

# Engineering Tripos Part IA - Mathematics

Morărescu Mihnea-Theodor

January 15, 2024

# Contents

<b>1</b>	<b>Vectors</b>	<b>4</b>
1.1	Vector algebra . . . . .	4
1.2	Analytic geometry in $\mathbb{R}^3$ . . . . .	9
1.3	Vector equations . . . . .	14
<b>2</b>	<b>Functions and series approximation</b>	<b>15</b>
2.1	Curve sketching . . . . .	15
2.2	Hyperbolic functions . . . . .	16
2.3	Taylor series expansion . . . . .	18
<b>3</b>	<b>Complex numbers</b>	<b>20</b>
3.1	The Argand plane . . . . .	20
3.2	Euler's formula . . . . .	21
3.3	Trigonometric and hyperbolic functions . . . . .	23
3.4	The complex natural logarithm . . . . .	24
3.5	Roots of a complex number . . . . .	25
3.6	Roots of a polynomial with real coefficients . . . . .	26
<b>4</b>	<b>Ordinary differential equations</b>	<b>27</b>
4.1	First-order ordinary differential equations . . . . .	27
4.1.1	Direct integration . . . . .	27
4.1.2	Separable equations . . . . .	28
4.1.3	Integrating factor method . . . . .	29
4.1.4	Equations reducible to separable form and substitutions . . . . .	30
4.2	Second-order ordinary differential equation . . . . .	33
4.2.1	Constant-coefficients homogeneous second-order ODEs . . . . .	33
4.2.2	Constant-coefficients non-homogeneous second-order ODEs . . . . .	35
4.2.3	The Wronskian . . . . .	36
<b>5</b>	<b>Linear difference equations</b>	<b>37</b>
<b>6</b>	<b>Partial differentiation</b>	<b>39</b>
6.1	Partial derivatives . . . . .	39
6.2	Linearisation . . . . .	41
6.3	Directional derivatives and the chain rule . . . . .	42
6.4	Normal vector to a surface . . . . .	43
<b>7</b>	<b>Matrices and linear transformations</b>	<b>44</b>
7.1	Properties of matrices . . . . .	44
7.2	Mappings and transformations . . . . .	46
7.2.1	Rotation of a vector in two-dimensions . . . . .	46

---

7.3	Rotation about an axis in three dimensions . . . . .	48
7.4	Properties of rotation matrices . . . . .	49
7.5	Reflections . . . . .	50
7.6	Change of basis . . . . .	51
7.6.1	Vectors . . . . .	51
7.6.2	Matrices . . . . .	51
<b>8</b>	<b>Eigenvalues and eigenvectors</b>	<b>53</b>
8.1	The eigenspace of a matrix . . . . .	53
8.2	Properties of symmetric matrices . . . . .	55
8.3	Diagonalizing matrices . . . . .	57
8.4	Generalized eigenvalue problems . . . . .	59

# 1 Vectors

In this section of the course, we will turn our attention to analyzing vectors. We will first begin with elements of vector algebra and the various properties of these objects, and then move onto analytic geometry and equations using vectors.

## 1.1 Vector algebra

**Definition 1.1** (Vector space) A vector space over  $\mathbb{R}$  or  $\mathbb{C}$  is a set  $V \neq \emptyset$  with an internal operation (addition), and an external operation (multiplication of a vector with a scalar).

Vector addition has to satisfy:

- (a)  $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \forall \mathbf{a}, \mathbf{b} \in V$
- (b)  $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in V$
- (c) There exists an identity element  $\mathbf{0} \in V$  so that  $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}, \forall \mathbf{a} \in V$
- (d) For any vector  $\mathbf{a} \in V$ , there exists another vector  $-\mathbf{a} \in V$  so that  $\mathbf{a} + (-\mathbf{a}) = (-\mathbf{a}) + \mathbf{a} = \mathbf{0}$

In other words,  $(V, +)$  must be an abelian group. Furthermore, scalar multiplication has to satisfy:

- (a)  $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$
- (b)  $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$
- (c)  $\lambda(\mu\mathbf{a}) = (\lambda\mu)\mathbf{a}$
- (d)  $1_{\mathbb{F}}\mathbf{a} = \mathbf{a}$

**Definition 1.2** (Vector) For a vector space  $V$ , we say that any element  $\mathbf{v} \in V$  is a vector.

This is the formal definition of a vector. However, for the sake of keeping things simple, a vector is just a quantity that has both magnitude and direction.

**Definition 1.3** (Magnitude of a vector) Consider a vector  $\mathbf{v} \in \mathbb{R}^n$ , with  $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ . We

define the magnitude of a vector as:

$$|\mathbf{v}| = \sqrt{\sum_{i=1}^n v_i^2}$$

**Definition 1.4** (Unit vector) A unit vector  $v$  is a vector with unit length (length equal to 1). For any vector  $\mathbf{v} \in V$ , the unit vector that gives the direction of  $\mathbf{v}$  is denoted by  $\hat{\mathbf{v}}$  and is given by:

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

**Definition 1.5** (Vector addition) Consider two vectors  $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$  and  $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ . Then,

we define the sum of the two vectors as:

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

**Definition 1.6** (Scalar product) Consider two vectors  $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$  and  $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ . We

define the scalar product  $\mathbf{u} \cdot \mathbf{v}$  between the two vectors as:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u} = \sum_{i=1}^n u_i v_i$$

Note that the scalar (dot) product is commutative and is also distributive over addition and subtraction. Therefore:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$$

However, sometimes we might be interested in finding the scalar product between two vectors, yet we only have information about their magnitudes and the angle between them - how would we proceed in this case?

**Proposition 1.1** (Magnitude using scalar product) The magnitude of a vector  $\mathbf{v} \in V$  can be expressed, using the scalar product, as:

$$|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v}$$

**Theorem 1.1** (Scalar product equivalence) Consider two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ . Their scalar product is given by:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$$

Where  $\theta$  is the angle between  $\mathbf{a}$  and  $\mathbf{b}$ .

*Proof.* Consider two points in space with position vectors  $\mathbf{a}$  and  $\mathbf{b}$ . We define  $\mathbf{c} = \mathbf{a} - \mathbf{b}$ . Therefore:

$$|\mathbf{c}|^2 = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = \mathbf{a} \cdot \mathbf{a} - \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2\mathbf{a} \cdot \mathbf{b}$$

However, from the cosine rule, we obtain:

$$|\mathbf{c}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos \theta$$

Hence,  $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$ . □

**Proposition 1.2** (Angle between two vectors) For any two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , we can define the angle between the two vectors as:

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|}$$

**Definition 1.7** (Vector product) Consider two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ . We define the vector product between them as:

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

Using the properties of the determinant, from the definition it follows that:

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$$

Hence, the vector product is anti-commutative. Moreover, by setting  $\mathbf{b} \rightarrow \mathbf{a}$ , we obtain:

$$\mathbf{a} \times \mathbf{a} = -\mathbf{a} \times \mathbf{a} \Leftrightarrow \mathbf{a} \times \mathbf{a} = \mathbf{0}$$

Note that the vector product between  $\mathbf{a}$  and  $\mathbf{b}$  returns a vector  $\mathbf{c}$  that is perpendicular to both  $\mathbf{a}$  and  $\mathbf{b}$ .

**Theorem 1.2** (Magnitude of the vector product) For any two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ , the magnitude of their vector product is given by:

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

Where  $\theta$  is the angle between the two vectors. Also note that the vector product can only be defined in  $\mathbb{R}^3$  (it cannot be generalized for higher dimensions).

**Definition 1.8** (Basis) Let  $V \neq \emptyset$  be a vector space. We say that  $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  is a basis if every element of  $V$  can be written as a linear combination of the elements of  $B$ , i.e.  $\forall \mathbf{v} \in V, \exists \alpha_1, \dots, \alpha_n \in K$  (the field over which we defined the external operation), such that:

$$\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{e}_i$$

**Definition 1.9** (Linear independence) Let  $V \neq \emptyset$  be a vector space. We say that any vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  are linearly independent if  $\forall \lambda_1, \lambda_2, \dots, \lambda_n \in K$ , then:

$$\sum_{i=1}^n \lambda_i \mathbf{v}_i = \mathbf{0} \iff \lambda_i = 0, \forall 1 \leq i \leq n$$

**Theorem 1.3** (Generating bases in  $\mathbb{R}^3$ ) Consider two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$  such that  $\mathbf{a} \times \mathbf{b} \neq \mathbf{0}$ . Then, the set  $B = \{\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}\}$  forms a basis of  $\mathbb{R}^3$ . (Note that this is true provided the two vectors are not parallel).

Therefore,  $\forall \mathbf{x} \in \mathbb{R}^3, \exists \alpha, \beta, \gamma \in \mathbb{R}$  so that:

$$\mathbf{x} = \alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{a} \times \mathbf{b}$$

**Definition 1.10** (Scalar triple product) Consider three arbitrary vectors  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ . We define the scalar triple product between the vectors as:

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

The geometrical interpretation of the scalar triple product is that it is equal to the volume of the parallelepiped formed by the three vectors.

**Proposition 1.3** (Properties of the scalar triple product) Consider three vectors  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ . Then:

- (a)  $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$  (Cyclic property)
- (b)  $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$  (We can replace the dot with the cross)
- (c) If all vectors are co-planar, then  $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = 0$
- (d) If  $[\mathbf{a}, \mathbf{b}, \mathbf{c}] > 0$ , the three vectors form a right-handed set. Otherwise, they form a left-handed set.

$$(e) \ [\mathbf{a}, \mathbf{b}, \mathbf{c}] = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}$$

$$(f) \ \forall \lambda \in \mathbb{R}, [\lambda \mathbf{a}, \mathbf{b}, \mathbf{c}] = \lambda [\mathbf{a}, \mathbf{b}, \mathbf{c}]$$

**Definition 1.11** (Vector triple product) Consider three vectors  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ . We define the vector triple product between these vectors as:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$$

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}$$



## 1.2 Analytic geometry in $\mathbb{R}^3$

**Proposition 1.4** (Equation of line) Consider a line which contains a point with position vector  $\mathbf{a}$  and direction vector given by  $\mathbf{b}$ . The equation of the line is then given by:

$$\mathbf{r} = \mathbf{a} + \lambda \mathbf{b}, \lambda \in \mathbb{R}$$

The proof is trivial.

**Proposition 1.5** (Cartesian equivalent of line equation I) Consider a line given by  $\mathbf{r} = \mathbf{a} + \lambda \mathbf{b}$ . Then, the equation of this line in Cartesian form is given by:

$$\frac{x - a_x}{b_x} = \frac{y - a_y}{b_y} = \frac{z - a_z}{b_z}$$

*Proof.* We can rewrite the equation of line using matrix form as:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} + \lambda \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}$$

Therefore,  $x = a_x + \lambda b_x$ ,  $y = a_y + \lambda b_y$  and  $z = a_z + \lambda b_z$ . Note that this is also known as the parametric form of the line equation, where every component is expressed in terms of a parameter  $\lambda$ . By expressing  $\lambda$  for each one of these, we obtain the conclusion.  $\square$

Note that in order to obtain different points on a line, we need to vary  $\lambda$ .

**Proposition 1.6** (Equation of line between two points) Consider two points with position vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ . Therefore, the line between these two points is given by:

$$\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})$$

The conclusion follows because we can take  $\mathbf{b} - \mathbf{a}$  to be the direction vector of the line.

**Proposition 1.7** (Cartesian equivalent of line equation II) Consider a line given by  $\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})$ . Then, the equation of this line in Cartesian form is given by:

$$\frac{x - a_x}{b_x - a_x} = \frac{y - a_y}{b_y - a_y} = \frac{z - a_z}{b_z - a_z}$$

**Proposition 1.8** (Second form of an equation of a line) Consider a line given by  $\mathbf{r} = \mathbf{a} + \lambda \mathbf{b}$ . By considering  $\mathbf{b} \times \mathbf{r}$ , we obtain:

$$\mathbf{b} \times \mathbf{r} = \mathbf{b} \times \mathbf{a}$$

By defining a new vector  $\mathbf{c} = \mathbf{b} \times \mathbf{a}$ , we obtain the second form of a line equation:

$$\mathbf{b} \times \mathbf{r} = \mathbf{c}$$

**Proposition 1.9** (Intersection between two lines) Consider two lines given by  $\mathbf{r}_1 = \mathbf{a}_1 + \lambda_1 \mathbf{b}_1$  and  $\mathbf{r}_2 = \mathbf{a}_2 + \lambda_2 \mathbf{b}_2$ . We want to determine the intersection point between these two lines. Therefore, we begin by parametrizing the line as:  $x_1(\lambda_1) = a_{1x} + \lambda_1 b_{1x}$ ,  $x_2(\lambda_2) = a_{2x} + \lambda_2 b_{2x}$ , and  $y_1(\lambda_1) = a_{1y} + \lambda_1 b_{1y}$ ,  $y_2(\lambda_2) = a_{2y} + \lambda_2 b_{2y}$ . We then determine  $\lambda_1, \lambda_2$  so that  $x_1 = x_2$  and  $y_1 = y_2$ , and afterwards we might perform a sanity check to make sure that  $z_1 = z_2$ .

**Proposition 1.10** (Distance between point and line) Consider a line given by  $\mathbf{r} = \mathbf{a} + \lambda \mathbf{b}$  and a point with position vector  $\mathbf{p}$ . The distance between the point  $P$  and the line is given by:

$$\delta = \frac{|(\mathbf{p} - \mathbf{a}) \times \mathbf{b}|}{|\mathbf{b}|}$$

*Proof.* Consider the angle given between the line  $AP$  and the direction vector  $\mathbf{b}$ . We observe that  $\sin \theta = \frac{\delta}{|\mathbf{p} - \mathbf{a}|}$ . Therefore,  $\delta = |\mathbf{p} - \mathbf{a}| \sin \theta$ . We can then express the sine term as:

$$\delta = \frac{|(\mathbf{p} - \mathbf{a}) \times \mathbf{b}|}{|\mathbf{b}|}$$

Thus concluding the proof. □

Note that if we have to find the distance between two parallel lines, the problem is equivalent to finding the distance between any point on one line to the second line.

**Proposition 1.11** (Distance between two skew lines) Consider two skew lines given by equations  $\mathbf{r}_1 = \mathbf{a}_1 + \lambda \mathbf{b}_1$  and  $\mathbf{r}_2 = \mathbf{a}_2 + \mu \mathbf{b}_2$ . Therefore, the distance between the two lines is given by:

$$\delta = \frac{|(\mathbf{a}_2 - \mathbf{a}_1) \cdot (\mathbf{b}_1 \times \mathbf{b}_2)|}{|\mathbf{b}_1 \times \mathbf{b}_2|}$$

*Proof.* We first begin by determining a vector that is perpendicular to both lines. This is given by:

$$\mathbf{n} = \mathbf{b}_1 \times \mathbf{b}_2$$

We do not want to alter, however, the length of our distance. For this reason, we normalize this vector, and we obtain:

$$\hat{\mathbf{n}} = \frac{\mathbf{b}_1 \times \mathbf{b}_2}{|\mathbf{b}_1 \times \mathbf{b}_2|}$$


---

Hence, the distance between the two lines is given by the component in the direction of  $\hat{\mathbf{n}}$  of  $\mathbf{r}_2 - \mathbf{r}_1$ . We are then free to pick any two points on the line, hence:

$$\delta = |(\mathbf{r}_2 - \mathbf{r}_1) \cdot \hat{\mathbf{n}}| = \frac{|(\mathbf{a}_2 - \mathbf{a}_1) \cdot (\mathbf{b}_1 \times \mathbf{b}_2)|}{|\mathbf{b}_1 \times \mathbf{b}_2|}$$

□

**Proposition 1.12** (Equation of plane) Consider a plane  $\Pi$ , a point  $A \in \Pi$  with position vector  $\mathbf{a}$ , and a vector  $\mathbf{n}$  that is normal to the plane. Therefore, the equation of the plane is:

$$(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$$

The geometrical meaning is extremely intuitive: the equation above just tells us that the plane represents the set of points in space for which the line  $AP$  is perpendicular to  $\mathbf{n}$ .

This equation can be rearranged into:

$$\mathbf{r} \cdot \mathbf{n} - \mathbf{a} \cdot \mathbf{n} = 0 \Leftrightarrow \mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}$$

By denoting  $d = \mathbf{a} \cdot \mathbf{n}$  (note that this is the distance from the origin to the plane), we obtain a simplified version of the plane equation:

$$\mathbf{r} \cdot \mathbf{n} = d$$

**Proposition 1.13** (Cartesian equation of plane) Consider a plane given by  $\mathbf{r} \cdot \mathbf{n} = d$ . The equation in Cartesian coordinates is given by:

$$ax + by + cz = d$$

Where  $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ . By setting  $d \rightarrow -d$ , we can obtain an equivalent expression:

$$ax + by + cz + d = 0$$

**Proposition 1.14** (Equation of plane between three points) Consider three points  $A, B, C$  in space with position vector  $\mathbf{a}, \mathbf{b}, \mathbf{c}$ . We can obtain the expression of the plane determined by these three points by either considering the normal vector  $\mathbf{n} = (\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a})$  and then we obtain  $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$ , or by considering:

$$\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) + \mu(\mathbf{c} - \mathbf{a})$$

This is equivalent to:

$$\mathbf{r} = (1 - \lambda - \mu)\mathbf{a} + \lambda\mathbf{b} + \mu\mathbf{c}$$

By setting  $\alpha = 1 - \lambda - \mu$ ,  $\beta = \lambda$  and  $\gamma = \mu$ , we obtain the equivalent plane equation:

$$\mathbf{r} = \alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c}, \alpha + \beta + \gamma = 1$$

**Proposition 1.15** (Distance between point and plane) Consider a point  $P(x_0, y_0, z_0) \in \mathbb{R}^3$  and a plane given by  $\mathbf{r} \cdot \mathbf{n} = d$  and normal vector given by  $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ , the distance between  $P$  and the plane is:

$$\delta = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$$

Note that if we have to calculate the distance between a line and a plane, we distinguish between two cases:

- (a) If the line and the plane are not parallel, then they intersect, and therefore the distance is null
- (b) If the line and the plane are parallel, the problem is equivalent to obtaining the coordinates of a point on the line and calculating the distance from said point to the plane

**Proposition 1.16** (Intersection between a line and a plane) Consider a line given by  $\mathbf{r} = \mathbf{a} + \lambda \mathbf{b}$  and a plane given by  $\mathbf{r} \cdot \mathbf{n} = d$ . In order to find the intersection between a line and a plane, we need to proceed as follows:

- (a) Parametrize the line by expressing  $x = x(\lambda)$ ,  $y = y(\lambda)$  and  $z = z(\lambda)$
- (b) Express the plane in Cartesian form as  $ax + by + cz + d = 0$
- (c) By using the parametrization above, use the equation of the plane and determine  $\lambda$
- (d) We distinguish between three cases: (i) if we have a unique solution, then we have a point of intersection, (ii) if we obtain a truism (i.e.  $0 = 0$ ), the line is contained within the plane, (iii) if we obtain a false expression, then the line and the plane do not intersect

**Proposition 1.17** (Intersection between three planes) Consider two planes given by the equations  $\mathbf{r}_1 \cdot \mathbf{n}_1 = d_1$  and  $\mathbf{r}_2 \cdot \mathbf{n}_2 = d_2$ . The general procedure for determining the intersection between the planes is:

- (a) Derive the Cartesian equations for both planes, i.e.  $a_1x + b_1y + c_1z + d_1 = 0$  and  $a_2x + b_2y + c_2z + d_2 = 0$

- (b) Solve the given system of equations (which has either no solution or infinitely many solutions)
- (c) If the system has no solutions, the planes are parallel. Otherwise, the intersection of the two planes is a line

### 1.3 Vector equations

The general strategy when solving vector equations is to either take the scalar or vector product of the equation with suitable vectors in order to further reduce it. Otherwise, if we cannot perform this simplification or after the equation has been sufficiently simplified, we can write component-wise equations.

**Example 1.1** Simplify the following vector equation:  $\mathbf{x} - (\mathbf{x} \times \mathbf{a}) \times \mathbf{b} = \mathbf{c}$ .

The equation can be expanded using the triple vector product as:

$$\mathbf{x} - (\mathbf{x} \cdot \mathbf{b})\mathbf{a} + (\mathbf{a} \cdot \mathbf{b})\mathbf{x} = \mathbf{c}$$

We can dot the above equation with  $\mathbf{b}$  to obtain:

$$\mathbf{x} \cdot \mathbf{b} - (\mathbf{x} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{b}) + (\mathbf{a} \cdot \mathbf{b})(\mathbf{x} \cdot \mathbf{b}) = \mathbf{c} \cdot \mathbf{b}$$

Which can be further simplified to:

$$\mathbf{x} \cdot \mathbf{b} = \mathbf{c} \cdot \mathbf{b}$$

## 2 Functions and series approximation

In this section we are concerned with curve sketching, performing revision of functions and extending some concepts.

### 2.1 Curve sketching

**Definition 2.1** (Even function) We say that a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is even if  $\forall x \in \mathbb{R}, f(-x) = f(x)$ .

**Definition 2.2** (Odd function) We say that a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is odd if  $\forall x \in \mathbb{R}, f(-x) = -f(x)$ .

**Theorem 2.1** Any function  $f : \mathbb{R} \rightarrow \mathbb{R}$  can be written as the sum of an even and an odd function.

*Proof.* Choose  $g, h : \mathbb{R} \rightarrow \mathbb{R}$ , with  $g(x) = \frac{f(x)+f(-x)}{2}$  and  $h(x) = \frac{f(x)-f(-x)}{2}$ . Then,  $g(x)$  is an even function, while  $h(x)$  is odd, and then:

$$f(x) = g(x) + h(x), \forall x \in \mathbb{R}$$

□

**Theorem 2.2** Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a differentiable function. If  $f$  is odd, then  $f(0) = 0$ . If  $f$  is even, then  $f'(0) = 0$ .

We will now consider a general algorithm to follow whenever sketching curves.

- (a) Solve the equation  $f(x) = 0$  and find the intersection points with the  $Ox$  axis
- (b) Find the intersection with the  $Oy$  axis, i.e.  $f(0)$
- (c) Take the limits at the extreme points of the domain, i.e.  $\lim_{x \rightarrow \pm\infty} f(x)$  or otherwise
- (d) Determine all the points of discontinuity (singularities)
- (e) Determine  $f'(x)$  and set  $f'(x) = 0$  and determine the monotonicity intervals. Note that  $f'(x) \geq 0$  implies function is increasing, and  $f'(x) \leq 0$  implies function is decreasing. Note that for a point  $x_0 \in \mathbb{R}$ , if  $f'(x_0) \geq 0, \forall x \geq x_0$  and  $f'(x_0) \leq 0, \forall x \leq x_0$ , then  $x_0$  is a minima. If the inverse is true, then it is a point of maxima.
- (f) Determine  $f''(x)$  and determine the function's convexity intervals
- (g) Sketch the curve

## 2.2 Hyperbolic functions

The hyperbolic functions are functions expressed in terms of exponentials, that are closely related to the trigonometric functions. We will further explore this relationship in the next chapter, where we will thoroughly analyze complex numbers.

**Definition 2.3** (Hyperbolic functions) We define the hyperbolic sine and cosine as  $\sinh, \cosh : \mathbb{R} \rightarrow \mathbb{R}$ , with:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \cosh x = \frac{e^x + e^{-x}}{2}$$

From the above definition, we can easily deduce the hyperbolic tangent (and respectively the hyperbolic cotangent):

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Furthermore, by differentiating the hyperbolic functions, we obtain that  $\sinh'(x) = \cosh x$  and  $\cosh'(x) = \sinh x$ . These appear similar to the normal trigonometric functions, however we must note that the derivative of the hyperbolic cosine does not contain a minus sign. Furthermore,  $\tanh'(x) = \frac{1}{\cosh^2 x}$ .

Below, we will list some important identities of the hyperbolic functions:

- (a)  $\sinh x + \cosh x = e^x, \forall x \in \mathbb{R}$
- (b)  $\cosh x - \sinh x = e^{-x}, \forall x \in \mathbb{R}$
- (c)  $\cosh^2 x + \sinh^2 x = 1, \forall x \in \mathbb{R}$

We will prove the usual trigonometric identities in the complex numbers section of this course.

**Example 2.1** (Inverse hyperbolic sine) Compute  $\sinh^{-1} x$ . Let  $y = \sinh x$ . Therefore,  $y = \frac{e^x - e^{-x}}{2}$ , and by expanding this relationship, we obtain:

$$e^x - 2y - e^{-x} = 0$$

Multiplying the above by  $e^x$  (note that  $e^x > 0, \forall x \in \mathbb{R}$ ), we obtain:

$$e^{2x} - 2ye^x - 1 = 0$$

This is a quadratic equation in  $x$ , and we can simply proceed as per usual to obtain an expression for  $e^x$ . The discriminant is then  $\Delta = 4y^2 + 4$ , and therefore:

$$e^x = y \pm \sqrt{y^2 + 1}$$



Because  $e^x > 0, \forall x \in \mathbb{R}$  and because  $\sqrt{x^2 + 1} > x, \forall x \in \mathbb{R}$ , we can only allow for the positive solution. Therefore:

$$x = \ln(y + \sqrt{y^2 + 1})$$

And therefore we obtain:

$$\sinh^{-1} x = \ln(x + \sqrt{x^2 + 1})$$

We can proceed in a similar fashion to obtain expressions for the inverse hyperbolic cosine and tangent.

## 2.3 Taylor series expansion

**Theorem 2.3** (Taylor series) Consider an indefinitely differentiable function  $f : \mathbb{R} \rightarrow \mathbb{R}$ . The Taylor expansion of  $f(x)$  about a point  $x_0 \in \mathbb{R}$  is given by:

$$f(x) = f(x_0) + \sum_{n=1}^{\infty} \frac{f^{(n)}(x_0)(x - x_0)^n}{n!}$$

*Proof.* The Taylor series expansion can be proved by considering a polynomial expansion of  $f(x)$  of the form:

$$f(x) = \sum_{n=0}^{\infty} \alpha_n (x - x_0)^n$$

Setting  $x \rightarrow x_0$ , we obtain that  $\alpha_0 = f(x_0)$ . By differentiating once and setting  $x \rightarrow x_0$  again, we obtain  $\alpha_1 = f'(x_0)$ . In general, by differentiating  $n$  times and setting  $x \rightarrow x_0$ , we obtain:

$$\alpha_n = \frac{f^{(n)}(x_0)}{n!}$$

□

**Definition 2.4** (Maclaurin series) The Maclaurin series of a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is given by the Taylor series expansion of the function about  $x_0 = 0$ , i.e.:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

Sometimes, it is useful to expand the Taylor polynomial up until a certain power  $k \in \mathbb{N}$ , and ignore the other terms. This is called the big  $O$  notation, and we usually denote the remainder as:

$$f(x) = \sum_{n=0}^k \frac{f^{(n)}(x_0)(x - x_0)^n}{n!} + \mathcal{O}(x^{k+1})$$

This effectively means that  $\mathcal{O}(x^{k+1}) \rightarrow 0$ , or that all the terms above the power  $k$  can be neglected. Also note that:

- (a)  $\mathcal{O}(x^m) + \mathcal{O}(x^n) = \mathcal{O}(x^d)$ , where  $d = \min(m, n)$
- (b)  $\mathcal{O}(x^m)\mathcal{O}(x^n) = \mathcal{O}(x^{m+n})$

**Example 2.2** (Maclaurin series for usual functions) Below, we list a few usually encountered functions and their Maclaurin expansion:

- (a)  $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$
- (b)  $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$
- (c)  $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$
- (d)  $(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \dots$
- (e)  $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots$

**Proposition 2.1** (Gamma factorial) The factorial function can be extended to the real and complex numbers by the function:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

Note that  $\Gamma(n) = (n-1)!, \forall n \in \mathbb{N}$ .

**Proposition 2.2** (Stirling's approximation) For any  $n \in \mathbb{N}$ , the value of  $n!$  can be approximated by:

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2n\pi}$$

**Theorem 2.4** (L'Hôpital's rule) Let  $I \subset \mathbb{R}$  and  $f, g : I \rightarrow \mathbb{R}$  be two differentiable functions on  $I \setminus \{x_0\}$ , with  $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$  or  $\pm\infty$  and  $g'(x) \neq 0, \forall x \in I \setminus \{x_0\}$ . Then:

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$$

### 3 Complex numbers

In this section of the course, we are concerned with studying complex numbers and their various applications within multiple areas of Engineering.

#### 3.1 The Argand plane

**Definition 3.1** (Complex number) A complex number  $z \in \mathbb{C}$  can be represented as  $z = a + bi$ , where  $a, b \in \mathbb{R}$  and  $i$  is defined as  $i^2 = -1$ .

**Definition 3.2** (Modulus of a complex number) The modulus of a complex number  $z \in \mathbb{C}, z = a + bi, a, b \in \mathbb{R}$  is defined as:

$$|z| = \sqrt{a^2 + b^2}$$

**Definition 3.3** (Complex conjugate) Let  $z \in \mathbb{C}$ . The complex conjugate  $\bar{z}$  is defined as  $\bar{z} = a - bi$ , where  $z = a + bi$ .

Note the useful property that  $z\bar{z} = |z|^2$ .

**Proposition 3.1** (Trigonometric form of a complex number) Any complex number  $z \in \mathbb{C}$  can be represented in trigonometric form as:

$$z = r(\cos \theta + i \sin \theta)$$

Where  $r = |z| = \sqrt{a^2 + b^2}$  and  $\tan \theta = \frac{b}{a}$ . Note that  $\theta$  is also called the argument of  $z$ , and is denoted by  $\arg z = \theta$ .

Curves in the complex plane can be defined by the locus of points that satisfy a certain relationship. To determine the loci, we always let  $z = x + iy$  and rearrange the given equation to determine its Cartesian equation.

**Example 3.1** Determine the loci  $|z + 2| = 2|z - 1|$ .

Let  $z \in \mathbb{C}$ , with  $z = x + iy$ . Then, the equation becomes:

$$\begin{aligned} |z + 2|^2 &= 4|z - 1|^2 \Leftrightarrow \\ |(x + 2) + iy|^2 &= 4|(x - 1) + iy|^2 \Leftrightarrow \\ (x + 2)^2 + y^2 &= 4(x - 1)^2 + 4y^2 \end{aligned}$$

By rearrangement, we obtain:

$$(x - 2)^2 + y^2 = 4$$

Therefore, this is the equation of a circle with centre at  $C(2, 0)$  and radius 2.

## 3.2 Euler's formula

In this section, we are going to introduce one of the most important identities related to complex numbers, which we will further use to provide more insight into how they work and prove other theorems related to them.

**Theorem 3.1** (Euler's formula) Let  $\theta \in \mathbb{R}$ . Then:

$$e^{i\theta} = \cos \theta + i \sin \theta$$

*Proof.* We know that the Taylor series expansion of  $e^x$  is given by:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

We can now proceed forward and set  $x \rightarrow ix$ . This means that:

$$e^{ix} = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!}$$

Upon further inspection, we realize that the real part of this series is the Taylor series expansion of  $\cos x$ , and the imaginary part is just  $\sin x$ . Hence:

$$e^{ix} = \cos x + i \sin x$$

□

Since any complex number  $z$  can be written as  $z = r(\cos \theta + i \sin \theta)$ , by using the theorem above, we obtain:

$$z = re^{i\theta}$$

Where  $r = |z|$  and  $\theta = \arg z$ .

**Proposition 3.2** (De Moivre's theorem) Let  $z = r(\cos \theta + i \sin \theta)$ . Then:

$$z^\alpha = r^\alpha (\cos \alpha\theta + i \sin \alpha\theta)$$

This result can be proved using Euler's theorem for  $\theta \rightarrow \alpha\theta$  (noting that the modulus of the complex number also changes).

Furthermore, we can use Euler's formula to obtain trigonometric identities.

**Example 3.2** We wish to evaluate  $\sin(a+b)$  and  $\cos(a+b)$ . Consider the following expression:

$$e^{ia}e^{ib} = e^{i(a+b)} \Leftrightarrow$$
$$(\cos a + i \sin a)(\cos b + i \sin b) = \cos(a+b) + i \sin(a+b)$$

By expanding this expression and equating real and imaginary parts respectively, we obtain that  $\sin(a+b) = \sin a \cos b + \sin b \cos a$  and  $\cos(a+b) = \cos a \cos b - \sin a \sin b$ .

### 3.3 Trigonometric and hyperbolic functions

By coming back to Euler's formula, we can write that:

$$e^{ix} = \cos x + i \sin x$$

$$e^{-ix} = \cos x - i \sin x$$

By adding and subtracting these two relationships, we obtain:

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} = \cosh ix$$

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} = -i \sinh ix$$

Furthermore, the substitution  $x \rightarrow ix$  allows us to express the hyperbolic functions as complex trigonometric functions and vice-versa:

$$\cos(ix) = \cosh x$$

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

Using these relationships, we can now express the sine and cosine of complex numbers, allowing us to extend trigonometric functions towards the complex numbers. Therefore:

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

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

**Example 3.3** Determine  $\cosh^{-1}(0)$ .

Let  $\cosh^{-1}(0) = x + iy \Leftrightarrow \cosh(x + iy) = 0$ . This is equivalent to  $\cosh x \cos y + i \sinh x \sin y = 0$ . Note that  $\cosh x > 0$ , and hence  $\cos y = 0$ , so then  $y = \pm \frac{\pi}{2} + 2n\pi$ . Furthermore,  $\sinh x = 0$ , and therefore  $x = 0$ . Therefore:

$$\cosh^{-1}(0) = i(2n\pi \pm \frac{\pi}{2})$$

### 3.4 The complex natural logarithm

Consider a complex number  $z \in \mathbb{C}$ . Therefore, we can write  $z$  as  $z = re^{i\theta}$ . Hence, we can define the complex natural logarithm as:

$$\ln z = \ln |z| + i \arg z$$

**Example 3.4** Determine the value of  $\ln(1+i)$ .

We determine the modulus of the number and its argument. Hence,  $|1+i| = \sqrt{2}$  and  $\tan \theta = 1 \Leftrightarrow \theta = \frac{\pi}{4} + n\pi$ . Hence:

$$\ln(1+i) = \frac{1}{2} \ln 2 + i\left(\frac{\pi}{4} + n\pi\right)$$

If we want to find the principal value of the logarithm, we can just set  $n = 0$ .



### 3.5 Roots of a complex number

Suppose we have to solve an equation of the type  $z^n = w$ , where  $w \in \mathbb{C}$  is fixed. This is equivalent to:

$$z^n = re^{i\theta}$$

We observe that we need  $n$  solutions to this equation. Hence, by Euler's formula:

$$z = r^{\frac{1}{n}} e^{\frac{i\theta + 2k\pi}{n}}$$

Where  $k \in \{0, 1, \dots, n-1\}$ .

### 3.6 Roots of a polynomial with real coefficients

**Theorem 3.2** Consider a polynomial  $f \in \mathbb{R}[X]$  and  $z \in \mathbb{C}$  so that  $f(z) = 0$ . Then:

$$f(\bar{z}) = 0$$

*Proof.* Consider  $f \in \mathbb{R}[X]$  and  $z \in \mathbb{C}$  so that  $f(z) = 0$ . We can express  $f$  as:

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \dots a_1 z + a_0 = 0$$

Now, we can conjugate this expression:

$$\overline{f(z)} = \overline{a_n z^n + a_{n-1} z^{n-1} + \dots a_1 z + a_0} = 0$$

We can split the conjugation operation term by term likewise:

$$\overline{f(z)} = a_n \overline{z^n} + a_{n-1} \overline{z^{n-1}} + \dots a_1 \bar{z} + a_0 = 0$$

This is true since  $a_k \in \mathbb{R}, \forall 0 \leq k \leq n$ . Now, since  $\overline{z^k} = \bar{z}^k$ , we obtain:

$$f(\bar{z}) = 0$$

Thus concluding our proof. □

## 4 Ordinary differential equations

Most physical systems are governed by equations involving a function and its derivatives. These are called ordinary differential equations, and this chapter is dedicated to studying first and second order ODEs.

**Definition 4.1** (Ordinary differential equation) We say that an equation is an ordinary differential equation if it is of the form:

$$f(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots, \frac{d^ny}{dx^n}) = 0$$

**Definition 4.2** (Order) The order of an ordinary differential equation is the highest derivative present in the equation.

**Definition 4.3** (Linearity) An ordinary differential equation is linear if the maximum power of any of its derivatives is 1. Otherwise, we say that the differential equation is non-linear.

### 4.1 First-order ordinary differential equations

**Definition 4.4** (First-order ODE) We say that an ordinary differential equation is of first order if it is an equation of the type:

$$f(x, y, \frac{dy}{dx}) = 0$$

In the section below, we will consider different types of first-order ODEs and explain how to fully determine their solutions.

#### 4.1.1 Direct integration

We can apply direct integration if the differential equation is of the type:

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

We can proceed by integrating the above to obtain  $y = y(x)$  :

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

## 4.1.2 Separable equations

A separable equation is a differential equation of the type:

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

To solve an equation of this type, we just separate both differentials likewise:

$$f(y)dy = g(x)dx \Leftrightarrow \int f(y)dy = \int g(x)dx$$

**Example 4.1** Find the variation of velocity with time of a parachutist subject to a drag force  $F = kv^2$ .

We begin by applying Newton's second law:

$$ma = mg - kv^2 \Leftrightarrow m\frac{dv}{dt} = mg - kv^2 \Leftrightarrow \frac{dv}{dt} = g - \frac{k}{m}v^2$$

This can be rearranged as:

$$\frac{m}{k}\frac{dv}{dt} = \frac{mg}{k} - v^2$$

Let  $\alpha = \frac{mg}{k}$ . Therefore:

$$\frac{m}{k}\frac{dv}{dt} = \alpha^2 - v^2$$

This can be separated as:

$$\frac{m}{k}\frac{dv}{\alpha^2 - v^2} = dt$$

We can then integrate this to obtain:

$$\frac{m}{k} \int \frac{1}{\alpha^2 - v^2} dv = t + C$$

Hence:

$$\frac{m}{2k\alpha} \ln \left( \frac{\alpha + v}{\alpha - v} \right) = t + C$$

If we take  $v(0) = 0$  and by manipulating the above expression, we obtain our solution in the form:

$$v(t) = \alpha \frac{e^{2\alpha kt/m} - 1}{e^{2\alpha kt/m} + 1}$$

### 4.1.3 Integrating factor method

Consider a first-order differential equation of the type:

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

This is the most general type of a first-order linear ordinary differential equation. We want to determine a function  $g(x)$  so that when we multiply it by our equation, we can express the left side as the derivative of a product. Therefore, assuming  $g(x) \neq 0$ , we obtain:

$$\frac{dy}{dx}g(x) + p(x)g(x)y = q(x)g(x)$$

We impose the condition that:

$$\frac{d}{dx}(g(x)y) = g(x)\frac{dy}{dx} + g'(x)y = g(x)\frac{dy}{dx} + g(x)p(x)y$$

This yields:

$$g(x)p(x)y = g'(x)y$$

By our assumption that  $g(x) \neq 0$ , we obtain:

$$\frac{g'(x)}{g(x)} = p(x)$$

Integrating this yields us:

$$\int \frac{g'(x)}{g(x)} dx = \int p(x) dx$$

Therefore:

$$\ln g(x) = \int p(x) dx$$

And then, we deduce that the function we are looking for is:

$$g(x) = e^{\int p(x) dx}$$

Therefore, in order to solve equations of the type above, we must multiply by the integrating factor:

$$I = e^{\int p(x) dx}$$

**Example 4.2** Solve the following differential equation:  $\frac{dy}{dx} - 4x^3y = 4x^3$  with boundary conditions  $y(0) = 0$ .

We observe that  $p(x) = -4x^3$ , therefore:

$$I = e^{\int -4x^3 dx} = e^{-x^4}$$

Multiplying the above equation by  $I$ , we obtain:

$$e^{-x^4} \frac{dy}{dx} - 4x^3 e^{-x^4} y = 4x^3 e^{-x^4} \Leftrightarrow$$

$$\frac{d}{dx}(e^{-x^4} y) = 4x^3 e^{-x^4}$$

By integrating the relationship above, we obtain:

$$e^{-x^4} y = -e^{-x^4} + c$$

Therefore:

$$y(x) = Ce^{x^4} - 1$$

Since  $y(0) = 0$ , we deduce that  $C = 1$ , and then our solution is therefore:

$$y(x) = e^{x^4} - 1$$

## 4.1.4 Equations reducible to separable form and substitutions

Oftentimes, it is extremely useful to consider various substitutions when solving differential equations that do not fall under any of the categories above. These are dependent entirely on the equation in place, however, we might consider substitutions of the type  $u = \frac{1}{x}$ ,  $u = \frac{dy}{dx}$  and so on. These are better illustrated via examples.

**Example 4.3** (Bernoulli's differential equation) Consider the differential equation:

$$\frac{dy}{dx} + p(x)y = q(x)y^\alpha, \alpha \in \mathbb{R} \setminus \{0, 1\}$$

Substitute  $u = y^{1-\alpha}$ , and divide the equation above by  $y^\alpha$  to obtain:

$$\frac{dy}{dx} y^{-\alpha} + p(x)y^{1-\alpha} = q(x)$$

Now, we can write  $\frac{du}{dx} = (1 - \alpha)y^{-\alpha} \frac{dy}{dx}$ , and therefore:

$$\frac{dy}{dx} = \frac{du}{dx} \frac{y^\alpha}{1 - \alpha}$$

And by substituting everything into our original equation, we obtain:

$$\frac{1}{1 - \alpha} \frac{du}{dx} + p(x)u = q(x)$$

And then this becomes:

$$\frac{du}{dx} + (1 - \alpha)p(x)u = q(x)$$

And the initial equation simply degenerated to an integrating factor method. We then solve for  $u(x)$  and obtain  $y(x)$ .

**Example 4.4** (Substitutions of the type  $u = \frac{y}{x}$ ) Consider the following differential equation, subject to  $y(1) = 0$ , and  $x > 1$ :

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

We can forcefully divide the equation by  $x^2$ :

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

Substitute  $u = \frac{y}{x} \Leftrightarrow y = ux \Leftrightarrow \frac{dy}{dx} = x\frac{du}{dx} + u$ . Therefore:

$$\begin{aligned} x\frac{du}{dx} + u &= 1 - u + u^2 \Leftrightarrow \\ x\frac{du}{dx} &= (u - 1)^2 \end{aligned}$$

This equation can then be separated as:

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

Then, by integrating both sides:

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

By utilizing the boundary condition and coming back to  $y(x)$ , we deduce that the solution is:

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

**Example 4.5** (Differential substitutions) Determine the general solution to the following differential equation:

$$\frac{d^2y}{dx^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$

Initially, this equation does not appear to be of first order, and also does not appear to be linear. However, we can substitute  $u = \frac{dy}{dx}$  to obtain:

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

Now, we have reduced the equation to a simple, linear, first-order ODE. This equation is separable, and then we can obtain:

$$\begin{aligned}\frac{1}{\sqrt{1 + u^2}} du &= dx \Leftrightarrow \\ \sinh^{-1} u &= x + C \Leftrightarrow \\ u &= \sinh(x + C)\end{aligned}$$

Therefore:

$$\begin{aligned}\frac{dy}{dx} &= \sinh(x + C) \Leftrightarrow \\ dy &= \sinh(x + C) dx\end{aligned}$$

And lastly, we obtain:

$$y(x) = \cosh(x + C) + D$$



## 4.2 Second-order ordinary differential equation

**Definition 4.5** (Second-order ODE) We say that an ordinary differential equation is of second order if it is an equation of the type:

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

The most general form of a second-order linear ODE is:

$$\frac{d^2y}{dx^2} + p(x)\frac{dy}{dx} + q(x)y = r(x)$$

### 4.2.1 Constant-coefficients homogeneous second-order ODEs

**Definition 4.6** (Homogeneity) We say that an ODE is homogeneous if the constant term is equal to zero.

Consider the following equation:

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

This equation is homogeneous, as the constant term is null. In order to solve it, however, we must consider the characteristic equation of the differential equation:

$$a\lambda^2 + b\lambda + c = 0$$

We distinguish between the following three cases:

- (a)  $b^2 - 4ac > 0$ , then we have two separate solutions  $\lambda_1, \lambda_2$
- (b)  $b^2 - 4ac = 0$ , then we have one double solution  $\lambda$
- (c)  $b^2 - 4ac < 0$ , then we have complex conjugate solutions

We will tackle each of these cases individually, as problems might arise when attempting to solve them.

**Proposition 4.1** (Real distinct solution) Consider the case when  $b^2 - 4ac > 0$ . Therefore, by solving the equation, we obtain the solutions:

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Because exponential functions form the basis for all solutions, they are the eigenvectors of the the ODE system. Hence, we might assume a general solution of the type:

$$y(x) = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$$

Lastly, we can determine the constants  $A, B \in \mathbb{R}$  by utilizing the given boundary conditions.

**Proposition 4.2** (Repeated roots) By proceeding as in the previous case, we would get two equal functions  $e^{\lambda x}$ . However, we would need two separate vectors (functions) to span the entire vector spaces of functions. Hence, we look for another solution of the type:

$$y = f(x)e^{\lambda x}$$

By proceeding to substitute this solution in the differential equation and making use of the fact that  $b^2 - 4ac = 0$ , we are left with:

$$f''(x) = 0 \Leftrightarrow f(x) = \alpha x + \beta$$

Hence, the general solution to our ODE becomes:

$$y(x) = Ae^{\lambda x} + Bxe^{\lambda x}$$

Where  $A, B \in \mathbb{R}$  are constants that can be determined by using the boundary conditions.

**Proposition 4.3** (Complex conjugate roots) Suppose we are in the last case, where our roots are complex and conjugate. Therefore:

$$\lambda_{1,2} = \alpha \pm \beta i$$

Therefore:

$$y(x) = C_1 e^{(\alpha + \beta i)x} + C_2 e^{(\alpha - \beta i)x}$$

By factorizing  $e^{\alpha x}$  and utilizing Euler's identity, we deduce that the general solution is also equivalent to:

$$y(x) = e^{\alpha x} (A \sin \beta x + B \cos \beta x)$$

Where  $A, B \in \mathbb{R}$  are constants that can be determined by using the boundary conditions. Note that we used Euler's identity to deduce that:

$$e^{\beta i x} = \cos \beta x + i \sin \beta x$$

### 4.2.2 Constant-coefficients non-homogeneous second-order ODEs

Suppose we have to solve an equation of the type:

$$a \frac{d^2 y}{dx^2} + b \frac{dy}{dx} + cy = f(x)$$

Where  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a fixed function.

**Theorem 4.1** Let us consider a constant-coefficients non-homogeneous second-order ODEs. Consider we know a particular solution  $y_0(x)$  that satisfies the differential equation. Therefore, the general solution of the ODE is:

$$y(x) = y_C(x) + y_0(x)$$

Where  $y_C(x)$  is the solution to the homogeneous version of the same ODE.

*Proof.* The particular solution (also called the particular integral) satisfies the equation:

$$ay_0''(x) + by_0'(x) + cy_0(x) = f(x)$$

Also, we are looking for a general solution that satisfies:

$$ay''(x) + by'(x) + cy(x) = f(x)$$

By subtracting the two equations, we obtain:

$$a(y(x) - y_0(x))'' + b(y(x) - y_0(x))' + c(y(x) - y_0(x)) = 0$$

However, this is the homogeneous version of the same ODE, with general solution  $y_C(x)$  which we already know how to determine. Hence:

$$y_C(x) = y(x) - y_0(x)$$

Therefore, we obtain:

$$y(x) = y_C(x) + y_0(x)$$

□

In order to find a particular solution, we need to look at  $f(x)$  and try a particular solution that is similar to  $f(x)$ . Furthermore, if the complementary solution contains  $f(x)$ , then we need to consider a solution of the type:

$$y_0(x) = f(x)g(x)$$

And by substituting  $y_0(x)$  into the equation, we need to determine  $g(x)$ .

This case (where the complementary function contains  $f(x)$ ) has important physical significance - these types of systems are excited at their natural frequency, and this phenomenon is known as frequency.

### 4.2.3 The Wronskian

Suppose we need to solve a second-order linear ODE of the type:

$$a \frac{d^2 y}{dx^2} + b \frac{dy}{dx} + cy = f(x)$$

**Theorem 4.2** (The Wronskian technique) Suppose the attached homogeneous equation has complementary solution of the type:

$$y_C(x) = \alpha u_1(x) + \beta u_2(x)$$

We define the Wronskian determinant as:

$$W = \begin{vmatrix} u_1 & u_1' \\ u_2 & u_2' \end{vmatrix}$$

Therefore, a particular solution to our ODE is:

$$y_0(x) = A(x)u_1(x) + B(x)u_2(x)$$

Where we define:

$$A(x) = - \int \frac{f(x)}{W} u_2(x) dx \text{ and } B = \int \frac{f(x)}{W} u_1(x) dx$$

This is a particularly useful technique whenever  $f(x)$  is complicated and there is no easy way to guess a function similar to it, that would also satisfy the differential equation.

## 5 Linear difference equations

Some problems are concerned with quantities that do not vary continuously, but are rather defined discretely. This is commonly the case for finance problems where figures might appear on a discrete-interval basis. These types of problems often lead to linear difference equations.

**Definition 5.1** (Linear difference equation) Let  $f : \mathbb{N} \rightarrow \mathbb{R}$  and  $(x_n)_{n \geq 0}$  be a real sequence. A  $k$ -step linear difference equation is an equation of the form:

$$x_{n+k} + \alpha_{k-1}x_{n+k-1} + \cdots + \alpha_0x_n = f(n)$$

If  $f(n) \equiv 0$ , we say that the equation is homogeneous. Otherwise, the linear difference equation is non-homogeneous.

**Definition 5.2** (Characteristic equation) The characteristic equation of a  $k$ -step linear difference equation is:

$$\lambda^k + \alpha_{k-1}\lambda^{k-1} + \cdots + \alpha_1\lambda + \alpha_0 = 0$$

**Proposition 5.1** (Solution for homogeneous linear difference equations) We begin by finding the roots of the attached characteristic equation. If all roots are linearly independent, our solution is of the type:

$$x_n = \sum_{i=1}^k C_i \lambda_i^n$$

Moreover, if a solution  $\lambda_i$  has order of multiplicity  $j$ , then the  $j$  solutions are:

$$C_1 \lambda_i^n, C_2 n \lambda_i^n, \dots, C_j n^{j-1} \lambda_i^n$$

This is the same technique we utilized when solving ODEs - we are simply looking for a solution of the type  $g(n)\lambda_i^n$  instead.

**Proposition 5.2** (Solution for non-homogeneous linear difference equations) Suppose we have to solve a linear difference equation where the term  $f(n)$  is not null. In this case, we must proceed in a similar fashion to what we did for differential equations. Therefore, we must look for a particular solution  $x_{n_0}$  that is similar to the function  $f(n)$ . Therefore, our general solution becomes:

$$x_n = x_{n_C} + x_{n_0}$$

Where  $x_{n_C}$  is the solution to the homogeneous equation.

Furthermore, as with ODEs we need to consider the case where the complementary solution contains  $f(n)$ . In this case, we are looking for a solution of the type:

$$x_{n_0} = g(n)f(n)$$

And we need to determine an equation in  $g(n)$ . Lastly, we need to use the given boundary conditions to determine all the constants in our solution, just as we did for differential equations.

## 6 Partial differentiation

So far, we have only considered functions of a single variable. However, most Engineering applications involve functions of more than just a single variable, as it rarely happens that a system is fully governed by a single variable. For this reason, we must look into the analysis of functions of multiple variables. Note that more will be covered in a later section of this course, and also in Part IB Vector Calculus.

### 6.1 Partial derivatives

Recall that for a function of a single variable,  $f : \mathbb{R} \rightarrow \mathbb{R}$ , the derivative is defined as:

$$g'(x) = \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x}$$

This represents the instantaneous slope of a function at any given point. We can start from this definition and introduce the partial derivative of a function of multiple variables.

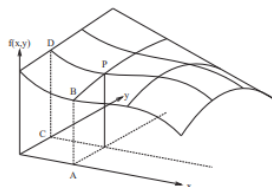
**Definition 6.1** (Partial differentiation) Consider a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . The partial derivative of  $f$  with respect to the variable  $x_i$  is defined as:

$$\frac{\partial f}{\partial x_i} = \lim_{\delta x_i \rightarrow 0} \frac{f(x_1, x_2, \dots, x_i + \delta x_i, \dots, x_n) - f(x_1, x_2, \dots, x_i, \dots, x_n)}{\delta x_i}$$

For a function of two variables  $f(x, y)$ , the partial derivatives degenerate into:

$$\begin{aligned} \frac{\partial f}{\partial x} &= \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x, y) - f(x, y)}{\delta x} \\ \frac{\partial f}{\partial y} &= \lim_{\delta y \rightarrow 0} \frac{f(x, y + \delta y) - f(x, y)}{\delta y} \end{aligned}$$

For any function of two variables, the partial derivatives  $\frac{\partial f}{\partial x}$  and  $\frac{\partial f}{\partial y}$ , can be interpreted by slicing through the surface at a point  $P$ . If we slice through using the plane  $CDP$ , then  $y$  is constant and the slope is  $\frac{\partial f}{\partial x}$ , whereas if we slice through with the plane  $ABP$ , the slope is  $\frac{\partial f}{\partial y}$ . Thus, for partial differentiation with respect to  $x$ ,  $y$  is being treated as a constant, whereas for  $y$ ,  $x$  is treated as a constant.



**Example 6.1** Compute the partial derivatives of the function  $f(x, y) = x^2 + y^3$ .

By treating  $y$  as a constant, we deduce that:

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

And by treating  $x$  as a constant, we deduce that:

$$\frac{\partial f}{\partial y} = 3y^2$$



## 6.2 Linearisation

Oftentimes we know the relationship that governs the dependency between a variable and different variables. However, we want to know what happens to our quantity if we vary the value of multiple variables. We can use partial differentiation as a way to estimate the change in a quantity when there is a known small change in the independent variables.

By using the Taylor series for a function of a single variable, we have seen that:

$$f(x + \delta x) = f(x) + f'(x)\delta x + \mathcal{O}(\delta x^2)$$

Therefore:

$$\delta f \approx f'(x)\delta x$$

This is true for a function of a single variable, where the change  $\delta x$  is small. This can then be extended to a function of multiple variables likewise:

$$\delta f \approx \sum_{i=1}^n \frac{\partial f}{\partial x_i} \delta x_i$$

**Definition 6.2** (Total differential) The total differential is defined as the infinitesimal change of the function when  $\delta x_i \rightarrow 0, 1 \leq i \leq n$ . Therefore:

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

**Definition 6.3** (Taylor series) Consider a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . The Taylor series of this function in the point  $(a_1, a_2, \dots, a_n)$  can be defined as:

$$f(x_1, x_2, \dots, x_n) = \sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} \frac{(x_1 - a_1)^{k_1} \dots (x_n - a_n)^{k_n}}{k_1! \dots k_n!} \left( \frac{\partial^{k_1 + \dots + k_n} f(a_1, \dots, a_n)}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \right)$$

### 6.3 Directional derivatives and the chain rule

**Definition 6.4** (Partial differentiation) Let us consider a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  and a vector  $\mathbf{u} \in \mathbb{R}^n$ . The directional derivative (the rate of change of  $f$  in the direction given by  $\mathbf{u}$ ) is defined as:

$$\begin{aligned} D_{\mathbf{u}}f(\mathbf{x}) &= \lim_{h \rightarrow 0} \frac{f(\mathbf{x} + h\mathbf{u}) - f(\mathbf{x})}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x_1 + hu_{x_1}, x_2 + hu_{x_2}, \dots, x_n + hu_{x_n}) - f(x_1, x_2, \dots, x_n)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x_1, x_2, \dots, x_n) + h \sum_{i=1}^n \frac{\partial f}{\partial x_i} u_{x_i} - f(x_1, x_2, \dots, x_n) + \mathcal{O}(x_1^2, x_2^2, \dots, x_n^2)}{h} \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i} u_{x_i} \end{aligned}$$

By defining the gradient of  $f$  as:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

The direction derivative is finally given by:

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}$$

**Theorem 6.1** (Chain rule I) Consider a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , with  $f = f(u)$  and  $u = u(x_1, x_2, \dots, x_n)$ . Therefore, the chain rule can be stated as:

$$\frac{\partial f}{\partial x_i} = \frac{df}{du} \frac{\partial u}{\partial x_i}$$

**Theorem 6.2** (Chain rule II) Consider a function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , with  $f = f(u(x, y), v(x, y))$ . Therefore, the chain rule can be stated as:

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} \\ \frac{\partial f}{\partial y} &= \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y} \end{aligned}$$

## 6.4 Normal vector to a surface

Consider a general surface given by:

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

By parametrizing  $x = x(s)$  and  $y = y(s)$ , we obtain:

$$f(x(s), y(s)) = c$$

We can now totally differentiate the above relationship, obtaining:

$$\frac{\partial f}{\partial x} \frac{dx}{ds} + \frac{\partial f}{\partial y} \frac{dy}{ds} = 0$$

We observe that the vector defined by  $\mathbf{t} = \begin{bmatrix} \frac{dx}{ds} \\ \frac{dy}{ds} \end{bmatrix}$  is the tangent vector to the surface.

Now, recalling our definition of  $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$ , we observe that the relationship above can be expressed as a dot product - the total derivative is essentially the directional derivative in the tangent direction to the surface. Hence:

$$\nabla f \cdot \mathbf{t} = 0$$

This tells us that the vectors  $\nabla f$  and  $\mathbf{t}$  are mutually orthogonal. Therefore, we obtain that  $\nabla f$  is a vector that is always perpendicular to the curve.

## 7 Matrices and linear transformations

In the following section, we will consider that the basic properties of matrices, such as the determinant, inverses, and so on are already known. We will only focus our attention on the new content, as presented in the Mathematics lectures.

### 7.1 Properties of matrices

**Definition 7.1** (Linear transformation) Consider a vector  $\mathbf{x} \in \mathbb{R}^n$  and a matrix  $\mathbf{A}$ . The vector  $\mathbf{x}'$  that is obtained by applying the matrix transformation is given by:

$$\mathbf{x}' = \mathbf{A}\mathbf{x}$$

This is called a linear transformation. In order for us to be able to study the nature of the transformation, we only need to study the properties of the matrix  $\mathbf{A}$ .

**Definition 7.2** (Einstein's notation) Consider we have two matrices  $\mathbf{A}, \mathbf{B}$  and we want to express each individual element of their product  $\mathbf{AB}$ . This could be done using a summation. However, for simplicity, we can write:

$$(\mathbf{AB})_{ij} = a_{ik}b_{kj} = \sum_{k=1}^n a_{ik}b_{kj}$$

It is understood here that the summation is done over the  $k$  variable, and  $i, j$  are fixed.

**Definition 7.3** (Transpose) Consider a matrix  $\mathbf{A}$ . We define the transpose of the matrix  $\mathbf{A}^T$  as:

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

**Proposition 7.1** For any two matrices  $\mathbf{A}, \mathbf{B}$ , the transpose of their product is given by:

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

*Proof.* Let us consider the transpose of their product. We will use Einstein's notation and the definition of the transpose.

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

**Proposition 7.2** (Scalar product equivalence) Consider two column vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ . Component-wise, we can express their scalar product as the sum:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=0}^n x_i y_i$$

In matrix form, this is equivalent to saying:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$$

Where  $\mathbf{x}^T$  denotes the transpose of the column vector  $\mathbf{x}$ , which is a row vector.

## 7.2 Mappings and transformations

Matrices are used to map one vector to another. If  $\mathbf{x}' = \mathbf{A}\mathbf{x}$ , we say that  $\mathbf{A}$  maps the vector  $\mathbf{x}$  onto  $\mathbf{x}'$ . In the following sections, we will study the two main types of transformations - rotations and reflections; and pure stretches. This can be further generalized to higher dimensions.

First, it would be useful to have a general method for determining the matrix  $\mathbf{A}$  corresponding to a given transformation. We have to consider the transformation given by  $\mathbf{A}$  on each unit vector of our coordinate system, and then combine those effects to effectively determine  $\mathbf{A}$ .

### 7.2.1 Rotation of a vector in two-dimensions

Consider a rotation of a vector through an angle  $\alpha$  in the positive (anti-clockwise) direction. We can observe that the vector  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  is going to be mapped onto  $\begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}$ . Furthermore, the vector  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$  is mapped to  $\begin{bmatrix} -\sin \alpha \\ \cos \alpha \end{bmatrix}$ . By combining the effects that the matrix has on both vectors, we obtain that the matrix representing a rotation of  $\alpha$  is given by:

$$\mathbf{A} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

And therefore, we can deduce that the transformation representing a rotation of  $\alpha$  can be written as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Furthermore, because the matrix  $\mathbf{A}$  represents an anti-clockwise rotation of  $\alpha$ , the matrix  $\mathbf{A}^{-1}$  represents a clockwise rotation of  $\alpha$ , and is given by:

$$\mathbf{A}^{-1} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}$$

This is trivial, because a clockwise rotation of  $\alpha$  is essentially an anti-clockwise rotation of  $-\alpha$ .

**Definition 7.4** (Orthogonal matrix) A matrix  $\mathbf{Q}$  is orthogonal if:

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$$

By taking the determinant of the above expression, we obtain:

$$\det(\mathbf{Q}\mathbf{Q}^T) = 1 \iff (\det \mathbf{Q})^2 = 1$$

And therefore, the determinant of an orthogonal matrix is:

$$\det \mathbf{Q} = \pm 1$$

### 7.3 Rotation about an axis in three dimensions

Consider the case where we want to determine the coordinates of a vector after we apply a linear transformation which rotates it by  $\alpha$  about a chosen axis. In order to determine the positive direction of the rotation, we have to refer to the cross product between the basis vectors. For  $\mathbb{R}^3$ , we have  $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ ,  $\mathbf{j} \times \mathbf{k} = \mathbf{i}$  and  $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ . We now have to find the corresponding relationship where the product is equal to the axis about we wish to rotate. Suppose we wish to determine the matrix that rotates a vector by  $\alpha$  about the  $y$  axis. Referring to the cross products, we see that:

$$\mathbf{k} \times \mathbf{i} = \mathbf{j}$$

Therefore, the positive direction of the rotation is from the  $z$  axis to the  $x$  axis. By considering each individual vector, we observe that  $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$  is mapped to  $\begin{bmatrix} \cos \alpha \\ 0 \\ -\sin \alpha \end{bmatrix}$ ,  $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$  is unchanged, and  $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$  is mapped to  $\begin{bmatrix} \sin \alpha \\ 0 \\ \cos \alpha \end{bmatrix}$ . Therefore, we can observe that a rotation of  $\alpha$  about the  $y$  axis is given by the linear transformation:

$$\mathbf{A} = \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix}$$



## 7.4 Properties of rotation matrices

**Proposition 7.3** (Orthogonality) Any rotation matrix  $\mathbf{Q}$  is orthogonal. Moreover, the components of its column space are mutually orthogonal. Furthermore, since the matrix  $\mathbf{Q}^T$  is also orthogonal by definition, we obtain that its columns are mutually orthogonal, i.e. the components of the row space of the matrix  $\mathbf{Q}$  are orthogonal.

Therefore, for an orthogonal matrix, both its row space and column space form an orthogonal basis.

**Proposition 7.4** (Inverse) Suppose  $\mathbf{Q}$  is an orthogonal matrix. Therefore:

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$$

Multiplying this relationship by  $\mathbf{Q}^{-1}$  to the left, we finally obtain:

$$\mathbf{Q}^{-1} = \mathbf{Q}^T$$

**Proposition 7.5** (Length invariance) Let  $\mathbf{Q}$  be an orthogonal matrix. The linear transformation determined by  $\mathbf{Q}$  preserves the length of any vector.

*Proof.* For any vector  $\mathbf{x}$ , the linear transformation above can be written as:

$$\mathbf{x}' = \mathbf{Q}\mathbf{x}$$

The modulus of any vector is equal to the dot product  $\mathbf{x} \cdot \mathbf{x}$ . Therefore, we can determine the length of  $\mathbf{x}'$ :

$$\mathbf{x}' \cdot \mathbf{x}' = (\mathbf{x}')^T \mathbf{x}' = \mathbf{x}^T \mathbf{Q}^T \mathbf{Q} \mathbf{x}$$

Since matrix multiplication is associative, and by noting that  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$ , we obtain:

$$\mathbf{x}' \cdot \mathbf{x}' = \mathbf{x}^T \mathbf{x} = \mathbf{x} \cdot \mathbf{x}$$

Therefore, since lengths are positive definite, we obtain:

$$|\mathbf{x}'| = |\mathbf{x}|$$

□

Similarly, we can prove that the scalar product between any two vectors is preserved, i.e. the angle between the vectors is preserved.

## 7.5 Reflections

Let us now consider the case of reflections. Consider an orthogonal matrix  $\mathbf{Q}$ . We observe that in  $\mathbf{R}^3$ :

$$\det \mathbf{Q} = \det \mathbf{Q}^T = \mathbf{q}_1 \cdot (\mathbf{q}_2 \times \mathbf{q}_3)$$

Where  $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$  represent the mutually orthogonal column vectors of the matrix. This is the scalar triple product, essentially the volume of a cube (as they are mutually orthogonal). We can observe that if the determinant is positive (the case of a rotation), the column vectors form a right-handed set. Otherwise, we obtain a reflection (or a number of rotations + an odd number of reflections), which forms a left-handed set.

**Proposition 7.6** (Reflection-rotation equivalence) Two reflections form a rotation.

**Example 7.1** Determine the orthogonal matrix  $\mathbf{Q}$  that represents a reflection in the plane  $x = 0$ .

By considering the unit vector of the  $x$  axis, we observe that  $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$  is mapped onto  $\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$ .

The unit vectors of the  $y$  and  $z$  axis are not affected by this reflection (they are left unchanged). Hence, the orthogonal matrix representing this transformation is:

$$\mathbf{Q} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

As a sanity check, we can observe that  $\det \mathbf{Q} = -1$ , which means that the matrix represents, in fact, a reflection.

## 7.6 Change of basis

The rotations considered so far have been physical transformations in which it is the vectors that have been rotated. A related problem is changing to a new set of coordinates. This is a common thing to do as many problems are greatly simplified by a judicious choice of direction for the axes. The physical meaning of the vector is not changed, just the components and basis with which it is represented.

### 7.6.1 Vectors

Consider a vector  $\mathbf{v}$  with components  $(v_1, v_2)$  in a coordinate system with basis vectors  $\mathbf{e}_1, \mathbf{e}_2$ . We want to find the components of this vector in a new coordinate system that is given by applying an orthogonal linear transformation  $\mathbf{Q}$  to the vectors  $\mathbf{e}_1, \mathbf{e}_2$ .

Initially, we know that  $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2$  and  $\mathbf{v}' = v'_1\mathbf{e}'_1 + v'_2\mathbf{e}'_2$ . Therefore, the components of the vector are:

$$v_i = \mathbf{v} \cdot \mathbf{e}_i \text{ and } v'_i = \mathbf{v} \cdot \mathbf{e}'_i$$

However, we know that  $\mathbf{e}'_i = \mathbf{Q}\mathbf{e}_i$ . Therefore:

$$v'_i = \mathbf{v} \cdot \mathbf{Q}\mathbf{e}_i = \mathbf{v}^T \mathbf{Q}\mathbf{e}_i = (\mathbf{Q}^T \mathbf{v})^T \mathbf{e}_i = (\mathbf{Q}^T \mathbf{v}) \cdot \mathbf{e}_i$$

Therefore, by condensing all the relationships together, if we apply a transformation  $\mathbf{Q}$  to the coordinate system, the same vector can be expressed in the new system by:

$$\mathbf{v} \rightarrow \mathbf{Q}^T \mathbf{v}$$

Therefore, if we decide to apply a transformation  $\mathbf{Q}$  to the system, the vector's components in the new system are given by applying the inverse of that transformation ( $\mathbf{Q}^T$ ) to the vector.

In general, in the case where  $\mathbf{Q}$  is not orthogonal, but is invertible, it can be proved that the new vector can be expressed as:

$$\mathbf{v}' = \mathbf{Q}^{-1} \mathbf{v}$$

### 7.6.2 Matrices

Consider a linear transformation  $\mathbf{A}$  that maps a vector  $\mathbf{b}$  into another vector  $\mathbf{c}$ . This relationship can be expressed as:

$$\mathbf{c} = \mathbf{A}\mathbf{b}$$

Suppose we now apply a transformation  $\mathbf{R}$  to the coordinate system. We now wish to determine the matrix  $\mathbf{A}'$  that relates the two new vectors. By the above, we deduce that the vectors will be linked by:  $\mathbf{c}' = \mathbf{R}^{-1}\mathbf{c}$  and  $\mathbf{b}' = \mathbf{R}^{-1}\mathbf{b}$ . Therefore:

$$\mathbf{c} = \mathbf{R}\mathbf{c}' \text{ and } \mathbf{b} = \mathbf{R}\mathbf{b}'$$

Therefore:

$$\mathbf{R}\mathbf{c}' = \mathbf{A}\mathbf{R}\mathbf{b}'$$

By multiplying with the inverse to the left, we obtain:

$$\mathbf{c}' = \mathbf{R}^{-1}\mathbf{A}\mathbf{R}\mathbf{b}'$$

Therefore, the matrix that relates the vectors in the new coordinate system is effectively given by:

$$\mathbf{A}' = \mathbf{R}^{-1}\mathbf{A}\mathbf{R}$$

Furthermore, if  $\mathbf{R}$  is orthogonal, the relationship above becomes:

$$\mathbf{A}' = \mathbf{R}^T\mathbf{A}\mathbf{R}$$

**Example 7.2** A certain transformation in two dimensions is given by  $\mathbf{A} = \begin{bmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{bmatrix}$ . Determine the matrix  $\mathbf{A}'$  that represents the same transformation in a coordinate system rotated by  $\frac{\pi}{4}$ .

A rotation of  $\frac{\pi}{4}$  can be represented by the matrix:

$$\mathbf{R} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

By the above theorem, the matrix  $\mathbf{A}'$  is given by:

$$\mathbf{A}' = \mathbf{R}^T\mathbf{A}\mathbf{R}$$

Hence:

$$\mathbf{A}' = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Finally, we obtain our matrix:

$$\mathbf{A}' = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

## 8 Eigenvalues and eigenvectors

### 8.1 The eigenspace of a matrix

In the previous chapter, we have discussed several methods for providing more insight into linear transformations. However, nearly all transformations leave certain characteristic directions unchanged. Such directions are known as the eigenvectors of the transformation, and are of particular interest to us.

**Definition 8.1** (Eigenvector) Consider a matrix  $\mathbf{A}$ . A non-zero vector  $\mathbf{x}$  that obeys the relationship  $\mathbf{Ax} = \lambda\mathbf{x}$  is said to be an eigenvector of  $\mathbf{A}$ , and the scalar  $\lambda$  is the eigenvector's corresponding eigenvalue.

If  $\mathbf{A}$  is regarded as a linear transformation, then the eigenvector  $\mathbf{x}$  is the direction left unchanged by the transformation, and  $\lambda$  is the scale factor. Essentially, by applying the transformation to these directions, we obtain the same vector, just scaled by  $\lambda$ .

Note that:

- (a) Whilst it is very important that the eigenvector  $\mathbf{x}$  is non-zero, the eigenvalue  $\lambda$  can be null
- (b) If  $\mathbf{x}$  is an eigenvector of  $\mathbf{A}$ , then so is  $k\mathbf{x}$ , where  $k \in \mathbb{R}$
- (c) It is customary, but not mandatory, to normalize the eigenvectors. It is, however, absolutely mandatory to normalize them whenever we want to diagonalize a matrix. We will look into this further down below.

**Proposition 8.1** (Determining the eigenvalues) Suppose we want to determine the eigenvalues of a matrix  $\mathbf{A}$ . Therefore:

$$\mathbf{Ax} = \lambda\mathbf{x}$$

This is equivalent to:

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

Since the vector  $\mathbf{x}$  is non-zero, the above can only be true if:

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

If  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , the above equation has  $n$  solutions  $\lambda_1, \lambda_2, \dots, \lambda_n$  that are not necessarily distinct. Furthermore, in order for us to determine the eigenvectors of  $\mathbf{A}$ , we simply have to solve:

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

This will give us an indeterminate system. We generally use Cramer's method (we determine the secondary unknown variables and set them equal to  $\alpha$  and determine the other variables in terms of  $\alpha$ ). Afterwards, we just set  $\alpha = 1$  (and normalize the eigenvector if necessary).

**Theorem 8.1** For any matrix  $\mathbf{A}$ , the product of its eigenvalues is equal to the determinant of the matrix, i.e.:

$$\det \mathbf{A} = \prod_{i=1}^n \lambda_i$$

**Theorem 8.2** For any matrix  $\mathbf{A}$ , the sum of its eigenvalues is equal to the trace of the matrix, i.e.:

$$\text{Tr } \mathbf{A} = \sum_{i=1}^n \lambda_i$$

**Theorem 8.3** Consider a matrix  $\mathbf{A}$ . If  $\lambda_i$  is an eigenvalue of  $\mathbf{A}$ , then  $\lambda_i^n$  is an eigenvalue of  $\mathbf{A}^n$ .

## 8.2 Properties of symmetric matrices

**Definition 8.2** (Symmetric matrix) A matrix  $\mathbf{A}$  is said to be symmetric if  $\mathbf{A} = \mathbf{A}^T$ . Also, a matrix is said to be anti-symmetric if  $\mathbf{A} = -\mathbf{A}^T$ .

**Theorem 8.4** Let  $\mathbf{S}$  be a real, symmetric matrix. Therefore, the eigenspace of  $\mathbf{S}$  is real.

*Proof.* Consider a real, symmetric matrix  $\mathbf{S}$ . Suppose  $\mathbf{S}$  has a complex eigenvalue  $\lambda = \alpha + i\beta$  and a corresponding complex eigenvector  $\mathbf{x} = \mathbf{u} + i\mathbf{v}$ . Therefore:

$$\mathbf{S}\mathbf{x} = \lambda\mathbf{x} \text{ and } \mathbf{S}\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$$

By dotting these equations with the eigenvectors, we obtain:

$$\bar{\mathbf{x}} \cdot (\mathbf{S}\mathbf{x}) = \lambda\bar{\mathbf{x}} \cdot \mathbf{x} = \lambda(\mathbf{u}^2 + \mathbf{v}^2)$$

$$\mathbf{x} \cdot (\mathbf{S}\bar{\mathbf{x}}) = \bar{\lambda}\mathbf{x} \cdot \bar{\mathbf{x}} = \bar{\lambda}(\mathbf{u}^2 + \mathbf{v}^2)$$

Let us define  $Y = \bar{\mathbf{x}} \cdot (\mathbf{S}\mathbf{x}) = \bar{\mathbf{x}}^T \mathbf{S}\mathbf{x}$ . Furthermore,  $Y^T = \mathbf{x}^T \mathbf{S}\bar{\mathbf{x}} = \mathbf{x} \cdot (\mathbf{S}\bar{\mathbf{x}})$ , because  $\mathbf{S}^T = \mathbf{S}$ . Therefore, the two scalar products defined above are equal, which means that:

$$\lambda = \bar{\lambda}$$

Therefore, we obtain:

$$\mathbf{S}\mathbf{x} = \lambda\mathbf{x} \text{ and } \mathbf{S}\bar{\mathbf{x}} = \lambda\bar{\mathbf{x}}$$

By summing these together:

$$\mathbf{S}(\mathbf{x} + \bar{\mathbf{x}}) = \lambda(\mathbf{x} + \bar{\mathbf{x}})$$

However,  $\mathbf{x} + \bar{\mathbf{x}} = \mathbf{u}$ , as defined before. This means that:

$$\mathbf{S}\mathbf{u} = \lambda\mathbf{u}$$

And since  $\mathbf{u}$  is a real eigenvector, we have concluded our proof.  $\square$

**Theorem 8.5** (Orthogonality of eigenvectors) Consider a real, symmetric matrix  $\mathbf{S}$ . Then, the eigenvectors of  $\mathbf{S}$  are orthogonal.

*Proof.* Consider a real, symmetric matrix  $\mathbf{S}$ . Let  $\lambda_1, \lambda_2$  be two distinct eigenvalues of  $\mathbf{S}$  and  $\mathbf{x}_1, \mathbf{x}_2$  be their corresponding eigenvectors. Therefore  $\mathbf{S}\mathbf{x}_1 = \lambda_1\mathbf{x}_1$  and  $\mathbf{S}\mathbf{x}_2 = \lambda_2\mathbf{x}_2$ . We define the quantity  $Y_1 = \mathbf{x}_2 \cdot (\mathbf{S}\mathbf{x}_1) = \lambda_1\mathbf{x}_1 \cdot \mathbf{x}_2$  and  $Y_2 = \mathbf{x}_1 \cdot (\mathbf{S}\mathbf{x}_2) = \lambda_2\mathbf{x}_1 \cdot \mathbf{x}_2$ . However:

$$Y_1 = \mathbf{x}_2^T \mathbf{S}\mathbf{x}_1 \iff Y_1^T = \mathbf{x}_1^T \mathbf{S}\mathbf{x}_2 = Y_2$$

However, since  $Y_1, Y_2$  are scalars, that means that  $Y_1 = Y_2$ , and hence:

$$\lambda_1\mathbf{x}_1 \cdot \mathbf{x}_2 = \lambda_2\mathbf{x}_1 \cdot \mathbf{x}_2$$

Given that  $\lambda_1 \neq \lambda_2$ , we must have:

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$$

Thus the two eigenvectors are orthogonal, therefore concluding our proof.  $\square$



### 8.3 Diagonalizing matrices

The eigenspace of a matrix allows us to write matrices with linearly independent eigenvectors in a convenient "diagonal" form. This allows us to easily raise matrices to certain powers, and to prove certain properties if two matrices have the same eigenvectors (for instance, commutativity).

Let  $\mathbf{A}$  be a matrix with linearly independent eigenvectors. Therefore, for each eigenpair:

$$\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i$$

By defining the matrix  $\mathbf{U} = \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix}$ , we can rewrite the equations in a compact form as:

$$\mathbf{A}\mathbf{U} = \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \lambda_1\mathbf{u}_1 & \lambda_2\mathbf{u}_2 & \dots & \lambda_n\mathbf{u}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix}$$

The matrix on the right hand side of the equation can be split into:

$$\begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \lambda_1\mathbf{u}_1 & \lambda_2\mathbf{u}_2 & \dots & \lambda_n\mathbf{u}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

By defining  $\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$ , the relationship can be expressed as:

$$\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{\Lambda}$$

Finally, if the matrix  $\mathbf{U}$  is invertible (which it is if the eigenvectors are independent), we obtain:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$$

Furthermore, if  $\mathbf{A}$  is a symmetric matrix, then the relationship becomes:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

Furthermore, we can express the matrix of eigenvalues as:

$$\mathbf{\Lambda} = \mathbf{U}^T\mathbf{A}\mathbf{U}$$

Therefore,  $\mathbf{\Lambda}$  represents the same linear transformation as  $\mathbf{A}$ , but in a basis that is aligned with the eigenvectors (assuming  $\mathbf{A}$  is symmetric). Note that when diagonalizing a matrix, we must always use normalized eigenvectors.

**Example 8.1** Suppose we have a matrix  $\mathbf{A}$  with linearly independent eigenvectors. Since we can write:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$$

By repeated multiplication, we obtain:

$$\mathbf{A}^k = \mathbf{U}\mathbf{\Lambda}^k\mathbf{U}^{-1}$$

The intricacy of this approach lies within the fact that  $\mathbf{\Lambda}$  is a diagonal matrix. Hence:

$$\mathbf{\Lambda}^k = \begin{bmatrix} \lambda_1^n & 0 & \dots & 0 \\ 0 & \lambda_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix}$$

## 8.4 Generalized eigenvalue problems

**Definition 8.3** A non-zero vector  $\mathbf{x}$  that satisfies  $\mathbf{Ax} = \lambda\mathbf{Mx}$  is an eigenvector of the above generalized problem.

If  $\mathbf{M}$  is invertible, this problem becomes a regular eigenvalue problem:

$$\mathbf{M}^{-1}\mathbf{Ax} = \lambda\mathbf{x}$$

Now, for the  $j^{\text{th}}$  eigenpair  $(\lambda_j, \mathbf{x}_j)$ :

$$\mathbf{Ax}_j = \lambda_j\mathbf{Mx}_j$$

Taking the scalar product with the eigenvector  $\mathbf{x}_i$ :

$$\mathbf{x}_i \cdot (\mathbf{Ax}_j) = \mathbf{x}_i^T \mathbf{Ax}_j = \lambda_j \mathbf{x}_i^T \mathbf{Mx}_j$$

In the case where  $\mathbf{A}, \mathbf{M}$  are symmetric matrices:

$$\mathbf{x}_i^T \mathbf{Ax}_j = \mathbf{x}_j^T \mathbf{Ax}_i \text{ and } \mathbf{x}_i^T \mathbf{Mx}_j = \mathbf{x}_j^T \mathbf{Mx}_i$$

This means that, when  $i \neq j$ :

$$\mathbf{x}_i \cdot (\mathbf{Ax}_j) = \mathbf{x}_i^T \mathbf{Ax}_j = 0$$

$$\mathbf{x}_i \cdot (\mathbf{Mx}_j) = \mathbf{x}_i^T \mathbf{Mx}_j = 0$$

Hence, the eigenvectors are orthogonal with respect to both  $\mathbf{A}$  and  $\mathbf{M}$ . Furthermore, for the case where  $i = j$ :

$$\mathbf{x}_i^T \mathbf{Ax}_i = \lambda_i \mathbf{x}_i^T \mathbf{Mx}_i$$

If the matrix  $\mathbf{M}$  is symmetric positive definite, meaning  $\mathbf{x}^T \mathbf{Mx} > 0$  for any non-zero  $\mathbf{x}$ , we can normalize our eigenvector using:

$$\mathbf{u}_i = \frac{\mathbf{x}_i}{\sqrt{\mathbf{x}_i^T \mathbf{Mx}_i}}$$

Therefore:

$$\mathbf{u}_i^T \mathbf{Mu}_i = 1$$

Lastly, we obtain:

$$\mathbf{u}_i^T \mathbf{Au}_i = \lambda_i \mathbf{u}_i^T \mathbf{Mu}_i = \lambda_i$$