol{b} \in \boldsymbol{G} \backslash < \boldsymbol{a} >$. Then $< \boldsymbol{a} > \triangleleft \boldsymbol{G}$ (index 2 in $\boldsymbol{G}$), so $\boldsymbol{bab^{-1}} \in < \boldsymbol{a} >$, say $\boldsymbol{bab^{-1}} = \boldsymbol{a^i}$. Then $\boldsymbol{i} = \pm \boldsymbol{1}$ as $\boldsymbol{bab^{-1}}$ is of order 4.
If $\boldsymbol{bab^{-1}} = \boldsymbol{a}$ then $\boldsymbol{G}$ would be abelian #
So $\boldsymbol{bab^{-1}} = \boldsymbol{a^{-1}}$. Finally, $\boldsymbol{b^2} \in < \boldsymbol{a} >$ as $< \boldsymbol{a} >$ has index 2 in $\boldsymbol{G}$
So $\boldsymbol{b^2} = \boldsymbol{e} \Leftrightarrow \boldsymbol{G} = < \boldsymbol{a}, \boldsymbol{b} : \boldsymbol{a^4} = \boldsymbol{e} = \boldsymbol{b^2}, \boldsymbol{bab^{-1}} = \boldsymbol{a^{-1}} > \simeq \boldsymbol{D_8}$
$\boldsymbol{b^2} = \boldsymbol{a^2} \Leftrightarrow \boldsymbol{G} = < \boldsymbol{a}, \boldsymbol{b} : \boldsymbol{a^4} = \boldsymbol{e}, \boldsymbol{b^2} = \boldsymbol{a^2}, \boldsymbol{bab^{-1}} = \boldsymbol{a^{-1}} > \simeq \boldsymbol{Q_8}$
This final set is equal to $\{\boldsymbol{a^i b^j} : \boldsymbol{0} \leq \boldsymbol{i} < \boldsymbol{4}, \boldsymbol{0} \leq \boldsymbol{j} < \boldsymbol{2}\}$
We have seen $\boldsymbol{Q_8}$ before, with $\boldsymbol{a} = \begin{pmatrix} \boldsymbol{i} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{-i} \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{1} \\ \boldsymbol{-1} & \boldsymbol{0} \end{pmatrix}$ as a subset of $\boldsymbol{SU_2} = \{\boldsymbol{A} \in \boldsymbol{SL_2}(\mathbb{C}) : \boldsymbol{A^{-1}} = \boldsymbol{A^T}\}$, the special unitary group. $\square$