Week 4

1.6.2 Orthogonal operators in $\mathbb{E}^3$ and rotations

In the previous section, we saw that the orthogonal operators of $\mathbb{E}^2$ that preserved orientation were rotation operators. The main result in this section will be Euler's Rotation Theorem that shows the same holds in $\mathbb{E}^3$. We will give the precise statement of Euler's Theorem at the end of this section. For now, we will just formulate a preliminary statement:

Theorem 1.75 *An orthogonal operator in $\mathbb{E}^3$ that preserves orientation is a rotation about an axis L by the angle φ.*

We shall slowly build up this statement by precisely defining rotations in $\mathbb{E}^3$ and how we can derive the axis and angle via standard linear algebra techniques.

Recall that a linear operator on $\mathbb{E}^2$ is a rotation by φ if it is of the form P_φ, see (1.22) and (1.23). The definition of a rotation is little bit more subtle in $\mathbb{E}^3$. Explicitly, a rotation in $\mathbb{E}^3$ occurs around an _______.

Let $\mathbf{n} \neq \mathbf{0}$ be an arbitrary non-zero vector in $\mathbb{E}^3$. Consider the line

$$L_{\mathbf{n}} = \operatorname{span}(\mathbf{n}) =$$

spanned by vector $\mathbf{n}$. We say $L_{\mathbf{n}}$ is the *axis* directed along the vector $\mathbf{n}$.

Note that $L_{\mathbf{n}}$ depends only on the _______________ of the vector $\mathbf{n}$, not the magnitude, *i.e.*, $L_{\mathbf{n}} = L_{\lambda\mathbf{n}}$ for all $\lambda \neq 0$. As a result, we shall often consider the *normalisation* of $\mathbf{n}$

Wikipedia: normalisation

$$\widehat{\mathbf{n}} = \frac{\mathbf{n}}{\|\mathbf{n}\|}, \tag{1.27}$$

i.e., the unit vector in the direction of $\mathbf{n}$. This will allow us to work with an orthonormal basis.

Definition 1.76 Let P be a linear operator on $\mathbb{E}^3$, and $\mathcal{B} = (\hat{\mathbf{n}}, \mathbf{f}, \mathbf{g})$ an orthonormal basis. We say P is a *rotation about the axis* $L_{\mathbf{n}}$ by the

Wikipedia: rotation about the axis

angle φ if it acts on $\mathcal{B}$ as follows:

$$P(\widehat{\mathbf{n}}) = \tag{1.28}$$

$$P(\mathbf{f}) = \tag{1.29}$$

$$P(\mathbf{g}) = \tag{1.30}$$

i.e., the matrix representation of P in the basis $\mathcal{B}$ is

$$[P]_{\mathcal{B}} = \tag{1.31}$$

Let's break down this definition and see what is happening geometrically.

- (1.28) implies that P fixes all vectors on the axis $L_{\mathbf{n}} = L_{\widehat{\mathbf{n}}}$, as

  $$P(\lambda\widehat{\mathbf{n}}) = \lambda P(\widehat{\mathbf{n}}) = \lambda\widehat{\mathbf{n}}$$

  for any scalar $\lambda \in \mathbb{R}$. Note that in linear algebra terms, this is equivalent to $\mathbf{n}$ (and, by Lemma 1.55, any scalar multiple of $\mathbf{n}$) being an __________________ of P with _________________.

- (1.29) and (1.30) state that if we restrict to the two dimensional plane spanned by $(\mathbf{f}, \mathbf{g})$, then P behaves exactly like a two dimensional rotation and rotates the plane by the angle φ. As the basis is orthonormal, this plane is orthogonal to $\widehat{\mathbf{n}}$, and so the plane rotates around the axis _____.

Recalling the definition of the trace of linear operator, we see from (1.31) that

$$\mathrm{Tr}(P) = 1 + 2\cos\varphi \tag{1.32}$$

where φ is angle of rotation. Recall by Proposition 1.50, that the trace of P does not depend on the choice of basis. This formula determines the cosine of the angle of rotation purely in terms of the operator rather than its matrix with respect to a basis. Furthermore, as $\cos(\varphi) = \cos(-\varphi)$ it determines the angle of rotation *up to sign*.

It is quick to check that P is orthogonal and preserves orientation. Euler's Theorem states that the converse is true: any orientation preserving orthogonal operator of $\mathbb{E}^3$ is a rotation.

$[P]_{\mathcal{B}}$ is an orthogonal matrix, therefore Proposition 1.58 implies P is orthogonal. It preserves orientation as $\det P = \det [P]_{\mathcal{B}} = 1$

Theorem 1.77 (Euler's Rotation Theorem) *Let P be an orientation preserving orthogonal operator of $\mathbb{E}^3$. Then P is a rotation around an axis L by the angle φ.*

Explicitly, the axis L is the one dimensional space of eigenvectors of

P with eigenvalue __, and the angle φ is determined up to sign by

$$\mathrm{Tr}(P) = 1 + 2\cos\varphi.$$

Not only does this theorem completely characterise which operators define rotations, it tells us precisely what the geometry of that rotation is using only linear algebra. Before we prove this, let's consider an example that demonstrates the power of Euler's Theorem.

Example 1.78 Consider the linear operator P we saw in Example 1.56 that maps the orthonormal basis $\mathcal{B} = (\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ to

$$P(\mathbf{e}_x) = \mathbf{e}_y\ ,\ P(\mathbf{e}_y) = \mathbf{e}_x\ ,\ P(\mathbf{e}_z) = -\mathbf{e}_z. \tag{1.33}$$

Is P an orientation preserving orthogonal operator? If so, what is the angle of rotation?

Remark 1.79 The identity operator I that leaves everything fixed, *i.e.*,

$$I(\mathbf{x}) = \mathbf{x} \text{ for all } \mathbf{x} \in \mathbb{E}^3$$

is an orthogonal operator that preserves orientation. However, intuitively we can see that I doesn't rotate anything, and this is reflected by Euler's Theorem. As $I(\mathbf{x}) = \mathbf{x}$ for all vectors, every vector is an eigenvector with eigenvalue 1, and so ________________ in $\mathbb{E}^3$ could be considered an axis of rotation. Furthermore, the trace of I is $\mathrm{Tr}(I) = 3$, therefore the angle φ is zero.

The proof of the Euler's Theorem has two parts: firstly that there exists an axis that P leaves fixed, and secondly that P rotates the remaining vectors around that axis. We prove these in the following two lemmas.

Lemma 1.80 *Let P be an orientation preserving orthogonal operator of $\mathbb{E}^3$. There exists an axis L such that $P(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in L$.*

Proof. If P has eigenvalue 1, then L is the one-dimensional space of eigenvectors $\mathbf{x}$ with eigenvalue 1, *i.e.*, $P(\mathbf{x}) = \mathbf{x}$. Therefore we show that P has an eigenvalue 1.

The polynomial $C_P(\lambda) = \det(\lambda I - P)$ is called the *characteristic polynomial of P*. The roots of this polynomial are exactly the eigenvalues of P.
Wikipedia: Characteristic polynomial

Recall from linear algebra that P has an eigenvalue λ if and only if $\det(\lambda I - P) = 0$, where I is the identity operator. Therefore it remains to show that $\det(I - P) = 0$. We show this using many of the determinant properties stated in Proposition 1.46.

$$\begin{aligned}
\det(I-P) &= \det(P)\det(I-P) && (P \text{ preserves orientation})\\
&= \det(P^\mathsf{T})\det(I-P) && (\det M = \det M^\mathsf{T})\\
&= \det(P^\mathsf{T} - P^\mathsf{T}P) && (\det MN = \det M \det N)\\
&= \det(P^\mathsf{T} - I) && (P \text{ orthogonal})\\
&= \det(P^\mathsf{T} - I^\mathsf{T}) &&\\
&= \det((P-I)^\mathsf{T}) && (M^\mathsf{T} + N^\mathsf{T} = (M+N)^\mathsf{T})\\
&= \det(P-I) &&\\
&= -\det(I-P) && (\det(-M) = (-1)^n \det(M))
\end{aligned}$$

Therefore $\det(I-P) = -\det(I-P) = 0$ and so 1 is an eigenvalue of P. By Lemma 1.55, the span of a corresponding eigenvector $\mathbf{n}$ forms a fixed axis $L_\mathbf{n}$. □

Lemma 1.81 *Let P be an orientation preserving orthogonal operator of $\mathbb{E}^3$ that fixes an axis L. Then P is a rotation about the axis L by some angle φ.*

Proof. Pick a unit vector $\widehat{\mathbf{n}}$ in L and pick an arbitrary orthonormal basis $(\widehat{\mathbf{n}}, \mathbf{f}, \mathbf{g})$ with $\widehat{\mathbf{n}}$ as the first basis vector. By Lemma 1.80 we know that $P(\widehat{\mathbf{n}}) = \widehat{\mathbf{n}}$, we wish to show that

$$P(\mathbf{f}) = 0 \cdot \widehat{\mathbf{n}} + \alpha\mathbf{f} + \beta\mathbf{g}, \quad P(\mathbf{g}) = 0 \cdot \widehat{\mathbf{n}} + \gamma\mathbf{f} + \delta\mathbf{g}, \quad \alpha, \beta, \gamma, \delta \in \mathbb{R}.$$

i.e., they have no $\widehat{\mathbf{n}}$ component.

Suppose $P(\mathbf{f}) = \mu\widehat{\mathbf{n}} + \alpha\mathbf{f} + \gamma\mathbf{g}$ and consider the inner product

$\langle P(\mathbf{f}), P(\widehat{\mathbf{n}})\rangle$. By inner product manipulation, we see that

$$\begin{aligned}\langle P(\mathbf{f}), P(\widehat{\mathbf{n}})\rangle &= \langle \mu\widehat{\mathbf{n}} + \alpha\mathbf{f} + \gamma\mathbf{g}, \widehat{\mathbf{n}}\rangle \\ &= \mu\langle \widehat{\mathbf{n}}, \widehat{\mathbf{n}}\rangle + \alpha\langle \mathbf{f}, \widehat{\mathbf{n}}\rangle + \gamma\langle \mathbf{g}, \widehat{\mathbf{n}}\rangle \\ &= \mu\end{aligned}$$

However, as P is orthogonal, we have $\langle P(\mathbf{f}), P(\widehat{\mathbf{n}})\rangle = \langle \mathbf{f}, \widehat{\mathbf{n}}\rangle = 0$. Therefore $\mu = 0$ and $P(\mathbf{f})$ has no $\widehat{\mathbf{n}}$ component. A similar calculation holds for $P(\mathbf{g})$. As a result, the matrix of P is

$$[P]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \gamma & \delta \end{bmatrix}$$

As the matrix is orthogonal and orientation preserving, we can repeat the calculations as in (1.20) and (1.22) to show that

$$P(\mathbf{f}) = \mathbf{f}\cos\varphi + \mathbf{g}\sin\varphi, \quad P(\mathbf{g}) = -\mathbf{f}\sin\varphi + \mathbf{g}\cos\varphi.$$

Therefore P is the rotation about the axis L by φ. □

1.7 Area, volume and determinant

You may have seen before that the area of a parallelogram and the volume of a parallelepiped can be calculated in terms of the vector (cross) product, which in turn is related to determinants. In this section we will give a rigorous definition of the vector product and prove the link to determinants. These formulas will help develop a geometrical understanding of the determinant of a linear operator.

1.7.1 Vector product in oriented $\mathbb{E}^3$

We begin with a formal definition of the vector product for an oriented 3-dimensional Euclidean vector space.

Definition 1.82 (Vector product) Let $V = \mathbb{E}^3$ be a 3-dimensional oriented Euclidean vector space. A *vector product* (also known as a *cross product*) is a map,

Wikipedia: vector product

$$- \times - \colon V \times V \to V$$

satisfying the following properties:

- The vector $\mathbf{x} \times \mathbf{y} \in V$ is orthogonal to $\mathbf{x}$ and $\mathbf{y}$, that is

(VP-⊥)

- The product is anticommutative:

(VP-AC)

- The product is linear:

(VP-Lin)

- For perpendicular vectors $\mathbf{x}$ and $\mathbf{y}$, the length of the vector product, $\mathbf{x} \times \mathbf{y}$, is equal to the area of the rectangle formed by $\mathbf{x}$ and $\mathbf{y}$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \quad \Rightarrow \qquad \text{(VP-Len)}$$

- For linearly independent vectors $\mathbf{x}$ and $\mathbf{y}$ the basis

$$(\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}) \quad \text{has the same orientation as } V. \qquad \text{(VP-O)}$$

Remark 1.83 As with the inner product (also known as scalar product), we can use the anticommutativity of the inner product together with linearity in the first variable to show linearity in the second variable:

$$\begin{aligned} \mathbf{z} \times (\lambda \mathbf{x} + \mu \mathbf{y}) &= -(\lambda \mathbf{x} + \mu \mathbf{y}) \times \mathbf{z} \\ &= -\lambda(\mathbf{x} \times \mathbf{z}) - \mu(\mathbf{y} \times \mathbf{z}) = \lambda(\mathbf{z} \times \mathbf{x}) + \mu(\mathbf{z} \times \mathbf{y}). \end{aligned}$$

Wikipedia: bilinear

This means that the vector product is actually *bilinear*.

Remark 1.84 Anticommutativity implies that $(\mathbf{x} \times \mathbf{x}) = -(\mathbf{x} \times \mathbf{x})$ so the self product is always _______. Moreover, if $\mathbf{x}$ and $\mathbf{y}$ are not linearly independent, so that for some λ we have $\mathbf{y} = \lambda \mathbf{x}$, then the vector product $\mathbf{x} \times \mathbf{y}$ is equal to:

$$\mathbf{x} \times \mathbf{y} =$$

Remark 1.85 Axiom (VP-⊥) means that if $\mathbf{x}$ and $\mathbf{y}$ are linearly independent then $\mathbf{x} \times \mathbf{y}$ is perpendicular to the plane spanned by $\mathbf{x}$ and $\mathbf{y}$. This implies that $(\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y})$ is a basis.

Remark 1.86 The orientation of V is important: if W is the oriented vector space that contains the same vector space as V but comes with opposite orientation, then for vectors $\mathbf{x}$ and $\mathbf{y}$, the basis $(\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y})$ has the same orientation as V by axiom (VP-O). This basis has the opposite orientation to W and hence the vector product on W would give a vector in the opposite direction, yet still perpendicular to both $\mathbf{x}$ and $\mathbf{y}$.

1.7.2 Existence and uniqueness of the vector product

At this point we have specified a list of axioms that a vector product should satisfy, it is not yet clear that such a function need necessarily exist. Our goal now is to both show that such a product exists and also that the vector product is unique.

We shall begin by showing that if a vector product exists for the oriented 3-dimensional Euclidean space V, then this product is unique. Let us fix an orthonormal basis $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ that shares the same orientation as V. We assume that a vector product $- \times -$ exists for V and may deduce the following facts:

$$\mathbf{e}_i \times \mathbf{e}_i = \_\_ \quad \forall i \in \{1, 2, 3\} \qquad \text{by Remark 1.84.}$$

$$\begin{aligned} \mathbf{e}_1 \times \mathbf{e}_2 &= \qquad \text{for some scalar } \lambda && \text{by axiom (VP-}\perp\text{)} \\ &= && \text{by axiom (VP-Len)} \\ &= && \text{by axiom (VP-O).} \end{aligned}$$

Similarly, since the bases $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, $(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1)$ and $(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2)$ all have the same orientation:

$$\mathbf{e}_2 \times \mathbf{e}_3 = \qquad \text{and} \quad \mathbf{e}_3 \times \mathbf{e}_1 = \qquad .$$

Finally by anticommutativity (axiom (VP-AC)) we know that

$$\mathbf{e}_2 \times \mathbf{e}_1 = \qquad, \ \mathbf{e}_3 \times \mathbf{e}_2 = \qquad, \ \text{and} \ \ \mathbf{e}_1 \times \mathbf{e}_3 = \qquad .$$

The facts above mean that the axioms determine the value of the vector product for any pair of elements in an orthonormal basis. From this we can determine the value for any pair of vectors by bilinearity. Explicitly, if $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3$ and $\mathbf{w} = w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3$ are arbitrary vectors then bilinearity gives:

$$\begin{aligned} \mathbf{v} \times \mathbf{w} &= (v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3) \times (w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3) \\ &= v_1w_1(\mathbf{e}_1 \times \mathbf{e}_1) + v_1w_2(\mathbf{e}_1 \times \mathbf{e}_2) + v_1w_3(\mathbf{e}_1 \times \mathbf{e}_3) \\ &\qquad + v_2w_1(\mathbf{e}_2 \times \mathbf{e}_1) + v_2w_2(\mathbf{e}_2 \times \mathbf{e}_2) + v_2w_3(\mathbf{e}_2 \times \mathbf{e}_3) \\ &\qquad + v_3w_1(\mathbf{e}_3 \times \mathbf{e}_1) + v_3w_2(\mathbf{e}_3 \times \mathbf{e}_2) + v_3w_3(\mathbf{e}_3 \times \mathbf{e}_3) \\ &= (v_2w_3 - v_3w_2)\mathbf{e}_1 + (v_3w_1 - v_1w_3)\mathbf{e}_2 + (v_1w_2 - v_2w_1)\mathbf{e}_3 \end{aligned}$$

It is convenient to represent this formula in the following very familiar way:

$$\mathbf{v} \times \mathbf{w} = \tag{1.34}$$

At this point we have shown that the value of the vector product for two arbitrary vectors is given by the determinant formula (1.34). It remains to show that this formula satisfies all of the required axioms (VP-⊥) to (VP-O). The proof of this is left as an extended exercise (see Exercise Sheet 4, Question 4).