Methods of Theoretical Physics
Vectors, Tensors and Continuum Mechanics

Willoughby Seago

September 12, 2021

These are my notes for the *vectors, tensors, and continuum mechanics* part of the *methods of theoretical physics* course from the University of Edinburgh as part of the third year of the theoretical physics degree. When I took this course in the 2020/21 academic year it was taught by Dr Mehdi Bouzid[1]. These notes are based on the lectures delivered as part of this course and the notes provided as part of this course which were originally written by Professor Alexander Morozov[2]. The content within is correct to the best of my knowledge but if you find a mistake or just disagree with something or think it could be improved please let me know.

These notes were produced using LaTeX[3]. Graphs where plotted using Matplotlib[4], NumPy[5], and SciPy[6]. Diagrams were drawn with tikz[7].

This is version 1.1 of these notes, which is up to date as of 07/03/2021.

Willoughby Seago

s1824487@ed.ac.uk

---

[1] https://www.ph.ed.ac.uk/people/mehdi-bouzid
[2] https://ww.ph.ed.ac.uk/people/alexander-morozov
[3] https://www.latex-project.org/
[4] https://matplotlib.org/
[5] https://numpy.org/
[6] https://scipy.org/scipylib/
[7] https://www.ctan.org/pkg/pgf

Contents

List of Figures

Part I

Vectors and Tensors

1   Basics of Vectors

1.1   Vectors and Index Notation

> **Definition 1: Vector, initial definition**
>
> A vector is an object with length and direction.

One important types of vectors is displacement vectors, which point between points, A and B, and are denoted $\boldsymbol{a} = \vec{a} = \overrightarrow{AB}$. Displacement vectors are free to move about under parallel displacements of both ends. Another similar type is a position vectors which are fixed with one end at the origin, O, and the other at some point, P. Position vectors are often denoted $\boldsymbol{x} = \vec{x} = \overrightarrow{OP}$.

Normally we consider vectors in some specific basis. In three dimensions a basis is formed of three linearly independent vectors. The most common choice is of an orthonormal basis, $\{\boldsymbol{e_i}\}$. A basis is orthonormal if $|\boldsymbol{e_i}| = 1$ and $\boldsymbol{e_i} \perp \boldsymbol{e_j}$ for all $i \neq j$. A basis is either right or left handed depending on the orientation of the vectors. Typically we work with a right handed system. Any vector can be expressed as a linear combination of basis vectors:

$$\boldsymbol{a} = a_1\boldsymbol{e_1} + a_2\boldsymbol{e_2} + a_3\boldsymbol{e_3} = (a_1, a_2, a_3).$$

The length of such a vector is then

$$|\boldsymbol{a}| = a = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

To save writing lots of sums we use the Einstein summation convention. Under this convention the following are equivalent:

$$\boldsymbol{a} = a_1\boldsymbol{e_1} + a_2\boldsymbol{e_2} + a_3\boldsymbol{e_3}$$
$$= \sum_{i=1}^{3} a_i\boldsymbol{e_i}$$
$$= a_i\boldsymbol{e_i}.$$

The rules for the Einstein summation convention are as follows:

- One index – No sum, the indices present are said to be free indices and can take any value in $\{1, 2, 3\}$. For example:

$$a_i \in \{a_1, a_2, a_3\}, \qquad \text{or} \qquad a_i = b_i \iff a_1 = b_1 \wedge a_2 = b_2 \wedge a_3 = b_3.$$

- Two indices repeated – Sum over the repeated index. For example:

$$\boldsymbol{a} = a_i\boldsymbol{e_i} = a_1\boldsymbol{e_1} + a_2\boldsymbol{e_2} + a_3\boldsymbol{e_3}, \qquad \text{or} \qquad (A\boldsymbol{b})_i = A_{ij}b_j = A_{i1}b1 + A_{i2}b_2 + A_{i3}b_3.$$

- Three indices repeated – Not allowed. For example

$$(\boldsymbol{a} \cdot \boldsymbol{b})c_i = a_jb_jc_i \neq a_ib_ic_i, \qquad \text{or} \qquad (\boldsymbol{a} \cdot \boldsymbol{b})(\boldsymbol{c} \cdot \boldsymbol{d}) = a_ib_ic_jd_j \neq a_ib_ic_id_i$$

1.2   Scalar Product

The **scalar product**, or **dot product**, is a linear map taking two vectors to a scalar. This map is symmetric and positive definite. Formally a scalar product is defined as below:

> ### Definition 2: Inner product
>
> A real **inner product** or scalar product is a function, $\cdot : V^2 \to \mathbb{R}$, where V is a real vector space. The map must have the following properties:
> - Linearity:
> $$(\alpha \boldsymbol{a}) \cdot (\beta \boldsymbol{b}) = (\alpha \beta)(\boldsymbol{a} \cdot \boldsymbol{b}) \qquad \text{and} \qquad (\boldsymbol{a} + \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot \boldsymbol{c} + \boldsymbol{b} \cdot \boldsymbol{c}.$$
> - Symmetry:
> $$\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}.$$
> - Positive-definiteness:
> $$\boldsymbol{a} \cdot \boldsymbol{a} \geq 0$$
> with equality only if $\boldsymbol{a} = \boldsymbol{0}$.
>
> One function that fits all of these properties, and is how we will define the scalar product, is
> $$\boldsymbol{a} \cdot \boldsymbol{b} = ab \cos \vartheta$$
> where ϑ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.

Notice that if $\{\boldsymbol{e_i}\}$ is an orthonormal basis then

$$\boldsymbol{e_1} \cdot \boldsymbol{e_1} = \boldsymbol{e_2} \cdot \boldsymbol{e_2} = \boldsymbol{e_3} \cdot \boldsymbol{e_3} = 1$$

and

$$\boldsymbol{e_1} \cdot \boldsymbol{e_2} = \boldsymbol{e_1} \cdot \boldsymbol{e_3} = \boldsymbol{e_2} \cdot \boldsymbol{e_3} = 0.$$

Using this and the properties of the scalar product we have that

$$\boldsymbol{a} \cdot \boldsymbol{b} = (a_i \boldsymbol{e_i}) \cdot (b_j \boldsymbol{e_j}) = a_i b_j \boldsymbol{e_i} \cdot \boldsymbol{e_j}.$$

We define the **Kronecker delta** as

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Which allows us to compactly write

$$\boldsymbol{e_i} \cdot \boldsymbol{e_j} = \delta_{ij}$$

and

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_i b_j \delta_{ij}$$

If $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors and $\boldsymbol{c} = \boldsymbol{b} - \boldsymbol{a}$ then

$$c^2 = \boldsymbol{c} \cdot \boldsymbol{c} = (\boldsymbol{b} - \boldsymbol{a}) \cdot (\boldsymbol{b} - \boldsymbol{a}) = b^2 + a^2 - 2ab \cos \vartheta.$$

This is the **cosine rule**. If we expand this in a different way then we get

$$\boldsymbol{c} \cdot \boldsymbol{c} = \boldsymbol{b} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{a} - 2\boldsymbol{a} \cdot \boldsymbol{b}$$

$$\implies \boldsymbol{a} \cdot \boldsymbol{b} = \frac{1}{2}[\boldsymbol{b} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{a} - \boldsymbol{c} \cdot \boldsymbol{c}]$$

$$\implies a_i b_j \delta_{ij} = \frac{1}{2}[b_1^2 + b_2^2 + b_3^2 + a_1^2 + a_2^2 + a_3^2 - (b_1 - a_1)^2 - (b_2 - a_2)^2 - (b_3 - a_3)^2] = a_1 b_1 + a_2 b_2 + a_3 b_3 = a_i b_i.$$

Comparing the left and right hand side of this we see that

$$\delta_{ij} b_j = b_i.$$

1.3   Direction Cosines

> **Definition 3: Direction cosines**
>
> Let ϑ_i be the angle between $\boldsymbol{a}$ and $\boldsymbol{e_i}$. Then the **direction cosines** of $\boldsymbol{a}$ are
>
> $$l_i = \cos \vartheta_i = \frac{a_i}{a}.$$
>
> The last equality is a result of $a_i = \boldsymbol{a} \cdot \boldsymbol{e_i} = a \cos \vartheta$.

Let $\{l_i\}$ be the direction cosines of $\boldsymbol{a}$ and $\{m_i\}$ be the direction cosines of $\boldsymbol{b}$. Then

$$\cos \vartheta = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{ab} = \frac{a_i}{a}\frac{b_i}{b} = l_i m_i.$$

By this we see that $l_i l_i = \boldsymbol{a} \cdot \boldsymbol{a}/a^2 = \cos(0) = 1$.

1.4   Vector Product

> **Definition 4: Vector product**
>
> The **vector product**, or **cross product**, of $\boldsymbol{a}$ and $\boldsymbol{b}$ is a linear map, $\times \colon V^2 \to V$ with the following properties:
> - $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{0}$ if and only if $\boldsymbol{a}$ and $\boldsymbol{b}$ are collinear.
> - $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$.
> - The vector $\boldsymbol{a} \times \boldsymbol{b}$ is perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$.
>
> One function that fits this, which we shall use as the definition of the vector product, is
>
> $$\boldsymbol{a} \times \boldsymbol{b} = ab \sin(\vartheta)\hat{\boldsymbol{n}}$$
>
> where $\hat{\boldsymbol{n}}$ is a unit vector perpendicular to $\boldsymbol{a}$ and $\boldsymbol{b}$ in such a way that $(\boldsymbol{a}, \boldsymbol{b}, \hat{\boldsymbol{n}})$ is a right handed system.

Notice that if $\{\boldsymbol{e_i}\}$ is an orthonormal basis then

$$\boldsymbol{e_1} \times \boldsymbol{e_2} = \boldsymbol{e_3}, \qquad \boldsymbol{e_2} \times \boldsymbol{e_3} = \boldsymbol{e_1}, \qquad \boldsymbol{e_3} \times \boldsymbol{e_1} = \boldsymbol{e_2}$$

and

$$\boldsymbol{e_1} \times \boldsymbol{e_1} = \boldsymbol{e_2} \times \boldsymbol{e_2} = \boldsymbol{e_3} \times \boldsymbol{e_3} = \boldsymbol{0}.$$

If we define the **Levi–Civita** symbol as

$$\varepsilon_{ijk} = \begin{cases} 1, & (i,j,k) \text{ is an even permutation of } (1,2,3), \\ -1, & (i,j,k) \text{ is an odd permutation of } (1,2,3), \\ 0, & \text{otherwise}, \end{cases}$$

then we can compactly write

$$\boldsymbol{e_i} \times \boldsymbol{e_j} = \varepsilon_{ijk}\boldsymbol{e_k}.$$

For example

$$\begin{aligned} \boldsymbol{e_1} \times \boldsymbol{e_2} &= \varepsilon_{12k}\boldsymbol{e_k} \\ &= \varepsilon_{121}\boldsymbol{e_1} + \varepsilon_{122}\boldsymbol{e_2} + \varepsilon_{123}\boldsymbol{e_3} \\ &= \boldsymbol{e_3}, \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{e_2} \times \boldsymbol{e_1} &= \varepsilon_{21k}\boldsymbol{e_k} \\ &= \varepsilon_{211}\boldsymbol{e_1} + \varepsilon_{212}\boldsymbol{e_2} + \varepsilon_{213}\boldsymbol{e_3} \\ &= -\boldsymbol{e_3}. \end{aligned}$$

The cross product of two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, is

$$\boldsymbol{a} \times \boldsymbol{b} = (a_1\boldsymbol{e_1} + a_2\boldsymbol{e_2} + a_3\boldsymbol{e_3}) \times (b_1\boldsymbol{e_1} + b_2\boldsymbol{e_2} + b_3\boldsymbol{e_3})$$

$$= (a_2 b_3 - a_3 b_2)\boldsymbol{e_1} + (a_3 b_1 - a_1 b_3)\boldsymbol{e_2} + (a_1 b_2 - a_2 b_1)\boldsymbol{e_3}$$

$$= \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

$$= \begin{vmatrix} \boldsymbol{e_1} & \boldsymbol{e_2} & \boldsymbol{e_3} \\ a_1 & a_2 & a_3 \\ b_1 & b_1 & b_3 \end{vmatrix}$$

$$= (a_i \boldsymbol{e_i}) \times (b_j \boldsymbol{e_j})$$

$$= \varepsilon_{ijk} a_i b_j \boldsymbol{e_k}$$

Using the last part if we rename $(i, j, k) \to (j, k, i)$ we get $\boldsymbol{a} \times \boldsymbol{b} = \varepsilon_{jki} a_j b_k \boldsymbol{e_i} = \varepsilon_{ijk} a_j b_k \boldsymbol{e_i}$ as (j, k, i) is an even permutation of (i, j, k). Hence

$$(\boldsymbol{a} \times \boldsymbol{b})_i = \varepsilon_{ijk} a_j b_k.$$

For example:

$$\begin{aligned}
(\boldsymbol{a} \times \boldsymbol{b})_i &= \varepsilon_{ijk} a_j b_k \\
&= \varepsilon_{111} a_1 b_1 + \varepsilon_{112} a_1 b_2 + \varepsilon_{113} a_1 b_3 \\
&\quad + \varepsilon_{121} a_2 b_1 + \varepsilon_{122} a_2 b_2 + \varepsilon_{123} a_2 b_3 \\
&\quad + \varepsilon_{131} a_3 b_1 + \varepsilon_{132} a_3 b_2 + \varepsilon_{133} a_3 b_3 \\
&= \varepsilon_{123} a_2 b_3 + \varepsilon_{132} a_3 b_2 \\
&= a_2 b_3 - a_3 b_2.
\end{aligned}$$

2 Basics of Vector Calculus

2.1 More on the Vector Product

The following are important properties of the vector product:

- $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{0}$ if and only if $\boldsymbol{a}$ and $\boldsymbol{b}$ are co-linear.

- $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$, i.e. the vector product is anticommutative.

- $\boldsymbol{a} \times \boldsymbol{b}$ is perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$.

- $(\boldsymbol{a} + \boldsymbol{b}) \times \boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{c} + \boldsymbol{b} \times \boldsymbol{c}$, i.e. the vector product distributes over addition.

- $(\alpha \boldsymbol{a}) \times (\beta \boldsymbol{b}) = (\alpha \beta)(\boldsymbol{a} \times \boldsymbol{b})$ for all $\alpha, \beta \in \mathbb{R}$. This and the previous property make $\times$ a linear operation.

- $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) \neq (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}$, i.e. the vector product isn't associative.

- If $\boldsymbol{w} = \boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ then $\boldsymbol{w}$ is in the plane defined by $\boldsymbol{b}$ and $\boldsymbol{c}$.

- If $\boldsymbol{w'} = (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}$ then $\boldsymbol{w'}$ is in the plane defined by $\boldsymbol{a}$ and $\boldsymbol{b}$.

In the last two points by 'is in the plane defined by $\boldsymbol{a}$ and $\boldsymbol{b}$' we mean 'is a linear combination of $\boldsymbol{a}$ and $\boldsymbol{b}$'. Exactly what the coefficients are is given by the next theorem:

Theorem 1: Vector triple product

Let $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$. Then

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

and

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

Proof. The vector $\boldsymbol{b} \times \boldsymbol{c}$ is perpendicular to the plane defined by $\boldsymbol{b}$ and $\boldsymbol{c}$. This means that $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ is perpendicular to the perpendicular to the plane defined by $\boldsymbol{b}$ and $\boldsymbol{c}$. This means that $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ is in

the plane defined by $\boldsymbol{b}$ and $\boldsymbol{c}$. Therefore there exists $\beta, \gamma \in \mathbb{R}$ such that

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \beta\boldsymbol{b} + \gamma\boldsymbol{c}.$$

Taking a scalar product with $\boldsymbol{a}$ the left hand side of this becomes zero as $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ is perpendicular to $\boldsymbol{a}$. The right hand side becomes

$$0 = \beta(\boldsymbol{a} \cdot \boldsymbol{b}) + \gamma(\boldsymbol{a} \cdot \boldsymbol{c}) \implies \beta(\boldsymbol{a} \cdot \boldsymbol{b}) = -\gamma(\boldsymbol{a} \cdot \boldsymbol{c})$$

Thus there exists $\alpha \in \mathbb{R}$ such that $\beta = \alpha(\boldsymbol{a} \cdot \boldsymbol{c})$ and also $\gamma = -\alpha(\boldsymbol{a} \cdot \boldsymbol{b})$. Therefore

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \alpha[(\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}].$$

We can determine the value of α by considering the specific case where $\boldsymbol{a} = a\boldsymbol{e_y}$, $\boldsymbol{b} = b\boldsymbol{e_x}$, and $\boldsymbol{c} = c\boldsymbol{e_y}$. Then

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = abc\boldsymbol{e_y} \times (\boldsymbol{e_x} \times \boldsymbol{e_y}) = abc\boldsymbol{e_y} \times \boldsymbol{e_z} = abc\boldsymbol{e_x}$$

and

$$\alpha[(\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}] = \alpha abc[(\boldsymbol{e_y} \cdot \boldsymbol{e_y})\boldsymbol{e_x} - (\boldsymbol{e_x} \cdot \boldsymbol{e_y})\boldsymbol{e_y}] = \alpha abc\boldsymbol{e_x}.$$

Comparing these we see that $\alpha = 1$ and hence

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

Finally

$$\begin{aligned}
(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} &= -\boldsymbol{c} \times (\boldsymbol{a} \times \boldsymbol{b}) \\
&= (\boldsymbol{c} \cdot \boldsymbol{a})\boldsymbol{b} - (\boldsymbol{c} \cdot \boldsymbol{b})\boldsymbol{a} \\
&= (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{a}.
\end{aligned}$$

$\square$

Corollary 1: Kronecker delta and Levi–Civita symbol relation

The following holds for all $i, j, l, m \in \{1, 2, 3\}$:

$$\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}.$$

Proof. Let $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$, then

$$\begin{aligned}
[\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})]_i &= \varepsilon_{ijk}a_j(\boldsymbol{b} \times \boldsymbol{c})_k \\
&= \varepsilon_{ijk}a_j(\varepsilon_{klm}b_lc_m) \\
&= \varepsilon_{ijk}\varepsilon_{klm}a_jb_lc_m.
\end{aligned}$$

Also by theorem 1 we know that

$$\begin{aligned}
[\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})]_i &= [(\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}]_i \\
&= a_jc_jb_i - a_jb_jc_i \\
&= a_j\delta_{jm}c_mb_i - a_jb_j\delta_{im}c_m \\
&= a_j\delta_{jm}c_m\delta_{il}b_l - a_j\delta_{jl}b_l\delta_{im}c_m \\
&= (\delta_{il}\delta_{jm} = \delta_{im}\delta_{jl})a_jb_lc_m.
\end{aligned}$$

Comparing these two expressions we see that

$$\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}.$$

$\square$

2.2   Fields

> **Definition 5: Scalar and vector fields**
>
> A **field** is any function of position, $f \colon \mathbb{R}^3 \to S$. We classify fields based on the type of their output. For example if $S = \mathbb{R}$ then we say f is a scalar field but if $S = \mathbb{R}^3$ then we say f is a vector field. That is a scalar (vector) field assigns a scalar (vector) to every point in space.

Fields are incredibly common in physics. For example the temperature at position $\boldsymbol{r}$ is a field, $T(\boldsymbol{r})$. A common example of a vector field is the electric field, $\boldsymbol{E}(\boldsymbol{r})$.

2.3 Vector Calculus

> **Definition 6: Gradient**
>
> Let $\varphi \colon \mathbb{R}^3 \to \mathbb{R}$ be a scalar field. Then the **gradient** of φ is the vector field, $\nabla\varphi \colon \mathbb{R}^3 \to \mathbb{R}^3$, defined by
> $$\nabla\varphi = \boldsymbol{e_x}\frac{\partial\varphi}{\partial x} + \boldsymbol{e_y}\frac{\partial\varphi}{\partial y} + \boldsymbol{e_z}\frac{\partial\varphi}{\partial z} = \boldsymbol{e_i}\partial_i\varphi.$$

We can view ∇ as a vector with components $(\nabla)_i = \partial_i$. This makes it natural to consider scalar and vector products of ∇ with other vectors.

> **Definition 7: Divergence**
>
> Let $\boldsymbol{a} \colon \mathbb{R}^3 \to \mathbb{R}^3$ be a vector field. Then the **divergence** of $\boldsymbol{a}$ is the scalar field, $\nabla \cdot \boldsymbol{a} \colon \mathbb{R}^3 \to \mathbb{R}$, defined by
> $$\nabla \cdot \boldsymbol{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} = \partial_i a_i.$$

> **Definition 8: Curl**
>
> Let $\boldsymbol{a} \colon \mathbb{R}^3 \to \mathbb{R}^3$ be a vector field. Then the **curl** of $\boldsymbol{a}$ is the vector field, $\nabla \times \boldsymbol{a} \colon \mathbb{R}^3 \to \mathbb{R}^3$, defined by
> $$\nabla \times \boldsymbol{a} = \left(\frac{\partial a_y}{\partial z} - \frac{\partial a_z}{\partial y}\right)\boldsymbol{e_x} + \left(\frac{\partial a_z}{\partial x} - \frac{\partial a_x}{\partial z}\right)\boldsymbol{e_y} + \left(\frac{\partial a_x}{\partial y} - \frac{\partial a_y}{\partial x}\right)\boldsymbol{e_z}$$
> $$= \begin{vmatrix} \boldsymbol{e_x} & \boldsymbol{e_y} & \boldsymbol{e_z} \\ \partial_x & \partial_y & \partial_z \\ a_x & a_y & a_z \end{vmatrix}$$
> $$= \varepsilon_{ijk}\boldsymbol{e_i}\partial_j a_k$$
>
> or in terms of its components
> $$(\nabla \times \boldsymbol{a})_i = \varepsilon_{ijk}\partial_j a_k.$$

We can also combine ∇ with itself, for example $\nabla \cdot \nabla = \nabla^2$.

> **Definition 9: Laplacian**
>
> Let $\varphi \colon \mathbb{R}^3 \to \mathbb{R}$ be a scalar field. Then the **Laplacian** of φ is the scalar field, $\nabla^2\varphi \colon \mathbb{R}^3 \to \mathbb{R}$, defined by
> $$\nabla^2\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} = \partial_i\partial_i\varphi.$$
>
> If $\boldsymbol{a} \colon \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field then the **Laplacian** of $\boldsymbol{a}$ is the vector field, $\nabla^2\boldsymbol{a} \colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by
> $$\nabla^2\boldsymbol{a} = \boldsymbol{e_i}\nabla^2 a_i$$
>
> Here $\nabla^2 a_i$ is the Laplacian of a scalar field as defined above.

Example 2.1. *What is* ∇r?

We will work out the components:

$$
\begin{aligned}
(\nabla r)_i &= \partial_i r \\
&= \partial_i (x_1^2 + x_2^2 + x_3^2)^{1/2} \\
&= 2x_i \frac{1}{2} (x_1^2 + x_2^2 + x_3^2)^{-1/2} \\
&= \frac{x_i}{r}.
\end{aligned}
$$

Therefore

$$
\nabla r = \boldsymbol{e_i} \partial_i r = \frac{x_i}{r} \boldsymbol{e_i} = \frac{\boldsymbol{r}}{r} = \hat{\boldsymbol{r}}.
$$

Find $\nabla \times (\nabla \times \boldsymbol{a})$.

Again we work out the components:

$$
\begin{aligned}
[\nabla \times (\nabla \times \boldsymbol{a})]_i &= \varepsilon_{ijk} \partial_j (\nabla \times \boldsymbol{a})_k \\
&= \varepsilon_{ijk} \varepsilon_{klm} \partial_j \partial_l a_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) \partial_j \partial_l a_m \\
&= \delta_{il}\delta_{jm} \partial_j \partial_l a_m - \delta_{im}\delta_{jl} \partial_j \partial_l a_m \\
&= \partial_j \partial_i a_j - \partial_j \partial_j a_i \\
&= \partial_i (\nabla \cdot \boldsymbol{a}) - (\nabla \cdot \nabla) a_i \\
&= [\nabla(\nabla \cdot \boldsymbol{a}) - \nabla^2 \boldsymbol{a}]_i.
\end{aligned}
$$

Hence

$$
\nabla \times (\nabla \times \boldsymbol{a}) = \nabla(\nabla \cdot \boldsymbol{a}) - \nabla^2 \boldsymbol{a}.
$$

This gives us a useful identity for the Laplacian of a vector field, $\boldsymbol{a}$:

$$
\nabla^2 \boldsymbol{a} = \nabla(\nabla \cdot \boldsymbol{a}) - \nabla \times (\nabla \times \boldsymbol{a}).
$$

What is $\nabla \times (\boldsymbol{a} \times \boldsymbol{b})$?

Working out the components:

$$
\begin{aligned}
[\nabla \times (\boldsymbol{a} \times \boldsymbol{b})]_i &= \varepsilon_{ijk} \partial_j (\boldsymbol{a} \times \boldsymbol{b})_k \\
&= \varepsilon_{ijk} \varepsilon_{klm} \partial_j a_l b_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) \partial_j a_l b_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) a_l \partial_j b_m + (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) b_m \partial_j a_l \\
&= \delta_{il}\delta_{jm} a_l \partial_j b_m - \delta_{im}\delta_{jl} a_l \partial_j b_m + \delta_{il}\delta_{jm} b_m \partial_j a_l - \delta_{im}\delta_{jl} b_m \partial_j a_l \\
&= a_i \partial_j b_j - a_j \partial_j b_i + b_j \partial_j a_i - b_i \partial_j a_j \\
&= a_i (\nabla \cdot \boldsymbol{b}) - (\boldsymbol{a} \cdot \nabla) b_i + (\boldsymbol{b} \cdot \nabla) a_i - b_i (\nabla \cdot \boldsymbol{a}) \\
&= [\boldsymbol{a}(\nabla \cdot \boldsymbol{b}) - (\boldsymbol{a} \cdot \nabla)\boldsymbol{b} + (\boldsymbol{b} \cdot \nabla)\boldsymbol{a} - \boldsymbol{b}(\nabla \cdot \boldsymbol{a})]_i.
\end{aligned}
$$

Hence

$$
\nabla \times (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a}(\nabla \cdot \boldsymbol{b}) + (\boldsymbol{b} \cdot \nabla)\boldsymbol{a} - \boldsymbol{b}(\nabla \cdot \boldsymbol{a}) - (\boldsymbol{a} \cdot \nabla)\boldsymbol{b}.
$$

3   Matrices

Definition 10: Matrix

A real $n \times m$ **matrix**, A, is a rectangular array of real numbers with n rows and m columns.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m-1,1} & a_{m-1,2} & \dots & a_{m-1,n-1} & a_{m-1,n} \\ a_{m1} & a_m2 & \dots & a_{m,n-1} & a_{mn} \end{pmatrix} = (a_{ij})$$

Notation

In these notes I will use the notation $\mathcal{M}_{n \times m}(S)$ to denote the set of $n \times m$ matrices with entries from the set S. This set will almost always be the real numbers and most of the time we consider square 3×3 matrices so most matrices will be in $\mathcal{M}_{3 \times 3}(\mathbb{R})$.

Notation

Matrices will normally be written with capital letter, such as $A \in \mathcal{M}_{n \times m}(\mathbb{R})$. The elements will be written with the same letter but lowercase or the same letter uppercase with indices, such as a_{ij} or A_{ij}. If we want to define A by its components, a_{ij}, then we will write $A = (a_{ij})$.

Definition 11: Addition and scalar multiplication of matrices

Let $A, B \in \mathcal{M}_{n \times m}(\mathbb{R})$. Then the **sum**, $A + B$, is the $n \times m$ matrix with elements

$$(A + B)_{ij} = a_{ij} + b_{ij}.$$

Subtraction is similarly defined as the $n \times m$ matrix with elements

$$(A - B)_{ij} = a_{ij} - b_{ij}.$$

So we see that for addition and subtraction we simply add or subtract matching entries. Addition and subtraction aren't defined if A and B have different numbers of rows or columns. **Scalar multiplication** by $\lambda \in \mathbb{R}$ is defined as

$$(\lambda A)_{ij} = \lambda a_{ij}.$$

So for scalar multiplication we simply multiply all entries by the scalar.

Definition 12: Square matrix

A **square matrix** is simply a a matrix in $\mathcal{M}_{n \times n}(\mathbb{R})$. Of particular interest is the set of 3×3 matrices, $\mathcal{M}_{3 \times 3}(\mathbb{R})$ as these are very common to describe transformations in the three-dimensional space we live in.

Definition 13: Identity matrix

The **identity matrix** or **unit matrix** in n-dimensions is the matrix

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = (\delta_{ij}).$$

In particular the three-dimensional identity matrix is

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

Notation

The identity matrix is variously denoted as I, $\mathbb{I}$, id, 1, and $\mathbb{1}$. Sometimes a subscript, n, or $n \times n$, is attached to denote the dimension of the matrix, such as I_3 or $I_{3 \times 3}$ for the 3 identity matrix.

Notation

A **diagonal matrix** is a square matrix where $a_{ij} = 0$ if $i \neq j$. These matrices are common enough that we define special notation, diag. If $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ has $a_{ii} = a_i$ then

$$\operatorname{diag}(a_1, a_2, \ldots, a_n) = \operatorname{diag}(\boldsymbol{a}) = \begin{pmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n \end{pmatrix}.$$

For example the three-dimensional identity matrix is

$$I = \operatorname{diag}(1, 1, 1).$$

Definition 14: Trace

The **trace** of a square matrix, $A \in \mathcal{M}_{n \times n}(\mathbb{R})$, is defined as

$$\operatorname{Tr} A = a_{ii} = \sum_{i=1}^{n} a_{ii}.$$

For example $\operatorname{Tr} I = \delta_{ii} = 3$.

Notation

The trace of A is sometimes denoted trace$A$, $\operatorname{tr} A$, or $\operatorname{Tr} A$.

Definition 15: Transpose

The **transpose** of $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ is $A^{\mathsf{T}} \in \mathcal{M}_{m \times n}(\mathbb{R})$ defined by

$$(A^{\mathsf{T}})_{ij} = A_{ji}.$$

For example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

If $A^{\mathsf{T}} = A$ we say that A is **symmetric**. If $A^{\mathsf{T}} = -A$ we say that A is **antisymmetric**. For example the following are symmetric and antisymmetric respectively:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 2 & 3 \\ -2 & 4 & 5 \\ -3 & -5 & 6 \end{pmatrix}.$$

> **Notation 1**
>
> The transpose of A is sometimes denoted A^T, A^{T}, A^{tr}, $^{\mathrm{T}}A$, or A^{T}.

Notice that $\mathrm{Tr}\!\left(A^{\mathsf{T}}\right) = \mathrm{Tr}(A)$ since the main diagonal is not changed by transposition.

3.1 Matrix Product

> **Definition 16: Matrix product**
>
> For $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ and $B \in \mathcal{M}_{m \times p}(\mathbb{R})$ the product $C = AB$ is defined as the matrix $C \in \mathcal{M}_{n \times p}(\mathbb{R})$ with entries
> $$c_{ij} = a_{ik}b_{kj}.$$
> Note that the sum is over $k = 1, \ldots, m$. Products of matrices where the number of columns of the first matrix is not the same as the number of rows of the second matrix are not defined.

> **Theorem 2: Matrix product associativity and distributivity**
>
> If A, B, and C are matrices such that the products $A(BC)$ and $(AB)C$ are defined then the product is associative:
> $$A(BC) = (AB)C,$$
> meaning we can unambiguously write ABC. If A, B, and C are matrices such that the sum $B+C$ and the products AB and AC are defined then the matrix product is distributive over addition:
> $$A(B + C) = AB + AC.$$

Proof. Assuming that the necessary products exist we have
$$\begin{aligned}
[A(BC)]_{ij} &= a_{ik}[BC]_{kj} \\
&= a_{ik}(b_{kl}c_{lj}) \\
&= (a_{ik}b_{kl})c_{lj} \\
&= [AB]_{il}c_{lj} \\
&= [(AB)C]_{ij}
\end{aligned}$$

Assuming that multiplication of the elements is associative this shows $A(BC) = (AB)C$.

Again assuming that A, B, and C are such that all operations are defined we have
$$\begin{aligned}
[A(B + C)]_{ij} &= a_{ik}[B + C]_{kj} \\
&= a_{ik}(b_{kj} + c_{kj}) \\
&= a_{ik}b_{kj} + a_{ij}c_{kj} \\
&= [AB + AC]_{ij}.
\end{aligned}$$

Assuming that multiplication of the elements distributes over addition this shows that $A(B + C) = AB + AC$. $\qquad\square$

Matrix multiplication is not commutative. For example if $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ and $B \in \mathcal{M}_{m \times p}(\mathbb{R})$ and $n \neq p$ then AB defined and BA isn't. Even in the case that $A, B \in \mathcal{M}_{n \times n}(\mathbb{R})$ it isn't always the case that $AB = BA$. For example
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix} \neq \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

> **Lemma 1: Transpose of a product**
>
> Let A and B be matrices such that AB is defined. Then $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$.

Proof. Let A and B be matrices such that AB is defined. Then

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

Assuming that multiplication of elements is commutative this shows that $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$. □

3.2   Determinants

> **Definition 17: Determinant**
>
> The **determinant** of a 3×3 matrix, A, is defined as
>
> $$\det A = |A| = \|A\| = \varepsilon_{ijk}a_{1i}a_{2j}a_{3k}.$$

The determinant can equivalently be defined by a recursive definition. First we define the determinant of a 2×2 matrix, B, to be

$$
\det B = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = b_{11}b_{22} - b_{21}b_{12}
$$

and then

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

The determinant generalises to $n \times n$ matrices fairly easily. If $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ then

$$\det A = \varepsilon_{i_1 i_2 \cdots i_n} a_{1i_1} a_{2i_2} \cdots a_{ni_n}$$

where $\varepsilon_{i_1 i_2 \cdots i_n}$ is the n-dimensional Levi–Civita symbol defined by

$$
\varepsilon_{i_1 i_2 \cdots i_n} = \begin{cases} +1, & (i_1, i_2, \ldots, i_n) \text{ is an even permutation of } (1, 2, \ldots, n), \\ -1, & (i_1, i_2, \ldots, i_n) \text{ is an odd permutation of } (1, 2, \ldots, n), \\ 0, & \text{otherwise.} \end{cases}
$$

The recursive definition of the determinant can also be generalised to higher dimensions. First we define the minor, $M^{(i,j)}$, of the matrix $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ to be the $(n-1) \times (n-1)$ matrix we get if we delete the ith row and jth column of A. Then

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det M^{(i,j)} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det M^{(i,j)}.$$

Notice that we are free to choose any value of j in the first sum or any value of i in the second sum. We can often use this to our benefit by expanding the sum in a row or column with a lot of zeros so that lots of terms vanish simplifying the calculation. Usually it is easier to calculate a specific determinant with the recursive definition and to prove something involving determinants with the ε_{ijk} definition.

4 Determinant Properties

In the last section we saw that for $A \in \mathcal{M}_{3\times 3}(\mathbb{R})$ the determinant is defined as

$$\det A = \varepsilon_{lmn} a_{1l} a_{2m} a_{3n}.$$

Now define

$$X_{ijk} = \varepsilon_{lmn} a_{il} a_{jm} a_{kn}.$$

Then we have

$$\begin{aligned}
X_{jik} &= \varepsilon_{lmn} a_{jl} a_{im} a_{kn} & l \leftrightarrow m \\
&= \varepsilon_{mln} a_{jm} a_{il} a_{kn} & \varepsilon_{mln} = -\varepsilon_{lmn} \\
&= -\varepsilon_{lmn} a_{il} a_{jm} a_{kn} \\
&= -X_{ijk}.
\end{aligned}$$

Similarly we can show that $X_{kji} = X_{ikj} = -X_{ijk}$ and $X_{jki} = X_{kij} = X_{ijk}$. We see that the symmetry of X_{ijk} is the same as the symmetry of ε_{ijk} so we can write $X_{ijk} = c\varepsilon_{ijk}$ for some $c \in \mathbb{R}$. To find c we consider the specific case $(i, j, k) = (1, 2, 3)$ and we see that

$$X_{123} = \varepsilon_{lmn} a_{1l} a_{2m} a_{3n} = \det A.$$

This means that $X_{ijk} = \varepsilon_{ijk} \det A$, we conclude

$$\varepsilon_{ijk} \det A = \varepsilon_{lmn} a_{il} a_{jm} a_{kn}.$$

If we then use $\varepsilon_{ijk}\varepsilon_{ijk} = 6$ we see that

$$\det A = \frac{1}{3!} \varepsilon_{ijk} \varepsilon_{lmn} a_{il} a_{jm} a_{kn}.$$

This is a symmetric form for the determinant. This generalises to $B \in \mathcal{M}_{n\times n}(\mathbb{R})$ as

$$\det B = \frac{1}{n!} \varepsilon_{i_1 i_2 \ldots i_n} \varepsilon_{j_1 j_2 \ldots j_n} b_{i_1 j_1} b_{i_2 j_2} \cdots b_{i_n j_n}.$$

If we define

$$Y_{lmn} = \varepsilon_{ijk} a_{il} a_{jm} a_{kn}$$

we see by the same argument that $Y_{lmn} = \varepsilon_{lmn} Y_{123}$. Further

$$\det A = \frac{1}{3!} \varepsilon_{ijk} \varepsilon_{lmn} a_{il} a_{jm} a_{kn} = \frac{1}{3!} \varepsilon_{lmn} Y_{lmn} = \frac{1}{3!} \varepsilon_{lmn} \varepsilon_{lmn} Y_{123} = Y_{123}.$$

So we have that

$$\det A = \varepsilon_{ijk} a_{i1} a_{j2} a_{k3}.$$

What this means is that we can expand along a column as well as a row when using the recursive definition of the determinant. As well as this since $Y_{lmn} = \varepsilon_{lmn} \det A$ we see that

$$\varepsilon_{lmn} \det A = \varepsilon_{ijk} a_{il} a_{jm} a_{kn}.$$

We will now list some properties of the determinant. We won't prove these but will give examples.

- If we multiply one row (or column) of A by a scalar, $\lambda \in \mathbb{R}$, to make the matrix A' then $\det A' = \lambda \det A$. For example suppose we multiply the first row by λ so $a'_{1i} = \lambda a_{1i}$, $a'_{ij} = a_{ij}$ for $i = 2, 3$. Then

$$\det A' = \varepsilon_{lmn} a'_{1l} a'_{2m} a'_{3n} = \varepsilon_{lmn} (\lambda a_{1l}) a_{2m} a_{3n} = \lambda \det A.$$

- Multiplying an $n \times n$ matrix by $\lambda \in \mathbb{R}$ multiplies the determinant by λ^n. This follows by applying the previous point to each row individually multiplied by λ.

- Adding a multiple of one row to another row (or a multiple of a column to another column) does not change the value of the determinant. For example if we add λ times the second row to the first row to get A' defined by $a'_{1i} = a_{1i} + \lambda a_{2i}$ and $a'_{ij} = a_{ij}$ for $i = 2, 3$ then

$$\det A' = \varepsilon_{lmn} a'_{1l} a'_{2m} a'_{3n} = \varepsilon_{lmn}(a_{1l} + \lambda a_{2l}) a_{2m} a_{3n} = \varepsilon_{lmn} a_{1l} a_{2m} a_{3n} + \lambda \varepsilon_{lmn} a_{2l} a_{2m} a_{3n}.$$

Considering only the last term we see that

$$\begin{aligned}
\varepsilon_{lmn} a_{2l} a_{2m} a_{3n} &= \varepsilon_{11n} a_{21} a_{21} a_{3n} + \varepsilon_{12n} a_{21} a_{22} a_{3n} + \varepsilon_{21n} a_{22} a_{12} a_{3n} + \varepsilon_{22n} a_{22} a_{22} a_{3n} \\
&= 0 + \varepsilon_{12n} a_{21} a_{22} a_{3n} - \varepsilon_{12n} a_{22} a_{21} a_{3n} + 0 \\
&= 0
\end{aligned}$$

So

$$\det A' = \varepsilon_{lmn} a_{1l} a_{2m} a_{3n} + \lambda \varepsilon_{lmn} a_{2l} a_{2m} a_{3n} = \varepsilon_{lmn} a_{1l} a_{2m} a_{3n} + 0 = \det A.$$

- Interchanging any two adjacent rows (or columns) of A to get A' gives $\det A' = -\det A$. For example if we swap the first and second row so $a'_{1j} = a_{2j}$, $a'_{2j} = a_{1j}$, and $a'_{3j} = a_{3j}$ then

$$\det A' = \varepsilon_{lmn} a'_{1l} a'_{2m} a'_{3n} = \varepsilon_{lmn} a_{2l} a_{1m} a_{3n} = -\varepsilon_{mln} a_{1m} a_{2l} a_{3n} \stackrel{l \leftrightarrow m}{=} -\varepsilon_{lmn} a_{1l} a_{2m} a_{3n} = -\det A.$$

- Finally

$$\varepsilon_{ijk} \varepsilon_{lmn} \det A = \begin{vmatrix} a_{il} & a_{im} & a_{in} \\ a_{jl} & a_{jm} & a_{jn} \\ a_{kl} & a_{km} & a_{kn} \end{vmatrix}. \tag{4.1}$$

This can be seen from the definition of $\det A$ and permuting rows and columns gives the same sign as the equivalent permutations in the indices of $\varepsilon_{ijk} \varepsilon_{lmn}$.

To show the above properties for columns instead of rows the method is the same but use $\det A = \varepsilon_{ijk} a_{i1} a_{j2} a_{k3}$ instead.

4.1   Linear Equations

Suppose we have a system of linear equations of the form

$$\begin{aligned}
a_{11} x_1 + a_{12} x_2 + a_{13} x_3 &= y_1 \\
a_{21} x_1 + a_{22} x_2 + a_{23} x_3 &= y_2 \\
a_{31} x_1 + a_{32} x_2 + a_{33} x_3 &= y_3.
\end{aligned}$$

We can write these more compactly as

$$a_{ij} x_j = y_j.$$

This makes it obvious that we can write these equations with a matrix and some vectors:

$$A\boldsymbol{x} = \boldsymbol{y}.$$

Strictly $\boldsymbol{x}$ and $\boldsymbol{y}$ are 3×1 matrices but nothing changes if we think of them as column vectors in $\mathbb{R}^3$. See appendix A. The solution to these equations is then

$$\boldsymbol{x} = A^{-1} \boldsymbol{y}, \qquad \text{or} \qquad x_i = (A^{-1})_{ij} y_j.$$

The question we now ask is what is A^{-1}? There are a few ways to calculate A^{-1}:

- If we solve the equations then we can easily find A^{-1}, although this sort of defeats the point of trying to find A^{-1} in the first place.

- It can be shown that if

$$(A^{-1})_{ij} = \frac{1}{2!} \frac{1}{\det A} \varepsilon_{imn} \varepsilon_{jpq} a_{pm} a_{qn}$$

then $AA^{-1} = A^{-1}A = I$.

- It can be shown that
$$A^{-1} = \frac{C^{\mathsf{T}}}{\det A}$$
satisfies these equations where C is the **co-factor** matrix of A which is defined by
$$c_{ij} = (-1)^{i+j} \det M^{(i,j)}$$
where $M^{(i,j)}$ is the **minor** matrix of A which is the matrix we get when we take A and remove the ith row and jth column.

Notice A^{-1} only exists if $\det A \neq 0$. These results generalise to $A \in \mathcal{M}_{n \times n}(\mathbb{R})$.

Calculating inverses is not particularly difficult but there are a lot of steps. For this reason it is often left to computers. Consider the case where
$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

Then
$$C = \begin{pmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ e & f \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{pmatrix}.$$

So A^{-1}, if it exists, is given by
$$A^{-1} = \frac{1}{a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}} \begin{pmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ \begin{vmatrix} d & e \\ g & h \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{pmatrix}.$$

$$A^{-1} = \frac{1}{aei - afh - bdi + bfg + cdh - ceg} \begin{pmatrix} ei - fh & -bi + ch & bf - ce \\ -di + fg & ai - cg & -af + cd \\ dh - eg & -ah + bg & ae - bd \end{pmatrix}.$$

> **Notation**
>
> The set of all invertible $n \times n$ matrices is is called the **general linear group** and is denoted
> $$\mathrm{GL}(n) = \mathrm{GL}(n, \mathbb{R}) = \mathrm{GL}_n(\mathbb{R}) = \{A \in \mathcal{M}_{n \times n}(\mathbb{R}) \mid \det A \neq 0\} \subseteq \mathcal{M}_{n \times n}(\mathbb{R}).$$

As the name suggests $\mathrm{GL}(n)$ is a group under matrix multiplication (see appendix B). The reason for its name is if we view matrices as linear transformations on $\mathbb{R}^n$ then this is the most general set of such transformations that forms a group.

4.2 More Determinant Properties

> **Lemma 2: Determinant of the transpose**
>
> Let $A \in \mathcal{M}_{n \times n}(\mathbb{R})$. Then $\det(A^{\mathsf{T}}) = \det A$.

Proof.
$$\begin{aligned} \det(A^{\mathsf{T}}) &= \varepsilon_{lmn}(A^{\mathsf{T}})_{1l}(A^{\mathsf{T}})_{2m}(A^{\mathsf{T}})_{3n} \\ &= \varepsilon_{lmn} a_{l1} a_{m2} a_{n3} \\ &= \det A \end{aligned}$$

$\qquad\square$

> ### Lemma 3: Determinant of a product
>
> The determinant of a product is the product of the determinants.

Proof. Let $A, B, C \in \mathcal{M}_{n \times n}(\mathbb{R})$ such that $C = AB$. Then $c_{ij} = a_{ik}b_{kj}$ so

$$
\begin{aligned}
\det C &= \varepsilon_{ijk}c_{1i}c_{2j}c_{3k} \\
&= \varepsilon_{ijk}a_{1l}b_{li}a_{2m}b_{mj}a_{3n}b_{nk} \\
&= [\varepsilon_{ijk}b_{li}b_{mj}b_{nk}]a_{1l}a_{2m}a_{3n} \\
&= \varepsilon_{lmn}\det(B)a_{1l}a_{2m}a_{3n} \\
&= \det(A)\det(B).
\end{aligned}
$$

$\square$

> ### Lemma 4: Product of two Levi–Civita symbols
>
> The following holds for all $i, j, k, l, m, n \in \{1, 2, 3\}$:
>
> $$
> \varepsilon_{ijk}\varepsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}.
> $$

Proof. Starting from equation 4.1 we know that

$$
\varepsilon_{ijk}\varepsilon_{lmn}\det A = \begin{vmatrix} a_{il} & a_{im} & a_{in} \\ a_{jl} & a_{jm} & a_{jn} \\ a_{kl} & a_{km} & a_{kn} \end{vmatrix}.
$$

Choosing $A = I$ and therefore $a_{ij} = \delta_{ij}$ since $\det I = 1$ we have

$$
\varepsilon_{ijk}\varepsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}.
$$

$\square$

This gives us $3^6 = 729$ identities. Of these 18 are of the form $1 = 1$, 18 are of the form $-1 = -1$ and the other 693 are of the form $0 = 0$.

> ### Corollary 2: Kronecker delta and Levi–Civita symbol relation (again)
>
> The following holds for all $i, j, l, m \in \{1, 2, 3\}$:
>
> $$
> \varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}.
> $$

Proof. Consider $\varepsilon_{ijk}\varepsilon_{klm}$. Using theorem **??** we see that

$$
\begin{aligned}
\varepsilon_{ijk}\varepsilon_{klm} &= \begin{vmatrix} \delta_{ik} & \delta_{il} & \delta_{im} \\ \delta_{jk} & \delta_{jl} & \delta_{jm} \\ \delta_{kk} & \delta_{kl} & \delta_{km} \end{vmatrix} \\
&= \delta_{ik}(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) - \delta_{il}(\delta_{jk}\delta_{km} - \delta_{kk}\delta_{jm}) + \delta_{im}(\delta_{jk}\delta_{kl} - \delta_{kk}\delta_{jl}) \\
&= \delta_{ik}(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) - \delta_{il}(\delta_{jk}\delta_{km} - 3\delta_{jm}) + \delta_{im}(\delta_{jk}\delta_{kl} - 3\delta_{jl}) \\
&= \delta_{jl}\delta_{im} - \delta_{jm}\delta_{il} - \delta_{il}\delta_{jm} + 3\delta_{il}\delta_{jm} + \delta_{im}\delta_{jl} - 3\delta_{im}\delta_{jl} \\
&= \delta_{il}\delta_{jm} - \delta_{jl}\delta_{im} \\
&= \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix}.
\end{aligned}
$$

$\square$

5   Passive Rotations

5.1   Orthogonal Matrices

> **Definition 18: Orthogonal matrix**
>
> Let $A \in \mathcal{M}_{n \times n}(\mathbb{R})$. Let $\boldsymbol{a}^{(i)} = (a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbb{R}^n$. Then A is an **orthogonal matrix** if
> $$\boldsymbol{a}^{(i)} \cdot \boldsymbol{a}^{(j)} = \delta_{ij}$$

That is A is an orthogonal matrix if its rows viewed as vectors are mutually orthonormal.

> **Theorem 3: Orthogonal matrix properties**
>
> Let A be an orthogonal matrix. Then
> $$A^{\mathsf{T}} A = A A^{\mathsf{T}} = I, \qquad \det A = \pm 1, \qquad \text{and} \qquad A^{\mathsf{T}} = A^{-1}.$$

Proof. Let A be an orthogonal matrix. Then using the notation of definition 18 we have

$$
\begin{aligned}
(AA^{\mathsf{T}})_{ij} &= A_{ik}(A^{\mathsf{T}})_{kj} \\
&= a_{ik} a_{jk} \\
&= a_k^{(i)} a_k^{(j)} \\
&= \boldsymbol{a}^{(i)} \cdot \boldsymbol{a}^{(j)} \\
&= \delta_{ij}
\end{aligned}
$$

Hence $AA^{\mathsf{T}} = I$. Recall that $\det I = 1$, $\det(BC) = \det(B)\det(C)$, and $\det(B^{\mathsf{T}}) = \det B$. Hence

$$\det(AA^{\mathsf{T}}) = \det(A)\det(A^{\mathsf{T}}) = \det(A)\det(A) = [\det A]^2$$

and

$$\det(AA^{\mathsf{T}}) = \det I = 1.$$

Combining these we have

$$[\det A]^2 = 1 \implies \det A = \pm 1.$$

Since $\det A \neq 0$ we know that A^{-1} and $(A^{\mathsf{T}})^{-1}$ always exists. Using this we see

$$
\begin{aligned}
AA^{\mathsf{T}} &= I \\
A^{-1} A A^{\mathsf{T}} &= A^{-1} I \\
A^{\mathsf{T}} &= A^{-1}.
\end{aligned}
$$

Finally

$$
\begin{aligned}
AA^{\mathsf{T}} &= I \\
A^{-1} A A^{\mathsf{T}} A (A^{\mathsf{T}})^{-1} &= A^{-1} I (A^{\mathsf{T}})^{-1} \\
I &= A^{-1}(A^{\mathsf{T}})^{-1} \\
I &= (A^{\mathsf{T}} A)^{-1} \\
I &= A^{\mathsf{T}} A
\end{aligned}
$$

where we use that $(BC)^{-1} = C^{-1}B^{-1}$ and also if $B^{-1} = I$ then $B = I$. $\square$

> **Corollary 3: Orthogonal transformations preserve the inner product**
>
> $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ preserves inner products if and only if A is orthogonal.

Proof. If A preserves inner products then this means that for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ $\boldsymbol{x} \cdot \boldsymbol{y}$ will be the same as $\boldsymbol{x}' \cdot \boldsymbol{y}'$ where $\boldsymbol{x}' = A\boldsymbol{x}$ and $\boldsymbol{y}' = A\boldsymbol{y}$. To see this consider

$$\boldsymbol{x}' \cdot \boldsymbol{y}' = (A\boldsymbol{x}) \cdot (A\boldsymbol{y}) = (A\boldsymbol{x})^{\mathsf{T}}(A\boldsymbol{y}) = \boldsymbol{x}^{\mathsf{T}} A^{\mathsf{T}} A\boldsymbol{y}.$$

If we require this to be the same as $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{y}$ then we must have that $A^{\mathsf{T}} A = I$ and so A is orthogonal. The same equation shows that if A is orthogonal then $\boldsymbol{x}' \cdot \boldsymbol{y}' = \boldsymbol{x}^{\mathsf{T}} A^{\mathsf{T}} A\boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}} I\boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y}$ so all orthogonal matrices preserve inner products. $\qquad\square$

Notation

The orthogonal group is the set of all $n \times n$ matrices. It is denoted

$$\mathrm{O}(n) = \{A \in \mathrm{GL}(n) \mid AA^{\mathsf{T}} = A^{\mathsf{T}} A = I\} \subseteq \mathrm{GL}(n) \subseteq \mathcal{M}_{n \times n}(\mathbb{R}).$$

As the name suggests $\mathrm{O}(n)$ is a group under matrix multiplication. It is a subgroup of $\mathrm{GL}(n)$.

5.2   Passive Rotations

A **passive** transformation is one which changes the basis or axes without changing any vectors. That is a passive transformation gives a new basis which we can express any vector in.

Given two right handed frames of reference, S and S', with bases $\{\boldsymbol{e_i}\}$ and $\boldsymbol{e'_i}$ respectively, then the vector $\boldsymbol{a} \in \mathbb{R}^3$ can be expressed as

$$\begin{aligned}
\boldsymbol{a} &= a_1\boldsymbol{e_1} + a_2\boldsymbol{e_2} + a_3\boldsymbol{e_3} = a_i\boldsymbol{e_i} \\
&= a'_1\boldsymbol{e'_1} + a'_2\boldsymbol{e'_2} + a'_3\boldsymbol{e'_3} = a'_i\boldsymbol{e'_i}.
\end{aligned}$$

As usual the components a'_i are given by $\boldsymbol{e'_i} \cdot \boldsymbol{a}$. However if we write $\boldsymbol{a}$ in frame S then we see that

$$a'_i = \boldsymbol{e'_i} \cdot \boldsymbol{a} = \boldsymbol{e'_i} \cdot (a_j\boldsymbol{e_j}) = (\boldsymbol{e'_i} \cdot \boldsymbol{e_j})a_j = l_{ij}a_j$$

where

$$L = (l_{ij}) = (\boldsymbol{e'_i} \cdot \boldsymbol{e_j}) = \begin{pmatrix} \boldsymbol{e'_x} \cdot \boldsymbol{e_x} & \boldsymbol{e'_x} \cdot \boldsymbol{e_y} & \boldsymbol{e'_x} \cdot \boldsymbol{e_z} \\ \boldsymbol{e'_y} \cdot \boldsymbol{e_x} & \boldsymbol{e'_y} \cdot \boldsymbol{e_y} & \boldsymbol{e'_y} \cdot \boldsymbol{e_z} \\ \boldsymbol{e'_z} \cdot \boldsymbol{e_x} & \boldsymbol{e'_z} \cdot \boldsymbol{e_y} & \boldsymbol{e'_z} \cdot \boldsymbol{e_z} \end{pmatrix}.$$

We see that the components of $\boldsymbol{a}$ transform as

$$a'_i = l_{ij}a_j.$$

Where L is the transformation matrix, in this case it is a rotation matrix. $l_{ij} = \boldsymbol{e'_i} \cdot \boldsymbol{e_j} = \cos \vartheta$ is the matrix of cosines of the angle between the ith axis of S' and the jth axis of S.

It is common to consider rotations around a particular axis as these are the simplest. Consider the rotation shown in figure 5.1. It can be seen from the geometry that

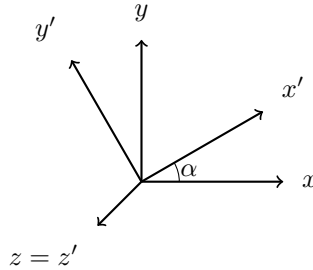Figure 5.1: A passive rotation about the z-axis through an angle α.

$$\begin{aligned}
\boldsymbol{e'_x} \cdot \boldsymbol{e_x} &= \cos \alpha, \\
\boldsymbol{e'_x} \cdot \boldsymbol{e_y} &= \cos\left(\frac{\pi}{2} - \alpha\right) = \sin \alpha,
\end{aligned}$$

$$\boldsymbol{e_y'} \cdot \boldsymbol{e_x} = \cos\left(\frac{\pi}{2} + \alpha\right) = -\sin\alpha,$$

$$\boldsymbol{e_y'} \cdot \boldsymbol{e_y} = \cos\alpha,$$

etc. and we see that a rotation through angle α about the z axis has the rotation matrix

$$L_z(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Similarly it can be shown that a rotation by angle β about the x axis and a rotation by angle γ about the y axis have the following rotation matrices:

$$L_x(\beta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{pmatrix}, \qquad \text{and} \qquad L_y(\gamma) = \begin{pmatrix} \cos\gamma & 0 & -\sin\gamma \\ 0 & 1 & 0 \\ \sin\gamma & 0 & \cos\gamma \end{pmatrix}.$$

Since $L_z(\alpha)$ represents a rotation we expect that its inverse is $L_z^{-1}(\alpha) = L_z(-\alpha)$, and indeed this is the case. By inspection using the fact that sin is odd so $\sin(-\alpha) = -\sin(\alpha)$ we can show that $L_z(\alpha)$ is an orthogonal matrix. This can be shown for a more general rotation matrix, $L = (l_{ij})$, using the fact that the length of a vector is invariant under rotation of the basis so $a^2 = \boldsymbol{a} \cdot \boldsymbol{a}$ should be the same before and after a rotation. Before a rotation we have $a^2 = a_i a_i = \delta_{ij} a_i a_j$. After a rotation each component is transformed so we have $a^2 = a_k' a_k' = (l_{ki}a_i)(l_{kj}a_j) = (l_{ki}l_{kj})a_i a_j$ so for a^2 to be invariant we require $l_{ki}l_{kj} = \delta_{ij}$. Thus $(L^\mathsf{T})_{ik}L_{kj} = (L^\mathsf{T}L)_{ij} = \delta_{ij}$ and so $L^\mathsf{T}L = I$ meaning that $L \in \mathrm{O}(3)$.

We can also consider how the basis vectors, $\{\boldsymbol{e_i}\}$, transform to the new basis vectors, $\{\boldsymbol{e_i'}\}$. Since a vector is not changed by a passive transform we have $a_i'\boldsymbol{e_i'} = a_j\boldsymbol{e_j}$. Inserting the transformation rule for a_i' we have $l_{ij}a_i\boldsymbol{e_i'} = a_j\boldsymbol{e_j}$ which must hold for all vectors, $\boldsymbol{a} \in \mathbb{R}^3$ and therefore must hold for all $a_i \in \mathbb{R}$. For this to be the case we must have $l_{ij}\boldsymbol{e_i'} = \boldsymbol{e_j}$. Multiplying through by l_{kj} this becomes $l_{kj}l_{ij}\boldsymbol{e_i'} = l_{kj}\boldsymbol{e_j}$. Now noting that $l_{kj}l_{ij} = L_{kj}(L^\mathsf{T})_{ji} = (LL^\mathsf{T})_{ki} = \delta_{ki}$ we have $\delta_{ki}\boldsymbol{e_i'} = \boldsymbol{e_{e_k'}} = l_{kj}\boldsymbol{e_j}$ which gives us the transformation rule for the basis vectors:

$$\boldsymbol{e_i'} = l_{ij}\boldsymbol{e_j}.$$

Definition 19: Proper and improper transformations

A transformation is called **proper** (or **improper**) if it preserves (or changes) the parity (i.e. left or right handed) of the system.

A pure rotation is a proper transformation as a rotation of a right handed system will again be right handed. An improper rotation also inverts the parity of the coordinate system. A proper transformation will have a positive determinant whereas an improper rotation will have a negative determinant. An improper rotation can be written as a proper rotation as well as a reflection in the origin.

Notation

The special orthogonal group, or rotation group, is the set of all proper rotations about the origin. It is denoted

$$\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det A = 1\} \subseteq \mathrm{O}(n) \subseteq \mathrm{GL}(n) \subseteq \mathcal{M}_{n \times n}(\mathbb{R}).$$

5.3   Composition of Rotations

Consider three bases, $\{\boldsymbol{e_i}\}$, $\{\boldsymbol{e_i'}\}$, and $\{\boldsymbol{e_i''}\}$, related by $\boldsymbol{e_i'} = l_{ij}^{(1)}\boldsymbol{e_j}$ and $\boldsymbol{e_i''} = l_{ij}^{(2)}\boldsymbol{e_j'}$. Then $\boldsymbol{e_i''} = l_{ij}\boldsymbol{e'i} = l_{ij}^{(2)}l_{jk}^{(1)}\boldsymbol{e_k} = (L^{(2)}L^{(1)})_{ik}\boldsymbol{e_k}$. The order is important. In general rotation A and then rotation B is not the same as rotation B then rotation A. This is reflected by the non-commutativity of matrix multiplication. The first transformation applied must be applied first and so is directly next to the object being transformed.

Any rotation about some arbitrary axis can be decomposed into three rotations around the Cartesian axis, first a rotation around the z-axis by angle α, then a rotation around the y'-axis by angle β, then a

rotation around the z''-axis by angle γ. These three angles are called the Euler angles and they encode the rotation matrix

$$L(\alpha, \beta, \gamma) = L_{z''}(\gamma) L_{y'}(\beta) L_z(\alpha)$$

$$= \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\beta\cos\alpha\cos\gamma - \sin\alpha\sin\gamma & \cos\beta\sin\alpha\cos\gamma + \cos\alpha\sin\gamma & -\sin\beta\cos\gamma \\ -\cos\beta\cos\alpha\sin\gamma - \sin\alpha\cos\gamma & -\cos\beta\sin\alpha\sin\gamma + \cos\alpha\cos\gamma & \sin\beta\sin\gamma \\ \sin\beta\cos\alpha & \sin\beta\sin\alpha & \cos\beta \end{pmatrix}.$$

6 Active Rotations

6.1 General Rotations

An active rotation keeps the basis fixed and rotates the vector. In this section we will derive the transformation matrix for a rotation about the unit vector, $\hat{\boldsymbol{n}}$, by an angle ϑ of a vector $\boldsymbol{x}$. Consider
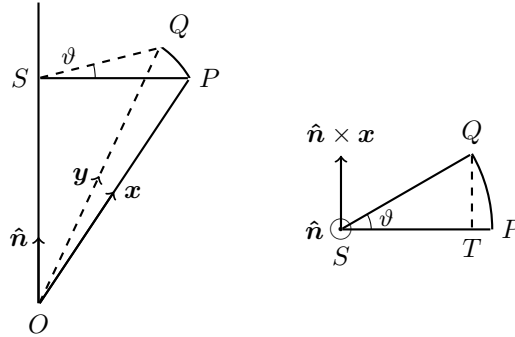
Figure 6.1: An active rotation of vector $\boldsymbol{x}$ into vector $\boldsymbol{y}$. Shown on the left is a view from the side. Shown on the right is a view from above.

figure 6.1. This shows the vector $\boldsymbol{x}$ which is the position vector of the point P. We consider a rotation about $\hat{\boldsymbol{n}}$ by angle ϑ to take the point P to the point Q which has position vector $\boldsymbol{y}$. We also define S as the projection of P or Q onto $\hat{\boldsymbol{n}}$. Then $\overrightarrow{OS}$ is the component of $\boldsymbol{x}$ in the $\hat{\boldsymbol{n}}$ direction, which is given by $\overrightarrow{OS} = (\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}$. Further let T be the projection of Q onto the vector $\overrightarrow{SP}$. This means that $\overrightarrow{TQ}$ is parallel to $\hat{\boldsymbol{n}} \times \boldsymbol{x}$. We can see that

$$\boldsymbol{y} = \overrightarrow{OS} + \overrightarrow{ST} + \overrightarrow{TQ}$$

$$= (\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}} + \underbrace{SQ\cos\vartheta}_{|\overrightarrow{ST}|}\underbrace{\frac{\boldsymbol{x} - (\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}}{SP}}_{\widehat{\overrightarrow{SP}}} + \underbrace{SQ\sin\vartheta}_{|\overrightarrow{TQ}|}\underbrace{\frac{\hat{\boldsymbol{n}} \times \boldsymbol{x}}{|\hat{\boldsymbol{n}} \times \boldsymbol{x}|}}_{\widehat{\overrightarrow{TQ}}}$$

$$= (\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}} + \cos(\vartheta)(\boldsymbol{x} - (\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}) + \sin(\vartheta)(\hat{\boldsymbol{n}} \times \boldsymbol{x}).$$

Here we have used the fact that $SP = SQ$ since the distance between the axis of rotation and a rotated point is maintained by definition in a rotation. Also $|\hat{\boldsymbol{n}} \times \boldsymbol{x}| = SP = SQ$. The result that we end up with is

$$\boldsymbol{y} = \boldsymbol{x}\cos\vartheta + (1 - \cos\vartheta)(\hat{\boldsymbol{n}} \cdot \boldsymbol{x})\hat{\boldsymbol{n}} + (\hat{\boldsymbol{n}} \times \boldsymbol{x})\sin\vartheta.$$

This can be written in index notation as

$$y_i = x_i\cos\vartheta + (1 - \cos\vartheta)n_j x_j n_i + \varepsilon_{ikj}n_k x_j \sin\vartheta$$

note the unusual ordering, (i, k, j), for the Levi–Civita symbol which is to match conventions later. We can rewrite this as

$$y_i = R_{ij}(\vartheta, \hat{\boldsymbol{n}}) = x_j$$

where

$$R_{ij}(\vartheta, \hat{\boldsymbol{n}}) = \delta_{ij}\cos\vartheta + (1 - \cos\vartheta)n_i n_j - \varepsilon_{ijk}n_k \sin\vartheta.$$

> **Definition 20: Rotation matrix**
>
> An active rotation of angle ϑ about the axis $\hat{\boldsymbol{n}}$ can be represented by the **rotation matrix** $R \in \mathcal{M}_{3\times3}(\mathbb{R})$ with components
>
> $$R_{ij} = \delta_{ij} \cos\vartheta + (1 - \cos\vartheta) n_i n_j - \varepsilon_{ijk} n_k \sin\vartheta.$$

> **Theorem 4: Rotation matrices are orthogonal**
>
> Rotation matrices are orthogonal.

Proof 2. A rotation preserves lengths. Since we can define the length of a vector, $\boldsymbol{x}$, to be $\sqrt{\boldsymbol{x}\cdot\boldsymbol{x}}$ this means a rotation preserves inner products. Therefore by corollary 3 rotations are orthogonal transformations. $\qquad\square$

> **Lemma 5: Rotation matrix properties**
>
> If R is a rotation matrix then
>
> $$\operatorname{Tr} R = 1 + 2\cos\vartheta, \qquad \text{and} \qquad n_k \sin\vartheta = -\frac{1}{2}\varepsilon_{ijk} R_{ij}.$$

Proof. Let R be a rotation matrix.
$$\delta_{ij} R_{ij} = R_{ii} = \operatorname{Tr} R$$

Also

$$\begin{aligned}
\delta_{ij} R_{ij} &= \delta_{ij}[\delta_{ij}\cos\vartheta + (1-\cos\vartheta)n_i n_j - \varepsilon_{ijk} n_k \sin\vartheta] \\
&= \delta_{ii}\cos\vartheta + (1-\cos\vartheta)\underbrace{n_i n_i}_{=\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}}=1} - \underbrace{\delta_{ij}\varepsilon_{ijk}}_{=\varepsilon_{iik}=0} n_k \sin\vartheta \\
&= 3\cos\vartheta + 1 - \cos\vartheta \\
&= 1 + 2\cos\vartheta.
\end{aligned}$$

Now consider

$$\begin{aligned}
\varepsilon_{ijk} R_{ij} &= \varepsilon_{ijk}[\delta_{ij}\cos\vartheta + (1-\cos\vartheta)n_i n_j - \varepsilon_{ijl} n_l \sin\vartheta] \\
&= \underbrace{\varepsilon_{ijk}\delta_{ij}}_{=\varepsilon_{iik}=0}\cos\vartheta + (1-\cos\vartheta)\underbrace{\varepsilon_{ijk}n_i n_j}_{(\hat{\boldsymbol{n}}\times\hat{\boldsymbol{n}})_k=0} - \underbrace{\varepsilon_{ijk}\varepsilon_{ijl}}_{=2\delta_{lk}} n_l \sin\vartheta \\
&= -2\delta_{lk} n_l \sin\vartheta \\
&= -2 n_k \sin\vartheta.
\end{aligned}$$

Rearranging we have

$$-\frac{1}{2}\varepsilon_{ijk} R_{ij} = n_k \sin\vartheta.$$

$\qquad\square$

We can use these two identities to find $\hat{\boldsymbol{n}}$ and ϑ given a rotation matrix, R.

As with passive transformation the product of two rotations is a rotation and the order matters. For example the vector given by rotating $\boldsymbol{x}$ by rotation R and then rotation S is $SR\boldsymbol{x}$ which is not necessarily equal to $RS\boldsymbol{x}$.

6.2 Rotational Symmetry

If we consider an infinitesimal rotation by $\delta\vartheta$ and we allow $\delta\vartheta \to 0$ then we can use the Taylor series,

$$\sin(\delta\vartheta) = \delta\vartheta + \mathcal{O}\big(\delta\vartheta^3\big) \qquad \text{and} \qquad \cos(\delta\vartheta) = 1 + \mathcal{O}\big(\delta\vartheta^2\big),$$

to write

$$R_{ij} = \delta_{ij} - \varepsilon_{ijk}n_k\delta\vartheta.$$

Going back to vectors from components we have

$$y_i = R_{ij}x_j = \delta_{ij}x_j - \varepsilon_{ijk}n_k x_j\delta\vartheta = x_i + \varepsilon_{ikj}n_k x_j\delta\vartheta$$

which gives

$$\boldsymbol{y} = R\boldsymbol{x} = \boldsymbol{x} + (\hat{\boldsymbol{n}} \times \boldsymbol{x})\delta\vartheta.$$

We can write the component form of the infinitesimal rotation matrices as

$$R_{ij} = \delta_{ij} - \varepsilon_{ijk}n_k\delta\vartheta = \delta_{ij} - in_k(T_k)_{ij}\delta\vartheta$$

Where we have defined three matrices T_k with components $(T_k)_{ij} = -i\varepsilon_{ijk}$. These matrices in full are

$$T_1 = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad T_2 = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad T_3 = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We say that the matrices T_i are the generators of rotations in the sense that any infinitesimal rotation matrix can be made from T_i and any rotation can be made from an infinite number of infinitesimal rotation matrices. We will see how shortly.

If we commit a small abuse of notation and define a 'vector' $\boldsymbol{T} = (T_1, T_2, T_3)$ then we can write R as

$$R(\delta\vartheta, \hat{\boldsymbol{n}}) = I - i\delta\vartheta(\hat{\boldsymbol{n}} \cdot \boldsymbol{T}).$$

Note that

$$\hat{\boldsymbol{n}} \cdot \boldsymbol{T} = n_1 T_1 + n_2 T_2 + n_3 T_3 \in \mathcal{M}_{3\times3}(\mathbb{C}).$$

> **Definition 21: Commutator**
>
> Let $A, B \in \mathcal{M}_{n\times n}(\mathbb{C})$ then we define the **commutator**,
>
> $$[\cdot, \cdot] \colon \mathcal{M}_{n\times n}(\mathbb{C}) \times \mathcal{M}_{n\times n}(\mathbb{C}) \to \mathcal{M}_{n\times n}(\mathbb{C})$$
>
> as
>
> $$[A, B] = AB - BA.$$
>
> This measures the degree to which A and B fail to commute.

It can be shown that

$$[T_i, T_j] = i\varepsilon_{ijk}T_k$$

which should be familiar from quantum mechanics as the commutation relation for angular momentum components (possibly with an extra factor of $\hbar$). There is a reason for this. Since T_i are the generators of rotations they lead to angular momentum in the same way that the generator of translations leads to linear momentum.

The commutation relations between the generators form part of the Lie[8] algebra, $\mathfrak{so}(3)$. This is the Lie algebra of the Lie group SO(3). What this means is that the elements of $\mathfrak{so}(3)$ generate SO(3). The elements of $\mathfrak{so}(3)$ are the are the commutator of linear combinations of T_i, for example

$$[T_1 + 2T_2, T_3] = [T_1, T_3] + 2[T_2, T_3] = -iT_2 + 2iT_1 \in \mathfrak{so}(3).$$

---

[8]pronounced 'lee'.

We say that T_i generate $\mathfrak{so}(3)$. To construct a finite rotation of angle ϑ about $\hat{\boldsymbol{n}}$ we consider N infinitesimal translations each taking us $\delta\vartheta = \vartheta/N$ and then take $N \to \infty$. So the rotation matrix is

$$R(\vartheta, \hat{\boldsymbol{n}}) = \lim_{N\to\infty} [R(\delta\vartheta, \hat{\boldsymbol{n}})]^N$$
$$= \lim_{N\to\infty} \left[I - i\frac{\vartheta}{N}\hat{\boldsymbol{n}} \cdot \boldsymbol{T} \right]^N$$
$$= \exp\left[-i\vartheta\hat{\boldsymbol{n}} \cdot \boldsymbol{T} \right].$$

Here we have used

$$e^x = \lim_{N\to\infty} \left[1 + \frac{x}{N} \right]^N$$

and assumed that this holds for exponentials containing matrices. Recall also that functions of matrices are defined through their power series and therefore

$$e^A = 1 + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots .$$

See appendix C for more detail on Lie algebras.

7   Tensors

7.1   Reflections

We will start this section with a definition and explain it after:

Definition 22: Reflection and Inversion

The **reflection** of a vector, $\boldsymbol{x}$, in a plane with surface normal $\hat{\boldsymbol{n}}$ is given by

$$\boldsymbol{y} = \boldsymbol{x} - 2(\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}$$

or

$$y_i = \sigma_{ij}x_j, \qquad \text{where} \qquad \sigma_{ij} = \delta_{ij} - 2n_i n_j.$$

An **inversion** of a vector, which is a reflection through the origin, is given by

$$\boldsymbol{y} = -\boldsymbol{x}$$

or

$$y_i = P_{ij}x_j, \qquad \text{where} \qquad P_{ij} = -\delta_{ij}.$$

P_{ij} is called the **parity operator**.

Consider the setup drawn in figure 7.1. From this it should be clear that $\boldsymbol{x} - (\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}$ is the projection of $\boldsymbol{x}$ onto the plane. Subtracting $(\boldsymbol{x} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}$ again gives us a vector the same distance below the plane again.
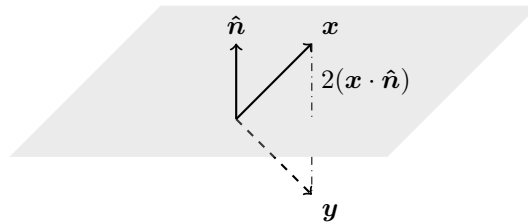
Figure 7.1: A vector, $\boldsymbol{x}$, and its reflection, $\boldsymbol{y}$, in the plane with surface normal $\hat{\boldsymbol{n}}$.

A reflection in the origin simply negates each component which is clearly the effect of the parity operator as described above. Further we can see this as a reflection where the vector being reflected is also the surface normal. Therefore the inverted vector is

$$\boldsymbol{y} = \boldsymbol{x} - 2(\boldsymbol{x} \cdot \hat{\boldsymbol{x}})\hat{\boldsymbol{x}} = \boldsymbol{x} - 2x\hat{\boldsymbol{x}} = \boldsymbol{x} - 2\boldsymbol{x} = -\boldsymbol{x}.$$

Reflections are self inverses, that is $P^2 = \sigma^2 = I$. They are also improper transformations, i.e. $\det P = \det \sigma = -1^9$. This means that reflections and inversions are included in $\mathrm{O}(n)$. In fact reflections generate $\mathrm{O}(n)$, by this we mean that any element of $\mathrm{O}(n)$, including rotations, can be made of some number of reflections. Rotations, as proper transformations, must be composed of an even number of reflections.

7.2   Projection Operators

> **Definition 23: Projection Operators**
>
> $P_{\boldsymbol{u}}$ is a **parallel projection operator** onto a vector $\boldsymbol{u}$ if
> $$P_{\boldsymbol{u}}\boldsymbol{u} = \boldsymbol{u}, \qquad \text{and} \qquad P_{\boldsymbol{u}}\boldsymbol{v} = \boldsymbol{0}$$
> where $\boldsymbol{v}$ is orthogonal to $\boldsymbol{u}$. $Q_{\boldsymbol{u}}$ is an **orthogonal projection operator** to $\boldsymbol{u}$ if
> $$Q_{\boldsymbol{u}}\boldsymbol{u} = \boldsymbol{0}, \qquad \text{and} \qquad Q_{\boldsymbol{u}}\boldsymbol{v} = \boldsymbol{v}$$
> where, again, $\boldsymbol{v}$ is orthogonal to $\boldsymbol{u}$.

Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be orthogonal vectors. Then
$$(I - P_{\boldsymbol{u}})\boldsymbol{u} = I\boldsymbol{u} - P_{\boldsymbol{u}}\boldsymbol{u} = \boldsymbol{u} - \boldsymbol{u} = \boldsymbol{0},$$
$$(I - P_{\boldsymbol{u}})\boldsymbol{v} = I\boldsymbol{v} - P_{\boldsymbol{u}}\boldsymbol{v} = \boldsymbol{v} - \boldsymbol{0} = \boldsymbol{v}.$$

We see that $Q_{\boldsymbol{u}} = I - P_{\boldsymbol{u}}$. Suitable operators are
$$P_{\boldsymbol{u}ij} = \frac{u_i u_j}{u^2}, \qquad \text{and} \qquad Q_{\boldsymbol{u}ij} = \delta_{ij} - \frac{u_i u_j}{u^2}.$$

> **Lemma 6: Uniqueness of projection operators**
>
> The projection operators for a given vector, $\boldsymbol{u} \in \mathbb{R}^3$, are unique.

Proof. Suppose $P_{\boldsymbol{u}}$ and $T_{\boldsymbol{u}}$ are both parallel projection operators for $\boldsymbol{u}$. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be orthogonal. Then $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{u} \times \boldsymbol{v}\}$ spans $\mathbb{R}^3$ and are all mutually orthogonal. Thus any vector, $\boldsymbol{w} \in \mathbb{R}^3$, can be written as a linear combination:
$$\boldsymbol{w} = \mu\boldsymbol{u} + \nu\boldsymbol{v} + \lambda(\boldsymbol{u} \times \boldsymbol{v}).$$

Consider now the operator $P_{\boldsymbol{u}} - T_{\boldsymbol{u}}$:
$$\begin{aligned}
(P_{\boldsymbol{u}} - T_{\boldsymbol{u}})\boldsymbol{w} &= (P_{\boldsymbol{u}} - T_{\boldsymbol{u}})[\mu\boldsymbol{u} + \nu\boldsymbol{v} + \lambda(\boldsymbol{u} \times \boldsymbol{v})] \\
&= \mu P_{\boldsymbol{u}}\boldsymbol{u} + \nu P_{\boldsymbol{u}}\boldsymbol{v} + \lambda P_{\boldsymbol{u}}(\boldsymbol{u} \times \boldsymbol{v}) - \mu T_{\boldsymbol{u}}\boldsymbol{u} - \nu T_{\boldsymbol{u}}\boldsymbol{v} - \lambda T_{\boldsymbol{u}}(\boldsymbol{u} \times \boldsymbol{v}) \\
&= \mu\boldsymbol{u} - \mu\boldsymbol{u} \\
&= \boldsymbol{0}
\end{aligned}$$

so $P_{\boldsymbol{u}}\boldsymbol{w} = T_{\boldsymbol{u}}\boldsymbol{w}$ for all $\boldsymbol{w} \in \mathbb{R}^3$ which means $P_{\boldsymbol{u}} = T_{\boldsymbol{u}}$.

Suppose now that $Q_{\boldsymbol{u}}$ and $S_{\boldsymbol{u}}$ are orthogonal projection operators for $\boldsymbol{u}$. Then
$$\begin{aligned}
(Q_{\boldsymbol{u}} - S_{\boldsymbol{u}})\boldsymbol{w} &= (Q_{\boldsymbol{u}} - S_{\boldsymbol{u}})[\mu\boldsymbol{u} + \nu\boldsymbol{v} + \lambda(\boldsymbol{u} \times \boldsymbol{v})] \\
&= \mu Q_{\boldsymbol{u}}\boldsymbol{u} + \nu Q_{\boldsymbol{u}}\boldsymbol{v} + \lambda Q_{\boldsymbol{u}}(\boldsymbol{u} \times \boldsymbol{v}) - \mu S_{\boldsymbol{u}}\boldsymbol{u} - \nu S_{\boldsymbol{u}}\boldsymbol{v} - \lambda S_{\boldsymbol{u}}(\boldsymbol{u} \times \boldsymbol{v}) \\
&= \nu\boldsymbol{v} + \lambda(\boldsymbol{u} \times \boldsymbol{v}) - \nu\boldsymbol{v} - \lambda(\boldsymbol{u} \times \boldsymbol{v}) \\
&= \boldsymbol{0}
\end{aligned}$$

Hence $Q_{\boldsymbol{u}} = S_{\boldsymbol{u}}$. $\square$

Note that projection operators are *not* orthogonal. For example if $\{\boldsymbol{e_i}\}$ is an orthonormal basis then $\boldsymbol{e_1} + \boldsymbol{e_2}$ has length $\sqrt{2}$ and $P_{\boldsymbol{e_1}}(\boldsymbol{e_1} + \boldsymbol{e_2}) = \boldsymbol{e_1}$ has length 1 so $P_{\boldsymbol{e_1}}$ doesn't preserve inner products so is not orthogonal.

---

[9]Inversion is only an improper transformation in an odd number of dimensions. In even dimensions all the negatives cancel to give a determinant of $+1$

7.3   Active and Passive Transformations

Recall that an active transformation involves changing the vector and a passive rotation involves changing the basis.

Consider the case when the basis, $\{e_i\}$ is transformed to $\{e_i'\}$ such that $\boldsymbol{x} = x_i\boldsymbol{e_i} = x_i'\boldsymbol{e_i'}$, with $x_i' = l_{ij}x_j$. This is a passive transformation but we can define a rotation matrix, $R_{ij} = l_{ij}$ and consider the vector $\boldsymbol{y} = R\boldsymbol{x}$ with components $y_i = R_{ij}x_j$. Clearly $y_i = l_P = ij = x_j'$. If we consider an active rotation about the z-axis then we have

$$R_{ij}(\vartheta, \boldsymbol{e_3}) = \delta_{ij} + (1 - \cos\vartheta)\delta_{i3}\delta_{j3} - \varepsilon_{ijk}\delta_{k3}\sin\vartheta,$$

$$R = \begin{pmatrix} \cos\vartheta & -\sin\vartheta & 0 \\ \sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Here we have used that $\boldsymbol{e_3} = n_i$ so $n_i = \delta_{i3}$. This matrix corresponds to an active transformation through the angle ϑ.

If instead we consider how the basis changes then we can readily see that

$$\boldsymbol{e_1'} = \cos\vartheta\boldsymbol{e_1} - \sin\vartheta\boldsymbol{e_2}$$
$$\boldsymbol{e_2'} = \sin\vartheta\boldsymbol{e_1} + \cos\vartheta\boldsymbol{e_2}$$
$$\boldsymbol{e_3'} = \boldsymbol{e_3}.$$

This is a passive rotation through angle $-\vartheta$.

In general an active rotation of $\boldsymbol{x}$ is the same as a passive rotation of basis vectors through an equal and opposite angle. The only thing that changes is what we are aiming to describe. An active transformation may represent a point on a rigid body moving whereas a passive transformation may represent a frame of reference rotating.

7.4   Tensors

7.4.1   Vectors

We have seen that under a rotation of the basis given by $L \in \mathrm{SO}(3)$ all vectors, $\boldsymbol{a} \in \mathbb{R}^3$, transform as

$$a_i' = l_{ij}a_j.$$

We now turn this on its head to give us a new definition of a vector:

> **Definition 24: Vector, tensor definition**
>
> A **vector**, $\boldsymbol{a} \in \mathbb{R}^d$, is an entity with d components, a_i, with respect to any basis. The components transform under a rotation, $L \in \mathrm{SO}(d)$, as
>
> $$a_i' = l_{ij}a_j.$$

7.4.2   Tensor Motivation

Consider two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$. The object a_ib_j has d^2 components. We want this object to transform in a way that means $(a_ib_j)' = a_i'b_j'$. If this is the case then

$$(a_ib_j)' = a_i'b_j' = l_{i\alpha}a_\alpha l_{j\beta}b_\beta = (l_{i\alpha}l_{j\beta})(a_\alpha b_\beta).$$

7.4.3   Tensor Definitions

Definition 25: Rank 2 tensor

A **rank 2 tensor**, T, in d dimensions is an object with d^2 components, T_{ij}, with respect to any basis. The components transform under a rotation, $L \in \mathrm{SO}(d)$, as

$$T'_{ij} = l_{i\alpha}l_{j\beta}T_{\alpha\beta}.$$

Notice from this definition that $a_i b_j$ as defined in the last section is a rank 2 tensor but not all rank two tensor components can be written as a product of vector components.

We can generalise this definition to a rank n tensor:

Definition 26: Rank n tensor

A **rank n tensor**, T, in d dimensions is an object with d^n components, $T_{i_1 i_2 \cdots i_n}$, with respect to any basis. The components transform under a rotation, $L \in \mathrm{SO}(d)$, as

$$T'_{i_1 i_2 \cdots i_n} = l_{i_1 \alpha_1} l_{i_2 \alpha_2} \cdots l_{i_n \alpha_n} T_{\alpha_1 \alpha_2 \cdots \alpha_n}.$$

The most common case for us will be $d = 3$ where a rank n tensor is an object with 3^n components, $T_{i_1 i_2 \cdots i_n}$, that transforms under a rotation, $L \in \mathrm{SO}(3)$, as

$$T'_{i_1 i_2 \cdots i_n} = l_{i_1 \alpha_1} l_{i_2 \alpha_2} \cdots l_{i_n \alpha_n} T_{\alpha_1 \alpha_2 \cdots \alpha_n}.$$

Note the distinction between rank and dimension. The rank is the number of indices whereas the dimension gives the possible values that each index can take, in general $i = 1, \ldots, d$ for a d-dimensional tensor.

From these definitions we can make two immediate observations:

- A vector is a rank 1 tensor, i.e. $a_i = l_{i\alpha}a_\alpha$.
- A scalar is a rank 2 tensor, i.e. $\varphi = \varphi'$.

While technically incorrect we will often say that $T_{i_1 i_2 \cdots i_n}$ is a rank n tensor when what we really mean is that $T_{i_1 i_2 \cdots i_n}$ are the components of a rank n tensor, T. In this course we make no distinction between covariant and contravariant indices. As such we use only lower indices since we don't need to separate the two concepts. Also since we are almost only interested in three dimensional Euclidean space we don't follow the normal convention where Greek indices run from 0 to 3 or 1 to 4 as is common when doing relativity. As a final note for this section while all vectors are tensors not all matrices are tensors. By this we mean that a matrix doesn't necessarily have an associated transformation law. This is the same as saying that although a vector can be written as a tuple not all tuples are vectors.

8   Properties of Tensors

8.1   Tensor Fields

Definition 27: Tensor field

Let $\mathcal{T}_n$ be the set of all rank n tensors. Then a **tensor field** is any function $T \colon \mathbb{R}^3 \to \mathcal{T}_n$. That is a tensor field assigns a tensor to every point in space.

An example of a tensor field is the strain in a material which is a rank 2 tensor field with components $E_{ij}(\boldsymbol{r})$. Since at any particular point $E_{ij}(\boldsymbol{r})$ are just the components of a tensor then E_{ij} transform as

$$E'_{ij}(\boldsymbol{r}) = l_{i\alpha}l_{j\beta}E_{\alpha\beta}(\boldsymbol{r}).$$

In the left hand side of this equation the components of $\boldsymbol{r}$ are considered in the $\{\boldsymbol{e'_i}\}$ basis whereas in the right hand side they are considered in the $\{\boldsymbol{e_i}\}$ basis. To emphasis this we often write $E'_{ij}(x'_k) = l_{i\alpha}l_{j\beta}E_{\alpha\beta}(x_k)$. Here x_k is shorthand for (x_1, x_2, x_3) rather than meaning a specific coordinate.

8.2   Dyadic Notation

Dyadic notation is a notation sometimes used when working with tensors of rank $n \leq 2$.

> **Notation**
>
> In dyadic notation scalars and vectors are represented as we are used to. Normally this means that scalars are simply letters, such as a or φ, and vectors are either bold or have an arrow or under line, such as $\boldsymbol{v}$, $\vec{v}$ or $\underline{v}$. Rank 2 tensors are then notated in a similar way with an extra arrow, underline or double headed arrow. For example $\vec{\vec{T}}$, $\overset{\leftrightarrow}{T}$ or $\underline{\underline{T}}$.

Dyadic notation is often less clear on meaning and therefore we won't use it here (apart from for vectors). We provide a short dictionary here for translation into index notation.

Dyadic		Index
$\vec{a}$	$\longleftrightarrow$	a_i
$\vec{a} \cdot \vec{b}$	$\longleftrightarrow$	$a_i b_i$
$\overset{\leftrightarrow}{T}$	$\longleftrightarrow$	T_{ij} or t_{ij}
$\vec{a}\overset{\leftrightarrow}{T}\vec{b}$	$\longleftrightarrow$	$a_i T_{ij} b_j$

8.3   Consistency of the Tensor Definition

Let $\mathcal{B} = \{\boldsymbol{e_i}\}$, $\mathcal{B}' = \boldsymbol{e'_i}\}$, and $\mathcal{B}'' = \boldsymbol{e''_i}\}$ be bases related by $\boldsymbol{e'_i} = l_{ij}\boldsymbol{e_j}$ and $\boldsymbol{e''_i} = m_{ij}\boldsymbol{e'_j}$ where $L, M \in \mathrm{SO}(3)$. We know that $\mathcal{B}$ and $\mathcal{B}''$ are related by the transformation matrix $N = ML$ so $\boldsymbol{e''_i} = n_{ij}\boldsymbol{e_j}$. For the definition of the transformation of a tensor to be consistent we need multiple transformations chained together to give the same components as a single equivalent transformation. We will show here that this is the case for a rank 2 tensor T which has components T_{ij} in $\mathcal{B}$, T'_{ij} in $\mathcal{B}'$ and T''_{ij} in $\mathcal{B}''$. This demonstration generalises easily to any given n.

$$
\begin{aligned}
T''_{ij} &= m_{i\alpha}m_{j\beta}T'_{\alpha\beta} \\
&= m_{i\alpha}m_{j\beta}(l_{\alpha r}l_{\beta s}T_{rs}) \\
&= (m_{i\alpha}l_{\alpha r})(m_{j\beta}l_{\beta s})T_{rs} \\
&= (ML)_{ir}(ML)_{js}T_{rs} \\
&= n_{ir}n_{js}T_{rs}
\end{aligned}
$$

which is exactly the same result we would expect going directly from $\mathcal{B}$ to $\mathcal{B}''$.

8.3.1   Properties of Cartesian Tensors

In this section we will prove various results about tensors. Most of these will be no surprise as they are simply generalisations of ideas we have met with vectors and matrices.

> **Lemma 7: Sum of tensors is a tensor**
>
> If T and U are rank n tensors then the object V with components
>
> $$V_{i_1 i_2 \cdots i_n} = T_{i_1 i_2 \cdots i_n} + U_{i_1 i_2 \cdots i_n}$$
>
> is also a tensor of rank n.

Proof. Let T and U be tensors of rank n with components $T_{i_1 i_2 \cdots i_n}$ and $U_{i_1 i_2 \cdots i_n}$ in $\mathcal{B} = \{\boldsymbol{e_i}\}$ and components $T'_{i_1 i_2 \cdots i_n}$ and $U'_{i_1 i_2 \cdots i_n}$ in basis $\mathcal{B}' = \{\boldsymbol{e'_i}\}$ where $\boldsymbol{e'_i} = l_{ij}\boldsymbol{e_j}$. Let V have components $V_{i_1 i_2 \cdots i_n} = T_{i_1 i_2 \cdots i_n} + U_{i_1 i_2 \cdots i_n}$ which transform as

$$
\begin{aligned}
V'_{i_1 i_2 \cdots i_n} &= T'_{i_1 i_2 \cdots i_n} + U'_{i_1 i_2 \cdots i_n} \\
&= l_{i_1 \alpha_1}l_{i_2 \alpha_2}\cdots l_{i_n \alpha_n}T_{\alpha_1 \alpha_2 \cdots \alpha_n} + l_{i_1 \alpha_1}l_{i_2 \alpha_2}\cdots l_{i_n \alpha_n}U_{\alpha_1 \alpha_2 \cdots \alpha_n} \\
&= l_{i_1 \alpha_1}l_{i_2 \alpha_2}\cdots l_{i_n \alpha_n}(T_{\alpha_1 \alpha_2 \cdots \alpha_n} + U_{\alpha_1 \alpha_2 \cdots \alpha_n}) \\
&= l_{i_1 \alpha_1}l_{i_2 \alpha_2}\cdots l_{i_n \alpha_n}V_{\alpha_1 \alpha_2 \cdots \alpha_n}.
\end{aligned}
$$

So V follows the expected transformation laws and is a rank n tensor.   □

Lemma 8: Product of tensors is a tensor

If T is a rank n tensor and U is a rank m tensor then the object V with components

$$V_{i_1 i_2 \cdots i_n j_1 j_2 \cdots j_m} = T_{i_1 i_2 \cdots i_n} U_{j_1 j_2 \cdots j_m}$$

is a tensor of rank $n + m$.

Proof. Let T and U be tensors of rank n and m respectively with components $T_{i_1 i_2 \cdots i_n}$ and $U_{i_1 i_2 \cdots i_m}$ in $\mathcal{B} = \{ \boldsymbol{e_i} \}$ and components $T'_{i_1 i_2 \cdots i_n}$ and $U'_{i_1 i_2 \cdots i_n}$ in basis $\mathcal{B}' = \{ \boldsymbol{e'_i} \}$ where $\boldsymbol{e'_i} = l_{ij} \boldsymbol{e_j}$. Let V have components $V_{i_1 i_2 \cdots i_n} = T_{i_1 i_2 \cdots i_n} U_{j_1 j_2 \cdots j_m}$ which transform as

$$
\begin{aligned}
V'_{i_1 i_2 \cdots i_n j_1 j_2 \cdots j_m} &= T'_{i_1 i_2 \cdots i_n} U'_{j_1 j_2 \cdots j_m} \\
&= (l_{i_1 \alpha_1} l_{i_2 \alpha_2} \cdots l_{i_n \alpha_n} T_{\alpha_1 \alpha_2 \cdots \alpha_n})(l_{j_1 \beta_1} l_{j_2 \beta_2} \cdots l_{j_m \beta_m} U_{\beta_1 \beta_2 \cdots \beta_m}) \\
&= l_{i_1 \alpha_1} l_{i_2 \alpha_2} \cdots l_{i_n \alpha_n} l_{j_1 \beta_1} l_{j_2 \beta_2} \cdots l_{j_n \beta_n} (T_{\alpha_1 \alpha_2 \cdots \alpha_n} U_{\beta_1 \beta_2 \cdots \beta_m}) \\
&= l_{i_1 \alpha_1} l_{i_2 \alpha_2} \cdots l_{i_n \alpha_n} l_{j_1 \beta_1} l_{j_2 \beta_2}.
\end{aligned}
$$

So V follows the expected transformation laws and is a rank $n + m$ tensor.   □

For example if λ is a scalar and $\boldsymbol{v}$ is a vector then $\lambda \boldsymbol{v}$ is a tensor of rank $0 + 1 = 1$ so is a vector.

Corollary 4: A linear combination of tensors is a tensor

Let $\{ T^{(i)} \mid i = 1, 2, \ldots N \}$ be rank n tensors and $\{ \lambda^{(i)} \mid i = 1, 2, \ldots, N \}$ be scalars. Then

$$T = \sum_{i=1}^{N} \lambda^{(i)} T^{(i)}.$$

is a rank n tensor. That is a finite linear combination of rank n tensors is a rank n tensor.

Proof. Consider first one particular term of the sum, $\lambda^{(i)} T^{(i)}$. By lemma 8 since this is a product of a rank n tensor and a rank 0 tensor it is a rank $n + 0 = n$ tensor. Thus the sum is a sum of rank n tensors. By lemma 7 considering terms pairwise we see that the sum must be a rank n tensor.   □

Definition 28: Tensor contraction

If T is a tensor of rank n with components $T_{ij \cdots r \cdots s \cdots k}$ then we define the tensor **contraction** across the indices r and s as the object with components

$$T_{ij \cdots r \cdots r \cdots k} = \sum_r T_{ij \cdots r \cdots r \cdots k}.$$

Lemma 9: A tensor contraction is a tensor

Let T be a rank n tensor with components $T_{ij \cdots r \cdots s \cdots k}$. Then the object with components $T_{ij \cdots r \cdots r \cdots k}$ is a rank $n - 2$ tensor.

Proof. Let T be a rank n tensor with components $T_{ij \cdots r \cdots s \cdots k}$ in basis $\mathcal{B} = \{ \boldsymbol{e_i} \}$ and components $T'_{ij \cdots r \cdots s \cdots k}$ in basis $\mathcal{B}' = \{ \boldsymbol{e'_i} \}$ where $\boldsymbol{e'_i} = l_{ij} \boldsymbol{e_j}$. Then the object with components $T_{ij \cdots r \cdots r \cdots k}$ transforms as

$$
\begin{aligned}
T'_{ij \cdots r \cdots r \cdots k} &= l_{i\alpha} l_{j\beta} \cdots l_{r\rho} \cdots l_{r\sigma} \cdots l_{k\gamma} T_{ij \cdots r \cdots r \cdots k} \\
&= l_{r\rho} l_{r\sigma} (l_{i\alpha} l_{j\beta} \cdots l_{k\gamma} T_{ij \cdots r \cdots r \cdots k}) \\
&= (L^{\mathsf{T}})_{\rho r} L_{r\sigma} (l_{i\alpha} l_{j\beta} \cdots l_{k\gamma} T_{ij \cdots r \cdots r \cdots k})
\end{aligned}
$$

$$= (L^\mathsf{T}L)_{\rho\sigma}(l_{i\alpha}l_{j\beta}\cdots l_{k\gamma}T_{ij\ldots r\ldots r\ldots k})$$
$$= (L^{-1}L)_{\rho\sigma}(l_{i\alpha}l_{j\beta}\cdots l_{k\gamma}T_{ij\ldots r\ldots r\ldots k})$$
$$= l_{i\alpha}l_{j\beta}\cdots l_{k\gamma}T_{ij\ldots r\ldots r\ldots k}.$$

So this follows the expected transformation laws and is a rank $n-2$ tensor.   □

The most common case of a tensor contraction comes from a rank 2 tensor with components a_ib_j which when contracted gives a_ib_i which is just the scalar product of $\boldsymbol{a}$ and $\boldsymbol{b}$. What we proved above shows therefore that the scalar product of two vectors (together one rank 2 tensor) gives a tensor of rank $2-2=0$, i.e. a scalar.

8.3.2   Properties of rank 2 tensors

In the following we will focus on rank two tensors although many of these ideas generalise to higher rank tensors.

> **Definition 29: Transpose**
>
> Let T be a rank 2 tensor with components T_{ij}. Then we define T^T as the object with components $(T^\mathsf{T})_{ij} = T_{ji}$.

In the specific case of a rank 2 tensor, T, we can write the transformation law as

$$T_{\mathcal{B}'} = LT_{\mathcal{B}}L^\mathsf{T}$$

where $T_{\mathcal{B}}$ is a matrix with elements T_{ij} which are the components of T in the basis $\mathcal{B} = \{\boldsymbol{e_i}\}$ and $T_{\mathcal{B}'}$ is a matrix with elements T'_{ij} which are the components of T in the basis $\mathcal{B}' = \boldsymbol{e'_i}$ and these two bases are related by $\boldsymbol{e'_i} = l_{ij}\boldsymbol{e_j}$. To see that this the case consider

$$T'_{ij} = l_{i\alpha}l_{j\beta}T_{\alpha\beta} = l_{i\alpha}T_{\alpha\beta}l_{j\beta} = l_{i\alpha}T_{\alpha\beta}(L^\mathsf{T})_{\beta j} = (LTL^\mathsf{T})_{ij}$$

Since L is orthogonal this can also be written as

$$T_{\mathcal{B}'} = LT_{\mathcal{B}}L^\mathsf{T}.$$

> **Lemma 10: Transpose is a tensor**
>
> Let T be a rank 2 tensor. Then T^T is a rank 2 tensor.

Proof. Let T be a rank 2 tensor with components T_{ij} in $\mathcal{B} = \{\boldsymbol{e_i}\}$ and components T'_{ij} in $\mathcal{B}' = \{\boldsymbol{e'_i}\}$ where $\boldsymbol{e'_i} = l_{ij}\boldsymbol{e_j}$. Then T^T has components $(T^\mathsf{T})_{ij} = T_{ji}$ which transform as

$$(T'^\mathsf{T})_{ij} = (T')_{ji}$$
$$= l_{j\beta}l_{i\alpha}T_{\beta\alpha}$$
$$= l_{i\alpha}l_{j\beta}(T^\mathsf{T})_{\alpha\beta}$$

So T^T follows the expected transformation laws and is a rank n tensor.   □

> **Definition 30: Symmetric and antisymmetric tensors**
>
> Let T be a rank 2 tensor. Then if $T = T^\mathsf{T}$ we say T is **symmetric** and if $T = -T^\mathsf{T}$ we say T is **antisymmetric**.

> **Lemma 11: symmetric decomposition**
>
> Any rank 2 tensor can always be decomposed into a sum of an odd and even tensor.

Proof. Let T be a rank 2 tensor. Then clearly

$$T = \frac{1}{2}(T + T^{\mathsf{T}}) + \frac{1}{2}(T - T^{\mathsf{T}}).$$

In index notation this becomes

$$T_{ij} = \frac{1}{2}(T_{ij} + T_{ji}) + \frac{1}{2}(T_{ij} - T_{ji}).$$

Let $S_{ij} = T_{ij} + T_{ji}$. Then

$$(S^{\mathsf{T}})_{ij} = S_{ji} = T_{ji} + T_{ij} = T_{ij} + T_{ji} = S_{ij}$$

so S is symmetric. Let $A_{ij} = T_{ij} - T_{ji}$. Then

$$(A^{\mathsf{T}})_{ij} = A_{ji} = T_{ji} - T_{ij} = -(T_{ij} - T_{ji}) = -A_{ij}$$

so A is antisymmetric.

Since S and A are defined by sums of the tensor T they are themselves tensors. Finally if $\lambda \neq 0$ is a scalar then if λU is symmetric (antisymmetric) then U is symmetric (antisymmetric). This means that $S/2$ and $A/2$ are symmetric and antisymmetric respectively so

$$T = \frac{1}{2}S + \frac{1}{2}A$$

is a sum of a symmetric and antisymmetric tensor.   □

8.3.3   Kronecker Delta

The Kronecker delta, defined by

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

in all bases, is a rank 2 tensor. This is easy to see if we require $\delta'_{ij} = \delta_{ij}$ then δ transforms as

$$\delta'_{ij} = l_{i\alpha}l_{j\beta}\delta'_{\alpha\beta} = l_{i\alpha}l_{j\beta}\delta_{\alpha\beta}.$$

Since δ has the same components in all bases we say that it is isotropic.

This is the first example of a rank 2 tensor we have seen that is not simple. A rank n tensor is simple if its components can be written as $a_i b_j \cdots c_n$ for vectors $\boldsymbol{a}, \boldsymbol{b}, \ldots, \boldsymbol{c}$. This isn't the case for δ as if its components could be written as its diagonal requires $a_i b_j$ be non-zero and its off-diagonal requires $a_i b_j$ be zero. Therefore the diagonal requires both a_i and b_j be non-zero whereas the off diagonal requires at least one of a_i and b_j to be zero.

8.4   Quotient Theorem

<div style="border:1px solid red;">

Theorem 5: Quotient theorem (rank 2)

Let T be an object with 9 components given by T_{ij} in some basis, $\mathcal{B} = \{\boldsymbol{e_i}\}$, and T'_{ij} in a different basis, $\mathcal{B}' = \{\boldsymbol{e'_i}\}$, where $\mathcal{B}$ and $\mathcal{B}'$ are related by $\boldsymbol{e'_i} = l_{ij}\boldsymbol{e_j}$. Let $\boldsymbol{a}$ be an arbitrary vector. Let b be an object with 3 components given by $b_i = T_{ij}a_j$ in basis $\mathcal{B}$ and $b'_i = T'_{ij}a'_j$ in basis $\mathcal{B}'$. Then T is a tensor if b always transforms as a vector.

</div>

Proof. In $\mathcal{B}'$ using the transformation law for $\boldsymbol{a}$'s components we have

$$b'_i = T'_{ij}a'_j = T'_{ij}l_{jk}a_k.$$

Now suppose that b transforms as a vector. Then using b's transformation law and the definition of b_j we have

$$b'_i = l_{ij}b_j = l_{ij}T_{jk}a_k.$$

Clearly both of these forms for b_i' must be equal so

$$T_{ij}'l_{jk}a_k = l_{ij}T_{jk}a_k.$$

This must hold for all vectors $\boldsymbol{a}$ meaning it holds for some specific vector where $a_k \neq 0$ and so

$$T_{ij}'l_{jk} = l_{ij}T_{jk}.$$

Multiplying both sides by l_{mk} this becomes

$$T_{ij}'l_{jk}l_{mk} = l_{ij}l_{mk}T_{jk}. \tag{8.1}$$

Now consider

$$l_{jk}l_{mk} = (L)_{jk}(L^{\mathsf{T}})_{km} = (LL^{\mathsf{T}})_{jm} = \delta_{jm}.$$

So 8.1 becomes

$$T_{ij}'\delta_{jm} = T_{im}' = l_{ij}l_{mk}T_{jk}.$$

This is exactly the transformation law that defines a rank 2 tensor so T is a rank 2 tensor. □

The quotient theorem generalises to

> **Theorem 6: Quotient theorem (rank n)**
>
> Let T be an object with 3^n components given by $T_{i_1 i_2 \cdots i_n}$ in some basis, $\mathcal{B} = \boldsymbol{e_i}$, and $T_{i_1 i_2 \cdots i_n}'$ in a different basis, $\mathcal{B}' = \{\boldsymbol{e_i'}\}$. Let R be an arbitrary rank $m < n$ tensor. Then T is a tensor if $T_{i_1 i_2 \cdots i_n}R_{i_1 i_2 \cdots i_m}$ are the components of a tensor of rank $n - m$.

It is actually common to take this to be the definition of a tensor and then derive the transformation laws from here. This is done in appendix D.

9   Pseudotensors

Up until now we have considered only transformations in SO(3). If instead we consider transformations in O(3) then recall that for $L \in$ O(3) either $\det L = 1$ for a rotation or $\det L = -1$ for a rotation and reflection.

> **Definition 31: Pseudotensor**
>
> Let $\mathcal{B} = \{\boldsymbol{e_i}\}$ and $\mathcal{B}' = \{\boldsymbol{e_i'}\}$ be bases related by $\boldsymbol{e_i'} = l_{ij}\boldsymbol{e_j}$ but now let $L = (l_{ij}) \in$ O(3). Then T is a rank n tensor if it transforms as
>
> $$T_{i_1 i_2 \cdots i_n} = l_{i_1 \alpha_1} l_{i_1 \alpha_2} \cdots l_{i_n \alpha_n} T_{\alpha_1 \alpha_2 \cdots \alpha_n}.$$
>
> If instead T transforms as
>
> $$T_{i_1 i_2 \cdots i_n} = \det(L)l_{i_1 \alpha_1} l_{i_1 \alpha_2} \cdots l_{i_n \alpha_n} T_{\alpha_1 \alpha_2 \cdots \alpha_n}$$
>
> then we say T is a **pseudotensor**. That is a pseudotensor picks up an extra minus sign if it is reflected.

9.1   Pseudovectors

Let $\boldsymbol{a}$ be a vector. Then under a reflection in the origin, given by $l_{ij} = -\delta_{ij}$ we have $a_i' = -\delta_{ij}a_j = -a_i$. We also have $\boldsymbol{e_i'} = -\delta_{ij}\boldsymbol{e_j} = -\boldsymbol{e_i}$. Hence

$$\boldsymbol{a} = a_i'\boldsymbol{e_i'} = (-a_i)(-\boldsymbol{e_i}) = a_i\boldsymbol{e_i}$$

so, as expected, vectors transform as rank 1 tensors under O(3) transformations.

A **pseudovector** is simply a rank 1 pseudotensor. These most commonly occur when we have a cross product. Consider $c = a \times b$ where a and b are vectors. Considering the components of c under a reflection in the origin we can easily show it is a pseudovector. For example

$$c_1' = a_2' b_3' - a_3' b_2' = (-a_2)(-b_3) - (-a_3)(-b_2) = a_2 b_3 - a_3 b_2 = c_1.$$

We find that in general $c_i' = c_i$ and so

$$c' = c_i' e_i' = c_i(-e_i) = -c_i e_i = -c.$$

So c is a pseudovector.

9.2   Levi–Civita Symbol

The Levi–Civita symbol is defined by

$$\varepsilon_{ijk} = \begin{cases} 1, & \text{if } (i,j,k) = (1,2,3), (2,3,1), (3,1,2) \\ -1, & \text{if } (i,j,k) = (1,3,2), (2,1,3), (3,2,1) \\ 0, & \text{otherwise} \end{cases}$$

in all bases. We can write the components of $c = a \times b$ as $c_i = \varepsilon_{ijk} a_j b_k$ in the basis $\mathcal{B} = \{e_i\}$ and as $c_i' = \varepsilon_{ijk} a_j' b_k'$ in the basis $\mathcal{B}' = \{e_i'\}$ defined by $e_i' = l_{ij} e_j$. From this we have

$$\begin{aligned} c_i' &= \varepsilon_{ijk} a_j' b_k' \\ &= \varepsilon_{ijk} l_{j\alpha} l_{k\beta} b_\beta \\ &= \varepsilon_{ijk} (l_{m\delta} l_{i\delta}) l_{j\alpha} l_{k\beta} a_\alpha b_\beta \\ &= \det(L) \varepsilon_{\delta\alpha\beta} l_{i\delta} a_\alpha b_\beta \\ &= \det(L) l_{i\delta} (a \times b)_\delta \\ &= \det(L) l_{i\delta} c_\delta. \end{aligned}$$

This shows (again) that $c = a \times b$ is a pseudovector.

The Levi–Civita symbols is a rank 3 pseudotensor. This is fairly easy to show. First we use the fact that ε_{ijk} is isotropic, that is the same in all frames, and also that for $L \in \mathrm{O}(3) \det(L)^2 = 1$:

$$\varepsilon_{ijk} = \varepsilon_{ijk} = \det(L) \det(L) \varepsilon_{ijk} = \det(L) l_{i\alpha} l_{j\beta} l_{k\gamma} \varepsilon_{\alpha\beta\gamma}.$$

In the last step we used the definition of the determinant. We can identify the last term here as the transformation law for a rank 3 pseudotensor.

We can use the Levi–Civita symbol to build pseudotensors of higher ranks. For example if a and b are vectors then $\varepsilon_{ijk} a_l b_m$ are the components of a rank 5 pseudotensor. Contracting across two pairs of indices, j, l and k, m, we have $\varepsilon_{ijk} a_j b_k$, which is simply $(a \times b)_i$ which is a pseudovector.

In general if an object has components written as a product of the components of tensors and pseudotensors then the object is a pseudotensor if the number of pseudotensors is odd and a tensor if the number of pseudotensors is even. Some examples of tensors and pseudotensors are given here:

- Position, r, is a vector.

- Velocity, $v = \dot{r}$, and acceleration, $a = \ddot{r}$, are vectors.

- Mass, m, is a scalar.

- Force, $F = ma$, is a vector.

- Charge, q, is a scalar.

- The electric field, $E = F/q$, is a vector.

- The torque, $\tau = r \times F$, is a pseudovector.

- The angular velocity, ω, is a pseudovector. This is because $v = \omega \times r$ and v and r are both vectors so ω must be a pseudotensor so that $\varepsilon_{ijk} \omega_j x_k$ are the components of a vector.

- Angular momentum, $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p} = \boldsymbol{r} \times m\boldsymbol{p}$, is a pseudovector.

- The magnetic field, $\boldsymbol{B}$, is a pseudovector. This can be seen as a consequence of the Biot–Savart law which defines
$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \oint_C \frac{I \, d\boldsymbol{r} \times (\boldsymbol{r} - \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^3}.$$
  So $\boldsymbol{B}$ is defined as the cross product of two position vectors so is a pseudovector.

- The quantity $\boldsymbol{E} \cdot \boldsymbol{B}$ is a pseudoscalar.

- The magnetic flux, $\boldsymbol{B} \cdot d\boldsymbol{S}$, is a pseudoscalar.

9.3   Isotropic Tensors

Definition 32: Isotropic, invariant

A tensor, T, is **isotropic** (or **invariant**), if it has the same components, $T_{i_1 i_2 \cdots i_n}$, in any Cartesian basis. That is
$$T_{i_1 i_2 \cdots i_n} = l_{i_1 \alpha_1} l_{i_2 \alpha_2} \cdots l_{i_n \alpha_n} T_{\alpha_1 \alpha_2 \cdots \alpha_n}.$$
Similarly a pseudotensor, T, is isotropic if
$$T_{i_1 i_2 \cdots i_n} = \det(L) l_{i_1 \alpha_1} l_{i_2 \alpha_2} \cdots l_{i_n \alpha_n} T_{\alpha_1 \alpha_2 \cdots \alpha_n}.$$

The word isotropic here refers to the fact that the components are necessarily the same in all directions as they don't change under a rotation of the basis. The word invariant simply applies to any quantity that is independent of the basis.

Theorem 7: Unique rank 2 isotropic tensor

There is, up to a scale factor, one isotropic rank two tensor, namely the Kronecker delta.

Proof. Let a_{ij} be a second rank isotropic tensor. Consider a rotation by $\pi/2$ about the z-axis given by

$$L = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The only non-zero elements are $l_{12} = l_{33} = -l_{21} = 1$. Thus we have

$$a_{11} = a'_{11} = l_{1\alpha} l_{1\beta} a_{\alpha\beta} = l_{12} l_{12} a_{22} = a_{22},$$
$$a_{13} = a'_{13} = l_{1\alpha} l_{3\beta} a_{\alpha\beta} = l_{12} l_{33} a_{22} = a_{23},$$
$$a_{23} = a'_{23} = l_{2\alpha} l_{3\beta} a_{\alpha\beta} = l_{21} l_{33} a_{13} = -a_{13}.$$

Hence

$$a_{11} = a_{22}, \qquad \text{and} \qquad a_{12} = a_{23} = 0.$$

Similarly considering a rotation of $\pi/2$ about the y-axis we can show that

$$a_{11} = a_{33}, \qquad \text{and} \qquad a_{12} = a_{32} = 0.$$

Finally a rotation of $\pi/2$ about the x-axis shows

$$a_{22} = a_{33}, \qquad \text{and} \qquad a_{21} = a_{31} = 0.$$

Hence we have $a_{ij} = \lambda \delta_{ij}$. $\square$

> **Theorem 8: Higher rank isotropic (pseudo)tensors**
>
> There is no (non-zero) isotropic tensor of rank 3. The most general isotropic pseudotensor of rank 3 has components
> $$a_{ijk} = \lambda \varepsilon_{ijk}.$$
> The most general rank 4 isotropic tensor has components
> $$a_{ijkl} = \lambda \delta_{ij}\delta_{kl} + \mu \delta_{ik}\delta_{jl} + \nu \delta_{il}\delta_{jk}.$$
> For odd n there are no isotropic rank n tensors. The most general rank n isotropic pseudotensor can be written as a linear combination of products of the Kronecker delta and Levi–Civita symbol. If n is even then the most general isotropic rank n tensor can be written as a product of Kronecker deltas (to have an even number of indices Levi–Civita symbols must appear in pairs and so can be written as products of Kronecker deltas).

Proof. Similar to the previous theorem but longer and not interesting.

10 Tensor Aspects of Gradient, Divergence, and Curl

The gradient, divergence, and curl operators generalise to act on tensors of various ranks. Many theorems involving these operators also generalise. In this section we will discuss these generalisations. As is normal in physics we will assume all objects are sufficiently 'nice' that there is no weird or undefined behaviour. This means assuming that sufficient derivatives exist and are sufficiently smooth as well as assuming that hypersurfaces are smooth, orientable, etc.

10.1 Gradient

Recall that the gradient operator, ∇, acts on a scalar field, φ, to give a vector, $\nabla\varphi$, with components $(\nabla\varphi)_i = \partial_i\varphi$. We can check that this is indeed a vector with our transformation definition. Consider two bases, $\mathcal{B} = \{e_i\}$ and $\mathcal{B}' = \{e'_i\}$ related by $x'_i = l_{i\alpha}x_\alpha$ for some $L = (l_{ij}) \in \mathrm{SO}(d)$. We can rewrite this relationship as $x_\alpha = x'_i l_{i\alpha}$. Hence

$$(\nabla\varphi)'_i = \frac{\partial \varphi}{\partial x'_i} = \frac{\partial x_\alpha}{\partial x'_i}\frac{\partial \varphi}{\partial x_\alpha} = l_{i\alpha}\frac{\partial \varphi}{\partial x_\alpha} = l_{i\alpha}(\nabla\varphi)_\alpha.$$

So $\nabla\varphi$ is indeed a vector.

> **Definition 33: Gradient of a tensor**
>
> Let T be a rank n tensor. The gradient of T is a rank $n+1$ tensor with components
> $$\frac{\partial T_{i_1 i_2 \cdots i_n}}{\partial x_j}.$$

For example $T = x_i(\boldsymbol{r} \cdot \boldsymbol{a}) = x_i x_j a_j$ for some constant vector $\boldsymbol{a}$ is a rank two tensor. The gradient, $\operatorname{grad} T$, is then a rank 2 tensor with components

$$(\operatorname{grad} T)_{ij} = (\nabla T_i)_j = \frac{\partial T_i}{\partial x_j} = \delta_{ij}(\boldsymbol{r} \cdot \boldsymbol{a}) + x_i a_j.$$

For a scalar field φ the fundamental theorem of multivariable calculus is

$$\int_{P_1}^{P_2} \mathrm{d}\boldsymbol{r} \cdot \nabla\varphi = \varphi(P_2) - \varphi(P_1).$$

This is because $\mathrm{d}\boldsymbol{r} \cdot \nabla\varphi = \mathrm{d}\varphi$ and so this is simply the normal fundamental theorem of calculus. Notice that the result depends only on the endpoints of the line integral, not on the path taken between the points. This generalises to a tensor, T, as

$$\int_{P_1}^{P_2} \mathrm{d}x_i\, \partial_i T_{j_1 j_2 \cdots j_n} = T_{j_1 j_2 \cdots j_n}(P_2) - T_{j_1 j_2 \cdots j_n}(P_1).$$

10.2   Divergence

Recall that the divergence operator, $\nabla\cdot$, acts on a vector field, $\boldsymbol{a}$, to give a scalar, $\partial_i a_i$.

> **Definition 34: Divergence of a tensor**
>
> Let T be a rank n tensor. The divergence of T is a rank $n-1$ tensor with components
>
> $$\frac{\partial T_{i_1 i_2 \cdots i_n}}{\partial x_{i_1}}.$$
>
> Note that we can differentiate with respect to any of x_{i_j} so be careful which we do.

The divergence theorem,

$$\int_V \nabla \cdot \boldsymbol{a} \, \mathrm{d}V = \int_S \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{S} \,,$$

generalises to

$$\int_V \frac{\partial}{\partial x_{i_1}} T_{i_1 i_2 \cdots i_n} \, \mathrm{d}V = \int_S T_{i_1 i_2 \cdots i_n} \, \mathrm{d}S_{i_1} \,.$$

10.3   Curl

Recall that the curl operator, $\nabla\times$, acts on a vector field, $\boldsymbol{a}$, to give a pseudovector, $\nabla \times \boldsymbol{a}$, with components $(\nabla \times \boldsymbol{a})_i = \varepsilon_{ijk}\partial_j a_k$. There isn't a simple generalisation of curl to a tensor field but there is a generalisation of Stokes' theorem. Stokes' theorem states

$$\int_S (\nabla \times \boldsymbol{a}) \cdot \mathrm{d}\boldsymbol{S} = \oint_C \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{r} \,.$$

This generalises to

$$\int_S \varepsilon_{ijk_1} \frac{\partial}{\partial x_j} T_{k_1 k_2 \cdots k_n} \, \mathrm{d}S_i = \oint_C T_{k_1 k_2 \cdots k_n} \, \mathrm{d}x_{k_1} \,.$$

For example if φ is a scalar field then

$$\begin{aligned}
\left[\int_S \mathrm{d}\boldsymbol{S} \times \nabla\varphi \right]_k &= \int_S \varepsilon_{kij} \, \mathrm{d}S_i \, \frac{\partial}{\partial x_j}\varphi \\
&= \int_S \varepsilon_{ijk} \, \mathrm{d}S_i \, \frac{\partial}{\partial x_j}\varphi \\
&= \oint_C \varphi \, \mathrm{d}x_k \\
&= \left[\oint_C \varphi \, \mathrm{d}\boldsymbol{r} \right]_k.
\end{aligned}$$

11   Taylor Expansions

Let f be a function with m continuous derivatives on $[a,b]$. Then by the fundamental theorem of calculus

$$\int_a^{x_1} f^{(m)}(x_0) = f^{(m-1)}(x_1) - f^{(m-1)}(a).$$

Continuing on and integrating m times we have

$$\begin{aligned}
I &= \int_a^{x_m} \cdots \int_a^{x_1} f^{(m)}(x_0) \, \mathrm{d}x_0 \cdots \mathrm{d}x_{m-1} \\
&= \int_a^{x_m} \cdots \int_a^{x_2} [f^{(m-1)}(x_1) - f^{(m-1)}(a)] \, \mathrm{d}x_1 \cdots \mathrm{d}x_{m-1} \\
&= \int_a^{x_m} \cdots \int_a^{x_3} [f^{(m-2)}(x_2) - f^{(m-2)}(a) - (x_2 - a)f^{(m-1)}(a)] \, \mathrm{d}x_2 \cdots \mathrm{d}x_{m-1}
\end{aligned}$$

$$\vdots$$

$$= f(x_m) - f(a) - (x_m - a)f'(a) - \frac{1}{2!}(x_m - a)^2 f''(a) - \cdots - \frac{1}{(m-1)!}(x_m - a)^{m-1} f^{(m-1)}(a).$$

Here we have used

$$\int_a^x (y-a)^{n-1} \, \mathrm{d}y = \frac{1}{n}(x-a)^n.$$

Now let $x_m = x$. Rearranging we have

$$f(x) = f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots + \frac{1}{(m-1)!}(x-a)^{m-1} f^{(m-1)}(a) + R_m(x)$$

where

$$R_m(x) = \int_a^x \cdots \int_a^{x_1} f^{(m)}(x_0) \, \mathrm{d}x_0 \cdots \mathrm{d}x_{m-1}.$$

The mean value theorem states that there exists some $\zeta \in [a,b]$ such that if $g \colon [a,b] \to \mathbb{R}$ is continuous then

$$\int_a^b g(x) \, \mathrm{d}x = (b-a)g(\zeta).$$

Applying this to $f^{(m)}$ we have

$$\int_a^x f^{(m)}(x_0) \, \mathrm{d}x_0 = (x-a)f^{(m)}(\zeta)$$

for some $\zeta \in [a,x]$. Hence

$$R_m(x) = \frac{1}{m!}(x-a)^m f^{(m)}(\zeta).$$

The same logic as above but considering integrals from x_i to a instead of from a to x_i gives us the same equations. If

$$\lim_{m \to \infty} R_m(x) = 0$$

then we have an infinite series that exactly recreates f. The set of values of x for which this expansion is valid is called the region of convergence.

Definition 35: Taylor Series

Let $f \colon \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function. Then we define the Taylor series of f as

$$\sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(x-a)^k.$$

In the special case $a = 0$ this is often called a Maclaurin series and takes the form

$$\sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)} x^k.$$

Under a variety of conditions the value of a function at a point will be the same as the value of the Taylor series at that point.

The physicist's 'proof' that of a Maclaurin series is as follows. Suppose there exists a series expansion for f of the form

$$f(x) = \sum_{k=0}^{\infty} a_k x^k.$$

Differentiating this n times we have

$$f^{(n)}(x) = 0 + 0 + \cdots + 0 + n!a_n + \frac{(n+1)!}{2}a_{n+1}x + \frac{(n+2)!}{2 \cdot 3}a_{n+2}x^2 + \cdots.$$

Evaluating this at 0 we have

$$f^{(n)}(0) = 0 + 0 + \cdots + 0 + n!a_n + \frac{(n+1)!}{2}a_{n+1} \cdot 0 + \frac{(n+2)!}{2 \cdot 3}a_{n+2} \cdot 0^2 + \cdots = n!a_n.$$

Rearranging this gives

$$a_n = \frac{1}{n!}f^{(n)}(0)$$

so

$$f(x) = \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} \frac{1}{k!}f^{(k)}(a)x^k.$$

A Taylor series is then simply a Maclaurin series shifted to a different point.

11.1   Examples

Let $f(x) = \sin(x)$. Find a Taylor series for f about $x = 0$.

$$
\begin{array}{ll}
f(x) = \sin(x), & f(0) = 0 \\
f'(x) = \cos(x), & f'(0) = 1 \\
f''(x) = -\sin(x), & f''(0) = 0 \\
f'''(x) = -\cos(x), & f'''(0) = -1 \\
f^{(4)}(x) = \sin(x), & f^{(4)}(0) = 0 \\
f^{(5)}(x) = \cos(x), & f^{(5)}(0) = 1 \\
f^{(6)}(x) = -\sin(x), & f^{(6)}(0) = 0
\end{array}
$$

Noticing a pattern we have

$$
\begin{array}{ll}
f^{(2n)}(x) = (-1)^n \sin(x), & f^{(2n)}(0) = 0 \\
f^{(2n+1)}(x) = (-1)^n \cos(x), & f^{(2n+1)}(0) = (-1)^n.
\end{array}
$$

Since $\left|f^{(m)}(\zeta)\right| \leq 1$ for all $\zeta \in \mathbb{R}$ and $m \in \mathbb{N}$ for some constant x as $m \to \infty$ we have

$$
\begin{aligned}
|R_m| &= \frac{1}{m!}|x^m||f^m(\zeta)| \\
&\leq \frac{1}{m!}x^m \\
&\to 0
\end{aligned}
$$

So $|R_m| \to 0$ meaning $R_m \to 0$ as required. Hence

$$
\begin{aligned}
\sin(x) &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\
&= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \frac{1}{9!}x^9 + \mathcal{O}(x^{-11}).
\end{aligned}
$$

Let $f(x) = (1+x)^\alpha$ for some $\alpha \in \mathbb{R}$.

$$
\begin{array}{ll}
f(x) = (1+x)^\alpha, & f(0) = 1 \\
f'(x) = \alpha(1+x)^{\alpha-1}, & f(0) = \alpha \\
f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2}, & f(0) = \alpha(\alpha-1) \\
f'''(x) = \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3}, & f(0) = \alpha(\alpha-1)(\alpha-2) \\
f^{(4)}(x) = \alpha(\alpha-1)(\alpha-2)(\alpha-3)(1+x)^{\alpha-4}, & f(0) = \alpha(\alpha-1)(\alpha-2)(\alpha-3) \\
f^{(5)}(x) = \alpha(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)(1+x)^{\alpha-5}, & f(0) = \alpha(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)
\end{array}
$$

Noticing a pattern we have

$$f^{(n)}(x) = (1+x)^{\alpha-n} \prod_{k=0}^{n-1} (\alpha-k), \qquad\qquad f^{(n)}(0) = \prod_{k=0}^{n-1} (\alpha-k).$$

It can be shown that for $\alpha \in \mathbb{R}$ then $R_m \to 0$ as $m \to \infty$ if $|x| < 1$. Thus if $|x| < 1$ then

$$(1+x)^\alpha = \sum_{n=0}^{\infty} x^n \prod_{k=0}^{n} (\alpha-k)$$

$$= 1 + \alpha x + \frac{1}{2!}\alpha(\alpha-1)x^2 + \frac{1}{3!}\alpha(\alpha-1)(\alpha-2)x^3 + \frac{1}{4!}\alpha(\alpha-1)(\alpha-2)(\alpha-3)x^4 + \mathcal{O}(x^5).$$

For the special case $\alpha \in \mathbb{N}$ we have

$$\prod_{k=0}^{n-1} (\alpha-k) = \frac{\alpha!}{(\alpha-n)!}$$

and it can be shown in this case that for all $x \in \mathbb{R}$

$$(1+x)^\alpha = \sum_{n=0}^{\alpha} \frac{\alpha!}{n!(\alpha-n)!} x^n = \sum_{n=0}^{\alpha} \binom{\alpha}{n} x^n.$$

So we recover the standard binomial expansion.

We should be careful as not all functions have Taylor series. For example

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

Here $f^{(m)}(0) = 0$ for all $m \in \mathbb{N}$ clearly $f(x) \neq 0$ for $x \neq 0$ so the Taylor series and the function do not have the same value. We say that f has an essential singularity at $x = 0$.

Since this is a physics course we will assume convergence most of the time and test the Taylor series for some values to see if it is valid. If asked on an exam to find a Taylor series in this course it is ok to assume it exists.

Some common Maclaurin series are

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \qquad = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \mathcal{O}(x^7), \qquad \forall x \in \mathbb{R},$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \qquad = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \mathcal{O}((x^6)), \qquad \forall x \in \mathbb{R},$$

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n \qquad = 1 + x + \frac{x^2}{2!} + \mathcal{O}(x^3), \qquad \forall x \in \mathbb{R}$$

$$\sinh(x) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} x^{2n+1} \qquad = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \mathcal{O}(x^7), \qquad \forall x \in \mathbb{R},$$

$$\cosh(x) = \sum_{n=0}^{\infty} \frac{1}{(2n)!} x^{2n} \qquad = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \mathcal{O}((x^6)), \qquad \forall x \in \mathbb{R},$$

$$(1+x)^\alpha = \sum_{n=0}^{\alpha} x^n \prod k = 0^{n-1}(\alpha-k) \qquad = 1 + \alpha x + \frac{x^2}{2!}\alpha(\alpha-1) + \mathcal{O}(x^3), \qquad \forall x \in (0,1),$$

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n \qquad = x - \frac{x^2}{2} + \frac{x^3}{3} - \mathcal{O}(x^4), \qquad \forall x \in (0,1).$$

11.1.1 Other Forms

Consider a function f. Define $g(a) = f(x+a)$ for some fixed x. Then using the Taylor series of g we have

$$f(x+a) = g(a) = \sum_{n=0}^{\infty} \frac{1}{n!} g^{(n)}(0) a^n.$$

Also

$$g^{(n)}(0) = f^{(n)}(x)$$

so

$$f(x + a) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x) a^n = \exp\left(a \frac{\mathrm{d}}{\mathrm{d}x}\right) f(x)$$

where we define exp through its Taylor series so

$$\exp\left(a \frac{\mathrm{d}}{\mathrm{d}x}\right) = 1 + a \frac{\mathrm{d}}{\mathrm{d}x} + \frac{a^2}{2!} \frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{a^3}{3!} \frac{\mathrm{d}^3}{\mathrm{d}x^3} + \cdots.$$

This is a standard way to extend functions to arguments other than numbers. For example we can define the exponential of a square matrix, A, as

$$\exp(A) = I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \cdots.$$

11.2   Three-Dimensional Taylor Series

Consider some field, φ. Let

$$F(t) = \varphi(\boldsymbol{r} + t\boldsymbol{a}) = \varphi(\boldsymbol{u})$$

where $\boldsymbol{r}$ is the position, $\boldsymbol{a}$ is some constant vector, $t \in \mathbb{R}$, and $\boldsymbol{u} = \boldsymbol{r} + t\boldsymbol{a}$. We can Taylor expand F as

$$F(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} F^{(n)}(0).$$

Our interest is in $F(0) = \varepsilon(\boldsymbol{r} + \boldsymbol{a})$. First consider

$$F'(t) = \frac{\partial \varphi}{\partial t} = \frac{\partial \varphi}{\partial u_i} \frac{\partial u_i}{\partial t} = \frac{\partial \varphi}{\partial u_i} \frac{\partial}{\partial t}(r_i + t a_i) = \frac{\partial \varphi}{\partial u_i} a_i = (\boldsymbol{a} \cdot \nabla_{\boldsymbol{u}}) \varphi(\boldsymbol{u})$$

where $(\nabla_{\boldsymbol{u}})_i = \partial_{u_i}$. Hence we can write

$$F^{(n)}(t) = (\boldsymbol{a} \cdot \nabla_{\boldsymbol{u}})^n \varphi(\boldsymbol{u})$$

and so

$$F^{(n)}(0) = (\boldsymbol{a} \cdot \nabla_{\boldsymbol{r}})^n \varphi(\boldsymbol{r}).$$

Hence

$$F(1) = \varphi(\boldsymbol{r} + \boldsymbol{a}) = \sum_{n=0}^{\infty} \frac{1}{n!} (\boldsymbol{a} \cdot \nabla_{\boldsymbol{r}}) \varphi(\boldsymbol{r}) = \exp(\boldsymbol{a} \cdot \nabla_{\boldsymbol{r}}) \varphi(\boldsymbol{r}).$$

Or in index notation:

$$\varphi(\boldsymbol{r} + \boldsymbol{a}) = \sum_{n=0}^{\infty} \frac{1}{n!} a_i \partial_i \varphi = \exp(a_i \partial_i) \varphi.$$

11.3   Multipole Expansion

Recall that the force on a charge, q, at $\boldsymbol{r}$ due to a charge, q_1, at $\boldsymbol{r_1}$ is

$$\boldsymbol{F} = \frac{1}{4\pi\varepsilon_0} \frac{q q_1 (\boldsymbol{r} - \boldsymbol{r_1})}{|\boldsymbol{r} - \boldsymbol{r_1}|^3}.$$

The electric field is then defined as

$$\boldsymbol{E}(\boldsymbol{r}) = \lim_{q \to 0^+} \frac{1}{q} \boldsymbol{F}.$$

So

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{q_1 (\boldsymbol{r} - \boldsymbol{r_1})}{|\boldsymbol{r} - \boldsymbol{r_1}|^3}.$$

If we add in more charges, q_i, at $\boldsymbol{r_i}$, then the resulting field is

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0} \sum_i \frac{q_i(\boldsymbol{r} - \boldsymbol{r_i})}{|\boldsymbol{r} - \boldsymbol{r_i}|^3}.$$

If the charge distribution becomes continuous then

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0} \int_V \mathrm{d}^3 r' \, \rho(\boldsymbol{r}') \frac{\boldsymbol{r} - \boldsymbol{r}'}{|\boldsymbol{r} - \boldsymbol{r}'|^3}.$$

Here $\rho(\boldsymbol{r}')$ is the charge density at $\boldsymbol{r}'$. We define the electrostatic potential, φ, as the scalar field such that

$$\boldsymbol{E} = -\nabla\varphi$$

which means

$$\varphi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0} \int_V \mathrm{d}^3 r' \, \frac{\rho(\boldsymbol{r}')}{|r - \boldsymbol{r}'|}.$$

The multipole expansion is a Taylor series expressing the electrostatic potential far from the charge distribution. We start by considering the expansion of $1/|\boldsymbol{r} + \boldsymbol{a}|$ for $r \gg a$:

$$\begin{aligned}
\frac{1}{|\boldsymbol{r} + \boldsymbol{a}|} &= \sum_{n=0}^{\infty} \frac{1}{n!} (\boldsymbol{a} \cdot \nabla_{\boldsymbol{r}})^n \frac{1}{r} \\
&= \frac{1}{r} + \frac{1}{1!} a_i \partial_i \frac{1}{r} + \frac{1}{2!} (a_i \partial_i a_j \partial_j) \frac{1}{r} + \cdots \\
&= \frac{1}{r} - \frac{\boldsymbol{a} \cdot \boldsymbol{r}}{r^3} + \frac{3(\boldsymbol{a} \cdot \boldsymbol{r}) - a^2 r^2}{2r^5} + \mathcal{O}(r^{-4}).
\end{aligned}$$

Hence

$$\begin{aligned}
\varphi(\boldsymbol{r}) &= \frac{1}{4\pi\varepsilon_0} \int_V \mathrm{d}^3 r' \, \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \\
&= \frac{1}{4\pi\varepsilon_0} \int_V \mathrm{d}^3 r' \, \rho(\boldsymbol{r}') \left[\frac{1}{r} + \frac{\boldsymbol{r}' \cdot \boldsymbol{r}}{r^3} + \frac{3(\boldsymbol{r}' \cdot \boldsymbol{r}) - r'^2 r^2}{r^5} + \cdots \right] \\
&= \frac{1}{4\pi\varepsilon_0} \left[\frac{Q}{r} + \frac{r_i p_i}{r^3} + \frac{r_i r_j \mathcal{Q}_{ij}}{2r^5} + \cdots \right].
\end{aligned}$$

Here we have defined the total charge or monopole term:

$$Q = \int_V \mathrm{d}^3 r' \, \rho(\boldsymbol{r}'),$$

the dipole moment:

$$p_i = \int_V \mathrm{d}^3 r' \, r_i' \rho(\boldsymbol{r}'),$$

and the quadrupole term:

$$\mathcal{Q}_{ij} = \int_V \mathrm{d}^3 r' \, (3 r_i' r_j' - r'^2 \delta_{ij}) \rho(\boldsymbol{r}').$$

This expansion, known as the multipole expansion, is valid in the far field, i.e. $r \gg r'$. It is possible to go to higher order and define an octopole moment and so on. The expansion is usually dominated by the lowest order non-zero term:

- If $Q \neq 0$ then far from the charge it approximates a point charge at the origin:

$$\varphi(\boldsymbol{r}) \approx \frac{1}{4\pi\varepsilon_0} \frac{Q}{r}.$$

- If $Q = 0$ and $\boldsymbol{p} \neq \boldsymbol{0}$ then the dipole term dominates and

$$\varphi(\boldsymbol{r}) \approx \frac{1}{4\pi\varepsilon_0} \frac{\boldsymbol{p} \cdot \boldsymbol{r}}{r^3}.$$

This is called a dipole moment as if we consider the charge distribution

$$\rho(\boldsymbol{r}) = q[\delta(\boldsymbol{r} - \boldsymbol{d}) - \delta(\boldsymbol{r})]$$

due to opposite charges of magnitude q at the origin and at $\boldsymbol{d}$ then we have

$$\boldsymbol{p} = \int_V \mathrm{d}^3 r' \, \boldsymbol{r}' q[\delta(\boldsymbol{r}' - \boldsymbol{d}) - \delta(\boldsymbol{r}')] = q\boldsymbol{d}$$

which is the dipole moment of a classic dipole.

- If $Q = 0$ and $\boldsymbol{p} = \boldsymbol{0}$ then the quadrupole term dominates and

$$\varphi(\boldsymbol{r}) \approx \frac{1}{4\pi\varepsilon_0} \frac{r_i r_j \mathcal{Q}_{ij}}{2r^5}.$$

$\mathcal{Q}$ is symmetric ($\mathcal{Q}_{ij} = \mathcal{Q}_{ji}$) and traceless ($\mathrm{Tr}\,\mathcal{Q} = 0$). The simplest example is four charges of magnitude q placed on the corners of a rectangle so that diagonally opposite charges have the same sign and neighbouring charges have opposite signs. This gives the charge distribution

$$\rho(\boldsymbol{r}) = q[\delta(\boldsymbol{r} - \boldsymbol{a}) - \delta(\boldsymbol{r}) - \delta(\boldsymbol{r} - \boldsymbol{a} - \boldsymbol{b}) + \delta(\boldsymbol{r} - \boldsymbol{b})]$$

which gives the quadrupole moment

$$\mathcal{Q}_{ij} = 2q \left[(\boldsymbol{a} \cdot \boldsymbol{b})\delta_{ij} - \frac{3}{2}(a_i b_j + a_j b_i) \right].$$

Due to similarities between the electric field and gravitational field the multipole expansion is also valid if we swap $\boldsymbol{E} \to \boldsymbol{g}$, $1/(4\pi\varepsilon_0) \to G$ and $q \to m$. However we lose the existence of negative mass and the analogy of a dipole breaks down even though we still get the same gravitational potential, φ.

Part II

Linear Elasticity

12   The Deformation Tensor

12.1   Deformation Types

There are many ways that a body can be deformed. Some of the most common are shears, compressions, and stretches. These are shown in figure 12.1. Very rarely does only one of these deformations occur. For example if you compress an object along one axis the matter has to go somewhere so it tends to increase in size along the other axes. We aim to come up with a way to describe a general deformation. In this course we are interested in linear elasticity which means that we assume all deformations are sufficiently small that things can safely be expanded only up to linear order in the relevant variable.

12.2   Geometry of Deformations

A body in equilibrium has its position, orientation, and shape set by the positions, $\{\boldsymbol{r}\}$, of its material points. What counts as a material point exactly isn't important, its simply a point that we can track as a body moves and deforms, it could be an atom or molecule or it might be a region of material.

Suppose the body is then deformed in such a way that a material point that was at $\boldsymbol{r}$ before the deformation is now at $\boldsymbol{r}'$. We define the **displacement vector**, $\boldsymbol{u}$, as the displacement of the point under this deformation:

$$\boldsymbol{u} = \boldsymbol{r} - \boldsymbol{r}'.$$

In a translation or rotation the shape of the body is fixed and therefore $\boldsymbol{u}$ is the same for all points. In a deformation however the shape of the body can change and therefore $\boldsymbol{u}$ is not the same for all points.
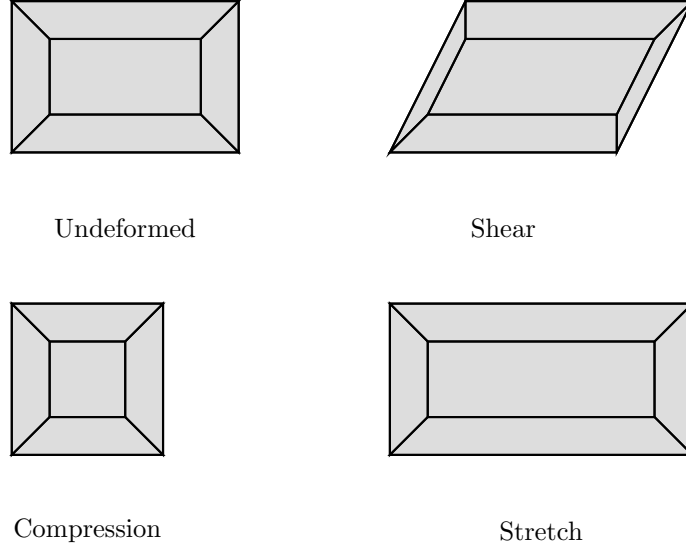
Figure 12.1: Common deformations

We can think of $\boldsymbol{u}$ as a vector field mapping points to there deformed points. This introduces a double dependence on $\boldsymbol{r}$: first $\boldsymbol{r}$ is the point that is being moved and second $\boldsymbol{r}$ is the point at which we are evaluating $\boldsymbol{u}(\boldsymbol{r})$. By simply rearranging the definition of $\boldsymbol{u}$ we have $\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r})$ and $\{\boldsymbol{r}'\}$ gives us the shape, orientation, and position of the deformed body.

Consider two points, $\boldsymbol{r}$ and $\boldsymbol{r} + \mathbf{d}\boldsymbol{r}$, on the undeformed body. The Cartesian coordinates of these points are then x_i and $x_i + \mathrm{d}x_i$ respectively. After the deformation we have

$$x_i \mapsto x_i + u_i(\boldsymbol{r}),$$
$$x_i + \mathrm{d}x_i \mapsto x_i + \mathrm{d}x_i + u_i(\boldsymbol{r} + \mathbf{d}\boldsymbol{r}).$$

The distance between the two points before the deformation is $\mathrm{d}l$, which by Pythagoras' theorem is simply

$$\mathrm{d}l^2 = \mathrm{d}x_1{}^2 + \mathrm{d}x_2{}^2 + \mathrm{d}x_3{}^2 = \mathrm{d}x_i{}^2 \,.$$

Note that even though the i only appears once since it appears in a squared term we view it as $\mathrm{d}x_i \, \mathrm{d}x_i$ so there is an implied summation. The distance between the two points after the deformation is $\mathrm{d}l'$. To calculate this we similarly need to sum the square of the differences between the coordinates:

$$\mathrm{d}l'^2 = [(x_i + \mathrm{d}x_i + u_i(\boldsymbol{r} + \mathbf{d}\boldsymbol{r})) - (x_i + u_i(\boldsymbol{r}))]^2$$
$$= [\mathrm{d}x_i + u_i(\boldsymbol{r} + \mathbf{d}\boldsymbol{r}) - u_i(\boldsymbol{r})]^2.$$

We can Taylor expand u_i about $\boldsymbol{r}$ to first order as

$$u_i(\boldsymbol{r} + \mathbf{d}\boldsymbol{r}) = u_i(\boldsymbol{r}) + \frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_k + \mathcal{O}\big(\mathrm{d}x_k{}^2\big) \implies u_i(\boldsymbol{r} + \mathbf{d}\boldsymbol{r}) - u_i(\boldsymbol{r}) \approx \frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_k \,.$$

Using this we have

$$\mathrm{d}l'^2 \approx \left[\mathrm{d}x_i + \frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_k\right]^2$$
$$= \mathrm{d}x_i{}^2 + 2\frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_i \, \mathrm{d}x_k + \frac{\partial u_i}{\partial x_k}\frac{\partial u_i}{\partial x_j} \, \mathrm{d}x_k \, \mathrm{d}x_j \,.$$

The first term is simply $\mathrm{d}l^2$. The second term can be expanded as

$$2\frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_i \, \mathrm{d}x_k = \frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_i \, \mathrm{d}x_k + \frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_i \, \mathrm{d}x_k$$

renaming $i \leftrightarrow k$ in the second term only we have

$$= \frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_i \, \mathrm{d}x_k + \frac{\partial u_k}{\partial x_k} \, \mathrm{d}x_k \, \mathrm{d}x_i$$

$$= \frac{\partial u_i}{\partial x_k} \, \mathrm{d}x_i \, \mathrm{d}x_k + \frac{\partial u_k}{\partial x_k} \, \mathrm{d}x_i \, \mathrm{d}x_k \,.$$

For the final term if we rename $i \to l$ and $j \to i$ we have

$$\frac{\partial u_i}{\partial x_k} \frac{\partial u_i}{\partial x_j} \, \mathrm{d}x_k \, \mathrm{d}x_j = \frac{\partial u_l}{\partial x_k} \frac{\partial u_l}{\partial x_k} \, \mathrm{d}x_k \, \mathrm{d}x_k$$

Putting this into $\mathrm{d}l'^2$ we have

$$\mathrm{d}l'^2 = \mathrm{d}l^2 + \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right) \mathrm{d}x_i \, \mathrm{d}x_k + \frac{\partial u_l}{\partial x_i} \frac{\partial u_l}{\partial x_k}$$

$$= \mathrm{d}l^2 + \left[\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \frac{\partial u_l}{\partial x_i} \frac{\partial u_l}{\partial x_k} \right] \mathrm{d}x_i \, \mathrm{d}x_k$$

$$= \mathrm{d}l^2 + 2u_{ik} \, \mathrm{d}x_i \, \mathrm{d}x_k$$

where

$$\boxed{u_{ik} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \frac{\partial u_l}{\partial x_k} \frac{\partial u_l}{\partial x_k} \right)}$$

are the coordinates of a rank 2 Cartesian tensor, u, called the **deformation tensor** or the **strain tensor**. Note that the factor of $1/2$ and therefore factor of 2 in $\mathrm{d}l'^2$ are simply a matter of convention. Note that sometimes the last term is dropped as it is of higher order so not as important for small deformations. The notation here can be a bit confusing as we use u for the displacement vector and the deformation tensor. The two uses can be distinguished by the rank of the object in question.

The components, u_{ik}, are symmetric in i and k ($u_{ik} = u_{ki}$). This means that there are only 6 free components and the other three are fixed by symmetry. One corollary of this si that u can always be diagonalised. Let $u^{(i)}$ be the eigenvalues, or principle components, of u. Then there in the eigenbasis, $\{u^{(i)}\}$, also known as the principle axes, we have

$$u = \begin{pmatrix} u^{(1)} & 0 & 0 \\ 0 & u^{(2)} & 0 \\ 0 & 0 & u^{(3)} \end{pmatrix}.$$

Suppose we work in the eigenbasis. Then

$$\mathrm{d}l'^2 = \mathrm{d}l^2 + 2u_{ik} \, \mathrm{d}x_i \, \mathrm{d}x_k$$

$$= \mathrm{d}x_i{}^2 + 2u_{ik} \, \mathrm{d}x_i \, \mathrm{d}x_k$$

$$= \delta_{ik} \, \mathrm{d}x_i \, \mathrm{d}x_k + 2u_{ik} \, \mathrm{d}x_i \, \mathrm{d}x_k$$

$$= (\delta_{ik} + 2u_{ik}) \, \mathrm{d}x_i \, \mathrm{d}x_k$$

$$= (1 + 2u_{11}) \, \mathrm{d}x_1{}^2 + (1 + 2u_{22}) \, \mathrm{d}x_2{}^2 + (1 + 2u_{33}) \, \mathrm{d}x_3{}^2$$

$$= (1 + 2u^{(1)}) \, \mathrm{d}x_1{}^2 + (1 + 2u^{(2)}) \, \mathrm{d}x_2{}^2 + (1 + 2u^{(3)}) \, \mathrm{d}x_3{}^2$$

So the proportional change of separation along each axis is

$$\mathrm{d}xi' = \sqrt{1 + 2u^{(i)}} \, \mathrm{d}x_i$$

and hence

$$\frac{\mathrm{d}x_i' = \mathrm{d}x_i}{\mathrm{d}x_i} = \sqrt{1 + 2u^{(i)}} - 1.$$

Since the deformation is small $u^{(i)}$ must be small and so we expand this in terms of $u^{(i)}$:

$$\sqrt{1 + 2u^{(i)}} = (1 + 2u^{(i)})^{1/2} \approx 1 + u^{(i)} + \mathcal{O}\left(u^{(i)2} \right)$$

so we have

$$\mathrm{d}x_i' \approx (1 + u^{(i)}) \, \mathrm{d}x_i \,, \qquad \text{and} \qquad \frac{\mathrm{d}x_i' - \mathrm{d}x_i}{\mathrm{d}x_i} \approx u^{(i)}.$$

Suppose we wish to know the volume change. Consider a small volume, $\mathrm{d}V = \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3$, in the undeformed body. After the deformation this becomes

$$
\begin{aligned}
\mathrm{d}V' &= (1 + u^{(1)})(1 + u^{(2)})(1 + u^{(3)})\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 \\
&\approx (1 + u^{(1)} + u^{(2)} + u^{(3)})\,\mathrm{d}V \\
&= [1 + \mathrm{Tr}(u)]\,\mathrm{d}V
\end{aligned}
$$

where we have dropped any terms of second order or higher. Hence we have

$$
\frac{\mathrm{d}V' - \mathrm{d}V}{\mathrm{d}V} = \mathrm{Tr}\,u.
$$

This is a useful result because the trace is invariant under a coordinate transform, this can be shown by considering a transformation $L \in \mathrm{SO}(3)$:

$$
A'_{ii} = l_{i\alpha}l_{i\beta}A_{\alpha\beta} = L^{\mathsf{T}}_{\alpha i}L_{i\beta}A_{\alpha\beta} = (L^{\mathsf{T}}L)_{\alpha\beta}A_{\alpha\beta} = \delta_{\alpha\beta}A_{\alpha\beta} = A_{\alpha\alpha} = \mathrm{Tr}(A).
$$

Example 12.1. *Consider a cuboid that is stretched along the x_1 axis so the length increases from l to $l + \Delta l$ for some fractional change, α. Suppose that this also causes the cuboid to contract along the other two axes by some amount such that the fractional change, β, is the same for both axes. What is the deformation tensor for this deformation?*

First consider the x_1 axis:

$$
\frac{\Delta l}{l} = \frac{x'_i - x_i}{x_i} = \frac{u^{(1)}}{x_i} = \alpha.
$$

So $u^{(1)} = \alpha x_1$. Now consider the x_2 axis. If we think of this as the height, h, of the cuboid then we have

$$
\frac{\Delta h}{h} = \frac{x'_2 - x_2}{x_2} = \frac{u^{(2)}}{x_2} = -\beta.
$$

So $u^{(2)} = -\beta x_2$, notice that this is negative since the cuboid is compressed along this axis. Similarly if we think about the x_3 axis as width, w, of the cuboid we have

$$
\frac{\Delta h}{h} = \frac{x'_3 - x_3}{x_3} = \frac{u^{(3)}}{x_3} = -\beta.
$$

So $u^{(3)} = -\beta x_3$. Hence the deformation tensor in the principle axis basis has components

$$
u = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & -\beta & 0 \\ 0 & 0 & -\beta \end{pmatrix}.
$$

---

13   The Stress Tensor

An undeformed body is in equilibrium with its surroundings. This means that all parts of the body are in equilibrium with each other. The total force and total torque the body are zero. When the body is deformed by external forces it is moved away from equilibrium creating internal forces in the body.

Consider a small portion of the body. Any forces acting on this portion must have a net effect of acting only at the surface as forces within the body cancel out. Let $\boldsymbol{F}$ be the force per unit volume, or force density, acting on the volume. The total force on the volume is simply

$$
\int_V \boldsymbol{F}\,\mathrm{d}V\,.
$$

Since the force acts only at the surface there exists a tensor, σ_{ij}, such that we can write the total force as a surface integral,

$$
\int_V F_i\,\mathrm{d}V = \int_{\partial V} \sigma_{ij}\,\mathrm{d}A_j\,, \tag{13.1}
$$

where ∂V is the boundary of the volume and $\mathrm{d}A_j = n_j \, \mathrm{d}A$ is an area element along the surface normal $\hat{\boldsymbol{n}}$. By the divergence theorem we can then write

$$\int_{\partial V} \sigma_{ij} \, \mathrm{d}A_j = \int_V \frac{\partial \sigma_{ij}}{\partial x_j} \, \mathrm{d}V \, .$$

Hence

$$\boxed{F_i = \frac{\partial \sigma_{ij}}{\partial x_j}.}$$

In conclusion if the only forces acting on a body act on its surface then the internal forces in the body can be written as the divergence of a second rank tensor, σ, which is called the **stress tensor**.

For a physical understanding of what σ represents consider equation 13.1. This gives us

$$\mathrm{d}F_i = \sigma_{ij} \, \mathrm{d}A_j = \sigma_{ij} n_j \, \mathrm{d}A$$

where $\mathrm{d}F_i = F_i \, \mathrm{d}V$ is the force on the volume element $\mathrm{d}V$ in the direction x_i. From this we can see that σ_{ij} is the force per unit area in the direction x_i acting on the surface element normal to x_j. This can be seen more clearly in figure 13.1 which shows the stress tensor on a cube. Each component of σ_{i3} is proportional to the force acting on the face normal to x_3 and the force acts in the x_i direction.
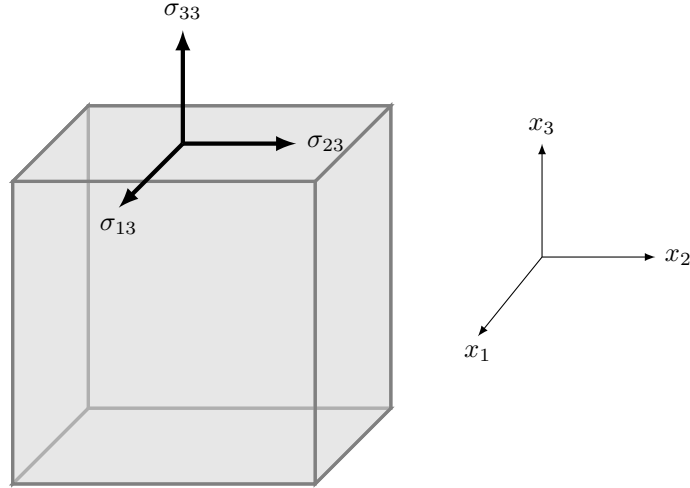

Figure 13.1: The stress tensor

Now consider the torque that acts on the volume as a result of the neighbouring volume elements inside the body. This is given by

$$\left[\int \boldsymbol{r} \times \boldsymbol{F} \, \mathrm{d}V \right]_i = \int \varepsilon_{ijk} x_j F_k \, \mathrm{d}V$$

$$= \int (x_2 F_3 - x_3 F_2)\delta_{i1} \, \mathrm{d}V + \int (x_3 F_1 - x_1 F_3)\delta_{i2} \, \mathrm{d}V + \int (x_1 F_2 - x_2 F_1)\delta_{i3} \, \mathrm{d}V \, .$$

Consider just one of these terms:

$$\int (x_i F_j - x_j F_i) \, \mathrm{d}V = \int \left(x_i \frac{\partial \sigma_{jk}}{\partial x_k} - x_j \frac{\partial \sigma_{ik}}{\partial x_k} \right) \mathrm{d}V$$

$$= \int \frac{\partial}{\partial x_k}(x_i \sigma_{ik} - x_j \sigma_{ik}) \, \mathrm{d}V - \int \left(\frac{\partial x_i}{\partial x_k}\sigma_{jk} - \frac{\partial x_j}{\partial x_k}\sigma_{ik} \right) \mathrm{d}V$$

$$= \int (x_i \sigma_{jk} - x_j \sigma_{ik}) \, \mathrm{d}A_k - \int (\delta_{ik}\sigma_{jk} - \delta_{jk}\sigma_{ik}) \, \mathrm{d}V$$

$$= \int (x_i \sigma_{jk} - x_j \sigma_{ik}) \, \mathrm{d}A_k - \int (\sigma_{ji} - \sigma_{ij}) \, \mathrm{d}V \, .$$

Here we have again used the divergence theorem to replace a volume integral with a surface integral. The same logic that we applied to the forces here leads us to the conclusion that the torque must be able to be written as a surface integral. Therefore we require the second term to be zero. Since this must be the case for an arbitrary volume element we have

$$\sigma_{ij} = \sigma_{ji}$$

so the stress tensor is symmetric.

In the absence of external forces there is no net force applied to any volume element and so

$$F_i = \frac{\partial \sigma_{ij}}{\partial x_j} = 0 \iff \operatorname{div} \sigma = 0.$$

If there is an external force that acts on all volume elements then

$$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0$$

where f_i is the force per unit volume in the x_i direction. For example if a body experiences gravity along the x_2 axis then we have

$$\frac{\partial \sigma_{1j}}{\partial x_j} = \frac{\partial \sigma_{2j}}{\partial x_j} - mg = \frac{\partial \sigma_{3j}}{\partial x_j} = 0.$$

The boundary equations for this differential equation come from specifying the force applied to each surface of the body. For each surface we must specify the force, P_i, applied to area A along x_j which is given by $P_i = \sigma_{ij} n_j A$ where $\hat{\boldsymbol{n}}$ is the surface normal.

Example 13.1. *Consider a cylinder stretched by two equal and opposite forces applied to its ends. What is the stress tensor?*

Figure 13.2: Cylinder being stretched.

Figure 13.2 shows the setup. Let a be the end with force $-P$ per unit area and b be the end with force P per unit area. Let the x_1-axis run from a to b, the x_2-axis be into the page and the x_3 axis be up. Then the boundary conditions are as follows:

- At b the outer normal is $\boldsymbol{e_1}$ and so $\sigma_{ij}(b)\delta_{j1}A = P\delta_{i1}$ where A is the area of the circle.

- At a the outer normal is $-\boldsymbol{e_1}$ and so $-\sigma_{ij}(a)\delta_{j1}A = -P\delta_{i1}$.

- On the curved surface of the cylinder $\hat{\boldsymbol{n}}$ is perpendicular to $\boldsymbol{e_1}$ and so $\hat{\boldsymbol{n}} = \alpha \boldsymbol{e_2} + \beta \boldsymbol{e_3}$. Since no force acts in this direction we have $\sigma_{ij} n_j \tilde{A} = 0$ where $\tilde{A}$ is the area of the element of the curved surface that we are currently considering. Expanding these we have

$$\alpha \sigma_{12} + \beta \sigma_{13} = 0,$$
$$\alpha \sigma_{22} + \beta \sigma_{23} = 0,$$
$$\alpha \sigma_{32} + \beta \sigma_{33} = 0,$$

Since, in general, there are surface elements where $\alpha, \beta \neq 0$ we must have $\sigma_{ij} = 0$ for $i, j \neq 1$. Hence

$$\sigma_{ij} = \begin{pmatrix} P/A & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Example 13.2. *What is the stress tensor for an isotropic compression?*

An isotropic compression involves the same force on each face. Let p is the force per unit area. For a compression p points into the surface, i.e. anti-parallel to the surface normal. Also the only force on a face is perpendicular to that face so the off diagonal elements of the stress tensor are zero. Hence $\sigma_{ij} n_j A = -p \delta_{ik} n_k A = -pI$.

$$\sigma_{ij} = \begin{pmatrix} -p & 0 & 0 \\ 0 & -p & 0 \\ 0 & 0 & -p \end{pmatrix}.$$

Example 13.3. *What is the stress tensor for a shear of the surface normal to x_2 along the x_1 direction?*

The only force is applied to the surface with surface normal $\boldsymbol{e_2}$. If the force is F and it is in the x_1 direction then we must have $\sigma_{12} = F$. Since the stress tensor is symmetric we must also have $\sigma_{21} = F$. Hence

$$\sigma_{ij} = \begin{pmatrix} 0 & F & 0 \\ F & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

13.1   Relation Between Deformation and Stress Tensors

Suppose we have a medium being deformed under some external force so that the stress tensor is σ and the deformation tensor is u. We want to find a relationship between σ and u. Imagine we perform some extra work on the system changing the displacement vector $\boldsymbol{u}$ to $\boldsymbol{u} + \boldsymbol{\delta u}$ (u_i to $u_i + \delta u_i$). We are free to choose how to do this work and so we choose to do it in a way such that δu_i vanishes at the surface. The work done by internal forces is given by

$$\int \delta W \, \mathrm{d}V = \int \boldsymbol{F} \cdot \boldsymbol{\delta u} \, \mathrm{d}V$$

where δW is the work done on a single volume element and $\boldsymbol{F}$ is the internal force that opposes the deformation. By the definition of the stress tensor we have

$$\int \delta W \, \mathrm{d}V = \int \frac{\partial \sigma_{ik}}{\partial x_k} \delta u_i \, \mathrm{d}V = \int \frac{\partial}{\partial x_k}(\delta u_i \sigma_{ik}) \, \mathrm{d}V - \int \sigma_{ik} \frac{\partial}{\partial x_k} \delta u_i.$$

We can write the first integral as a surface integral using the divergence theorem. We can write the second integral in a symmetric form.

$$\int \delta W \, \mathrm{d}V = \int \sigma_{ik} \delta u_i n_k \, \mathrm{d}A - \frac{1}{2} \int \left(\sigma_{ik} \frac{\partial}{\partial x_k} \delta u_i + \sigma_{ki} \frac{\partial}{\partial x_i} \delta u_k \right) \mathrm{d}V.$$

Since we chose δu_i to vanish at the surface the first integral vanishes. For the second integral we exploit the symmetry of the stress tensor:

$$\int \delta W \, \mathrm{d}V = -\frac{1}{2} \int \sigma_{ik} \left(\frac{\partial}{\partial x_k} \delta u_i + \frac{\partial}{\partial x_i} \delta u_k \right) \mathrm{d}V = -\int \sigma_{ik} \delta u_{ik} \, \mathrm{d}V$$

where

$$\delta u_{ik} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right).$$

Notice that this is simply the deformation tensor after dropping terms of second order.

14   Stress and Strain

Recall that in the last section we derived the result

$$\int \delta W \, \mathrm{d}V = -\int \sigma_{ik} \delta u_{ik} \, \mathrm{d}V.$$

The free energy is defined as

$$\mathcal{F} = U - TS$$

where U is the internal energy, T is the temperature, and S is the entropy. The value derived above is the work done, which is equal to the negative of the internal energy. The infinitesimal free energy change due to the deformation is

$$d\mathcal{F} = d(U - TS) = dU - S\,dT$$

using $dU = -\delta W$ we have

$$d\mathcal{F} = -\delta W - S\,dT = \sigma_{ik}\delta u_{ik} - S\,dT\,.$$

At constant temperature we have $dT = 0$ and so

$$d\mathcal{F} = \sigma_{ik}\delta u_{ik}.$$

Hence

$$\sigma_{ik} = \left(\frac{\partial \mathcal{F}}{\partial u_{ik}}\right)_T.$$

Hence $\mathcal{F}$ is a function of the deformation tensor, u. For a homogenous, isotropic medium, which is the only sort we are interested in here, there is no preferred direction and hence $\mathcal{F}$ can only depend on quantities that are invariant under rotations. These quantities are $\operatorname{Tr} u = u_{ii}$, $u_{ik}u_{ki}$, and $u_{ij}u_{jk}u_{ki}$. All higher order tensor invariants can be written in terms of these quantities. Further we consider only small deformations and therefore neglect any terms that appear higher than second order. This leaves us with a generalised version of Hooke's law, which states that the energy stored in a spring with spring constant k is $kx^2/2$ where x is the displacement of the spring from equilibrium. We can think of this simply as Taylor expanding $\mathcal{F}$ in terms of u_{ik}. Finally notice that two deformations with u_{ij} and $-u_{ij}$ produce the same free energy change and hence u can only appear in even powers since the transformation $u \to -u$ cannot be allowed to introduce a negative. Hence the free energy is given by

$$\mathcal{F} = \frac{\lambda}{2}u_{ii}^2 + \mu u_{ik}^2$$

where we have used the fact that u is symmetric so $u_{ik}u_{ki} = u_{ik}u_{ik} = u_{ik}^2$, note that we still sum over i and k. Here the constants of proportionality are $\lambda/2$ and μ. λ and μ are called the Lamè coefficients.

We saw in examples 13.2 and 13.3 that for an isotropic compression the stress tensor is diagonal. Similarly for a shear the stress tensor will have zeros on the diagonal. Using this we can split any stress tensor into two parts. A diagonal tensor for the compression component and a tensor with only non-zero elements off the diagonal for the shear component.

Now consider differentials of the free energy:

$$d\mathcal{F} = \frac{\lambda}{2}\,d(u_{ll}^2) + 2\mu\,du_{ik}^2$$
$$= \lambda u_{ll}\,du_{ll} + 2\mu u_{ik}\,du_{ik}$$

Consider the differential of the trace:

$$du_{ll} = \frac{\partial u_{ll}}{\partial u_{ik}}\,du_{ik}\,.$$

The derivative can be computed easily by first noting that

$$\frac{\partial u_{ab}}{\partial u_{ik}} = \delta_{ia}\delta_{kb}$$

and so

$$\frac{\partial u_{ll}}{\partial u_{ik}} = \delta_{il}\delta_{kl} = \delta_{ik}.$$

Hence

$$du_{ll} = \delta_{ik}\,du_{ik}\,.$$

Substituting this into the differential of the free energy we have

$$d\mathcal{F} = \lambda u_{ll}\delta_{ik}\,du_{ik} + 2\mu u_{ik}\,du_{ik}\,.$$

Dividing through by $\mathrm{d}u_{ik}$ and assuming convergence etc. we have

$$\sigma_{ik} = \frac{\partial \mathcal{F}}{\partial u_{ik}} = \lambda u_{ll}\delta_{ik} + 2\mu u_{ik}.$$

By defining $\lambda = K - 2\mu/3$ we can write this as

$$\sigma_{ik} = \left(K - \frac{2}{3}\mu\right)\delta_{ik}u_{ll} + 2\mu u_{ik}.$$

We will see a physical interpretation for the constant K later.

The equations above allow us to compute the stress tensor for a given deformation tensor assuming we know the values of the relevant constants. We would also like to be able to reverse this and work out what deformation caused a given stress. To this end consider the trace of the stress tensor:

$$\sigma_{ll} = \left[3\left(K - \frac{2}{3}\mu\right) + 2\mu\right]u_{ll} = 3Ku_{ll}.$$

Hence

$$u_{ll} = \frac{1}{3K}\sigma_{ll}.$$

Substituting this into the equation for σ_{ik} in terms of u and rearranging gives

$$u_{ik} = \frac{1}{2\mu}\left[\sigma_{ik} - \left(K - \frac{2}{3}\mu\right)\frac{1}{3K}\delta_{ik}\sigma_{ll}\right] = \frac{1}{2\mu}\sigma_{ik} - \frac{1}{6\mu}\delta_{ik}\sigma_{ll} + \frac{1}{9K}\delta_{ik}\sigma_{ll}.$$

At this point we note that a given deformation can be split into two parts, one which characterises the shear part of the deformation and the other characterises the compression/extension part. The shear part is given by all of the off diagonal elements whereas the compression is given by the diagonal parts:

$$u_{ik} = \underbrace{\left(u_{ik} - \frac{1}{3}\delta_{ik}u_{ll}\right)}_{\text{shear}} + \underbrace{\frac{1}{3}\delta_{ik}u_{ll}}_{\text{compression}}.$$

Notice the factor of $1/3$ which occurs since $\delta_{ii} = 3$. We can then write the free energy as

$$\mathcal{F} = \mu\left(u_{ik} - \frac{1}{3}\delta_{ik}u_{ll}\right)^2 + \frac{K}{2}u_{ll}^2.$$

This gives us a physical interpretation for K and μ. K, called the **bulk modulus** or **compressability**, characterises how hard it is to compress a material. Similarly μ, called the **shear modulus**, characterises how hard it is to shear the material.

We saw in example 13.2 that a compression of force p per unit area leads to a stress tensor $\sigma_{ik} = -p\delta_{ik}$. Hence the deformation tensor is

$$\begin{aligned}
u_{ik} &= \frac{1}{2\mu}\sigma_{ik} - \frac{1}{6\mu}\delta_{ik}\sigma_{ll} + \frac{1}{9K}\delta_{ik}\sigma_{ll} \\
&= \frac{1}{2\mu}(-p\delta_{ik}) - \frac{1}{6\mu}\delta_{ik}(-p\delta_{ll}) + \frac{1}{9K}\delta_{ik}(-p\delta_{ll}) \\
&= -\frac{1}{2\mu}p\delta_{ik} + \frac{1}{2\mu}\delta_{ik}p - \frac{1}{3K}\delta_{ik}p \\
&= -\frac{1}{3K}\delta_{ik}p.
\end{aligned}$$

So

$$u = \begin{pmatrix} -p/3K & 0 & 0 \\ 0 & -p/3K & 0 \\ 0 & 0 & -p/3K \end{pmatrix}.$$

Recall that the trace of the deformation tensor gives the relative volume change:

$$u_{ll} = -\frac{p}{K} = \frac{\delta V}{V}.$$

This shows that increasing K means that we need to increase the pressure to get the same volume change, which makes sense as K is the compressibility of the material. If we interpret this as a pressure change rather than an absolute pressure then we can rearrange to get

$$\frac{1}{K} = -\frac{1}{V}\frac{\partial V}{\partial P}$$

which should be familiar from thermodynamics. In this case the change in free energy is given by

$$\mathcal{F} = \frac{p^2}{2K}$$

where we have chosen to define zero free energy as the starting free energy of the system. Notice from this that K, and by extension μ and λ have units of energy per metre cubed which is the same as pascals.

15   Stress–Strain Examples

Example 15.1. *Consider a cylinder with its axis along the x_1 direction. Apply a force per unit area, P, to one end along the x_1 direction. What is the stress and strain?*

The stress is easy to compute, there is force only on one face and in one direction so the stress is zero for all components apart from $\sigma_{11} = P$. The deformation tensor in general is given by

$$u_{ik} = \frac{1}{2\mu}\left[\sigma_{ik} - \left(K - \frac{2}{3}\mu\right)\frac{1}{3K}\delta_{ik}\sigma_{ll}\right].$$

In the case of $\sigma_{11} = P$ being the only non-zero component then this reduces to

$$u_{ik} = \frac{1}{2\mu}\left[P\delta_{i1}\delta_{k1} - \left(K - \frac{2}{3}\mu\right)\frac{1}{3K}\delta_{ik}P\right],$$

or written in matrix form

$$u = \begin{pmatrix} \frac{1}{3}\left(\frac{1}{\mu} + \frac{1}{3K}\right)P & 0 & 0 \\ 0 & \left(\frac{1}{9K} - \frac{1}{\mu}\right)P & 0 \\ 0 & 0 & \left(\frac{1}{9K} - \frac{1}{\mu}\right)P \end{pmatrix}.$$

At this point it behoves us to introduce two new constants. The first is the **Young's modulus**,

$$E = \frac{9K\mu}{3K + \mu}.$$

It is a measure of how much a material will deform in the direction of the force. This can be seen from the fact that

$$u_{11} = \frac{P}{E}.$$

The other constant is **Poisson's ratio**,

$$\nu = \frac{1}{2}\frac{3K - 2\mu}{3K + \mu}.$$

It is a measure of deformation in the directions orthogonal to the forces which can be seen from the fact that

$$u_{22} = u_{33} = -\nu u_{11}.$$

The value of ν is restrained to $[-1, 1/2]$. Notice that the volume change is zero if $\nu = 1/2$, in this case we say that the material is incompressible. Since ν can be both positive and negative for some (most) materials a stretch in one direction will cause a contraction in the other directions but in some materials it will cause an expansion in all directions. ν is dimensionless and E has dimensions of force per unit area, or pascals, but the typical size of E is such that gigapascals are often the more appropriate unit.

Example 15.2. *Consider a block with its base (normal to the x_3 direction) fixed to the floor. Apply a force such that the deformation that results is $u_{33} = \gamma$ and all other components zero. What is the stress and free energy?*

The stress in general is given by

$$\sigma_{ik} = \frac{E}{1+\nu}\left[u_{ik} + \frac{\nu}{1-2\nu}u_{ll}\delta_{ik}\right].$$

In this case this means we have

$$\sigma_{33} = \frac{E(1-\nu)\gamma}{(1+\nu)(1-2\nu)} = -P$$

where we introduce the force per unit area, $-P$, which we take as negative as it acts out of the surface rather than into it. We also have

$$\sigma_{11} = \sigma_{22} = \frac{E\nu\gamma}{(1+\nu)(1-2\nu)} = -P\frac{\nu}{1-\nu}.$$

The free energy is

$$\mathcal{F} = \frac{E}{2(1+\nu)}\left[u_{ik}^2 + \frac{\nu}{1-2\nu}u_{ll}\right]$$

$$= \frac{P^2}{2E}\frac{(1+\nu)(1-2\nu)}{1-\nu}.$$

16   Point Force on an Infinite Medium

Consider an infinite elastic medium with a point force applied. Before we start with the maths what do we expect the result to be? We expect maximal deformation at the point where the force is applied and we expect a cylindrically symmetric deformation around this point which decreases as the distance from the point of application decreases. A point force like this is important as we consider linear deformations meaning that any force distribution can be thought of as a linear combination of, possibly an infinite number of, point forces.

Recall the force balance equation,

$$\frac{\partial \sigma_{ik}}{\partial x_k} + f_i = 0,$$

where σ is the stress tensor and $\boldsymbol{F}$ is the net external force. Writing σ in terms of the deformation tensor we have

$$-F_i = \frac{\partial}{\partial x_k}\left[\left(K - \frac{2}{3}\mu\right)\delta_{ik}u_{ll} + 2\mu u_{ik}\right]$$

$$= \left(K - \frac{2}{3}\mu\right)\partial_i u_{ll} + 2\mu\partial_k u_{ik}.$$

Now recall that, to second order,

$$u_{ik} = \frac{1}{2}(\partial_i u_k + \partial_k u_i)$$

where $\boldsymbol{u}$ is the displacement vector. Putting this into the force balance equation we have

$$-F_i = \left(K - \frac{2}{3}\mu\right)\partial_i\partial_l u_l + \mu\partial_k(\partial_i u_k + \partial_k u_i)$$

$$= \left(K - \frac{2}{3}\mu\right)\partial_i\partial_l u_l + \mu\partial_i\partial_k u_k + \partial_k\partial_k u_i$$

$$= \left(K - \frac{2}{3}\mu\right)\partial_i(\nabla \cdot \boldsymbol{u}) + \mu\partial_i(\nabla \cdot \boldsymbol{u}) + \nabla^2 u_i$$

$$= \left(K + \frac{1}{3}\mu \right) \partial_i (\nabla \cdot \boldsymbol{u}) + \nabla^2 u_i$$

or in vector form

$$\boldsymbol{0} = \mu \nabla^2 \boldsymbol{u} + \left(K + \frac{1}{3}\mu \right) \nabla(\nabla \cdot \boldsymbol{u}) + \boldsymbol{F}.$$

Since there is no other sensible choice we choose the origin to be at the point of application. A point force can be modelled with a Dirac delta distribution. So the force balance equation with a point force at the origin is

$$\boldsymbol{0} = \mu \nabla^2 \boldsymbol{u} + \left(K + \frac{1}{3}mu \right) \nabla(\nabla \cdot \boldsymbol{u}) + \boldsymbol{f}\delta(\boldsymbol{r}).$$

Where we choose f such that the force has the correct magnitude and direction.

To solve this problem we will move to Fourier space. Recall that the Fourier transform of a vector field, $\boldsymbol{v}$, is

$$\mathrm{FT}\{\boldsymbol{v}\}(\boldsymbol{k}) = \hat{\boldsymbol{v}}(\boldsymbol{k}) = \int e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \boldsymbol{v}(\boldsymbol{r}) \, \mathrm{d}^3 k \,,$$

and the inverse transform is

$$\mathrm{FT}^{-1}\{\hat{\boldsymbol{v}}\}(\boldsymbol{r}) = \boldsymbol{v}(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \boldsymbol{u}(\boldsymbol{r}) \, \mathrm{d}^3 r \,.$$

In particular using the sifting property of the Dirac delta distribution we have

$$\delta(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k \,.$$

Fourier transforming the entire force balance equation then gives us

$$\frac{1}{(2\pi)^3} \int \mathrm{d}^3 k \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \left[-\mu k^2 \hat{u}_a + \left(K + \frac{1}{3}\mu \right) i k_a k_b \hat{u}_b + f_a \right] = 0.$$

This must hold for any $\boldsymbol{r}$ and so we have

$$\mu k^2 \hat{u}_a + \left(K + \frac{1}{3}\mu \right) k_a k_b \hat{u}_b = f_a. \tag{16.1}$$

We want to solve this for $\hat{u}$ as this will allow us to find the deformation, u, in real space. Multiply by k_a and we get

$$\mu k^2 k_a \hat{u}_a + \left(K + \frac{1}{3}\mu \right) k_a k_a k_b \hat{u}_b = f_a k_a.$$

Renaming $a \to b$ in the first term and summing over a in the second term we have

$$\mu k^2 k_b \hat{u}_b + \left(K + \frac{1}{3}\mu \right) k^2 k_b \hat{u}_b = f_a k_a.$$

Rearranging gives

$$k_b \hat{u}_b = \frac{f_b k_b}{(K + 4\mu/3)k^2}.$$

We substitute this back into equation 16.1 and rearranging we have

$$\hat{u}_a = \frac{1}{\mu k^2} \left[\delta_{ab} - \frac{K + \mu/3}{K + 4\mu/3} \frac{k_a k_b}{k^2} \right] f_b.$$

We then take the inverse Fourier transform to get back to real space:

$$u_a(\boldsymbol{r}) = \frac{1}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \frac{1}{\mu k^2} \left[\delta_{ab} - \frac{K + \mu/3}{K + 4\mu/3} \frac{k_a k_b}{k^2} \right] f_b \, \mathrm{d}^3 k \,.$$

This can be split into two integrals:

$$I_1 = \frac{1}{(2\pi)^3} \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \frac{1}{k^2} \, \mathrm{d}^3 k$$

$$I_2 = \frac{1}{(2\pi)^3} \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \frac{k_a k_b}{k^4}$$

For the first integral we move to a spherical coordinate system, $\{k, \vartheta, \varphi\}$, with the z-axis ($\vartheta = 0$) aligned with $\boldsymbol{r}$ so that $\boldsymbol{k}\cdot\boldsymbol{r} = kr\cos\vartheta$. Then

$$
\begin{aligned}
I_1 &= \frac{1}{(2\pi)^3} \int_0^{2\pi} \mathrm{d}\varphi \int_0^\infty \int_0^\pi \sin\vartheta e^{ikr\cos\vartheta} \,\mathrm{d}\vartheta \,\mathrm{d}k \\
&= \frac{1}{(2\pi)^3} 2\pi \int_0^\infty \int_0^\pi \sin\vartheta e^{ikr\cos\vartheta} \,\mathrm{d}\vartheta \,\mathrm{d}k \\
&= \frac{1}{(2\pi)^2} \int_0^\infty \int_0^\pi \sin(\vartheta) e^{ikr\cos\vartheta} \,\mathrm{d}\vartheta \,\mathrm{d}k \,.
\end{aligned}
$$

First we compute the inner integral over ϑ. To do this first notice that

$$
\begin{aligned}
\frac{\partial}{\partial\vartheta} e^{ikr\cos\vartheta} &= \frac{\partial}{\partial\vartheta}[\cos\vartheta] \frac{\partial}{\partial\cos\vartheta} e^{ikr\cos\vartheta} \\
&= -ikr\sin(\vartheta) e^{ikr\cos\vartheta}
\end{aligned}
$$

so

$$\sin(\vartheta) e^{ikr\cos\vartheta} = -\frac{1}{ikr} \frac{\partial}{\partial\vartheta} e^{ikr\cos\vartheta}$$

and hence

$$
\begin{aligned}
I_1 &= \frac{1}{(2\pi)^2} \int_0^\infty \int_0^\pi \sin(\vartheta) e^{ikr\cos\vartheta} \,\mathrm{d}\vartheta \,\mathrm{d}k \\
&= \frac{1}{(2\pi)^2} \int_0^\infty \int_0^\pi -\frac{1}{ikr} \frac{\partial}{\partial\vartheta} e^{ikr\cos\vartheta} \,\mathrm{d}\vartheta \,\mathrm{d}k \\
&= \frac{1}{(2\pi)^2} \int_0^\infty -\frac{1}{ikr} \left[e^{ikr\cos\vartheta} \right]_0^\pi \,\mathrm{d}k \\
&= \frac{1}{(2\pi)^2} \int_0^\infty -\frac{1}{ikr} \left[e^{-ikr} - e^{ikr} \right] \,\mathrm{d}k \\
&= \frac{1}{(2\pi)^2} \int_0^\infty \frac{1}{kr} 2\sin(kr) \,\mathrm{d}k \\
&= \frac{1}{2\pi^2} \int_0^\infty \frac{\sin(kr)}{kr} \,\mathrm{d}k \,.
\end{aligned}
$$

Now let $\xi = kr$ so $\mathrm{d}k = \mathrm{d}\xi/r$ and we have

$$I_1 = \frac{1}{2\pi r^2} \int_0^\infty \frac{\sin\xi}{\xi} \,\mathrm{d}\xi = \frac{1}{4\pi r}$$

where we have used the standard integral

$$\int_0^\infty \frac{\sin\xi}{\xi} \,\mathrm{d}\xi = \frac{\pi}{2}\,.$$

For the second integral note that we can write

$$I_2 = \frac{1}{(2\pi)^3} \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \frac{k_a k_b}{k^4} \,\mathrm{d}^3 k = -\frac{1}{(2\pi)^3} \frac{\partial^2}{\partial x_a \partial x_b} \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \frac{1}{k^4} \,\mathrm{d}^3 k \,.$$

Then in the same spherical system as we used for the first integral we have

$$I_2 = -\frac{\partial^2}{\partial x_a \partial x_b} \frac{1}{2\pi^2} \int_0^\infty \frac{\sin(kr)}{k^3 r} \,\mathrm{d}k \,.$$

Taking the derivative and defining $\xi = kr$ we have

$$I_2 = \frac{1}{2\pi^2 r} \int_0^\infty \left[\delta_{ab} \left(\frac{\sin\xi}{\xi^3} - \frac{\cos\xi}{\xi^2} \right) + \frac{x_a x_b}{r^2} \left(\frac{\sin\xi}{\xi} + 3\frac{\cos\xi}{\xi} - 3\frac{\sin\xi}{\xi} \right) \right] \,\mathrm{d}\xi \,.$$

These integrals can be done using the integrals

$$\int_0^\xi \frac{1 - \cos \xi}{\xi^2} \, \mathrm{d}\xi = \frac{\pi}{2}, \qquad \text{and} \qquad \int_0^\infty \frac{\xi - \sin \xi}{\xi^3} = \frac{\pi}{4}.$$

Using this we have

$$I_2 = \frac{1}{8\pi r} \left(\delta_{ab} + \frac{x_a x_b}{r^2} \right).$$

Combining everything we have

$$\begin{aligned} u_a(\boldsymbol{r}) &= \frac{1}{8\pi r} \frac{K + \mu/3}{\mu(K + 4\mu/3)} \left[\left(1 + \frac{2\mu}{K + \mu/3} \right) \delta_{ab} + \frac{x_a x_b}{r^2} \right] f_b \\ &= \frac{1}{8\pi r E} \frac{1 + \nu}{1 - \nu} \left[(3 - 4\nu)\delta_{ab} + \frac{x_a x_b}{r^2} \right] f_b. \end{aligned}$$

The main feature of this result is that the deformation decays as $1/r$ meaning that it is still felt far from the point of application but it does decrease with distance.

Part III

Fluid Mechanics

17   Introduction To Fluid Mechanics

We can model a fluid as a collection of 'parcels' of fluid which are somewhere between molecules and macroscopic volumes. We assume that within each parcel the intrinsic properties, such as pressure and temperature, are constant and that parcels are small enough that neighbouring parcels have very similar intrinsic properties and therefore we can approximate these properties as changing continuously throughout the fluid.

17.1   Material and Spatial Variables

There are two ways of describing a fluid that are in common use. The first is called the **material** or **Lagrangian** description where we track each parcel as the fluid flows. We label each parcel, typically with its initial position, $\boldsymbol{a}$. We can then give, for example, the pressure in that parcel at time t as $p(\boldsymbol{a}, t)$.

The second description is the **spatial** or **Eulerian** description where we consider a fixed point in space, $\boldsymbol{x}$, and then the pressure at this point at time t is $p(\boldsymbol{x}, t)$. Notice that p is a different function here than in the Lagrangian description.

The relationship between the Lagrangian and Eulerian descriptions is given by a map, χ, which takes a parcel that starts at $\boldsymbol{a}$ and gives its position, $\boldsymbol{x}$, at time t by $\boldsymbol{x} = \boldsymbol{\chi}(\boldsymbol{a}, t)$. This is shown in figure 17.1 which depicts a control volume in the reference domain $\Omega(0)$ mapping to a volume in the current configuration, $\Omega(t)$. In the language of the earlier part of this course the two descriptions are simply different coordinate systems. The fixed spatial coordinate system is relative to some fixed object. A point described in this spatial coordinate will always be at the same place but the fluid packet that is at this point will change. The material description has a coordinate system that deforms with the fluid. A point in this coordinate system will always describe the same fluid parcel but its location in space will change.

17.2   Fluid Velocity

The velocity of fluid is an important quantity for obvious reasons. The problem is that it isn't straight forward to define. Clearly different parts of the fluid can move at different speeds so what is it we measure speed relative to? The obvious answer is we measure speed relative to a fixed point in the spatial description, however since we are interested in the velocity of a fluid parcel this means that we
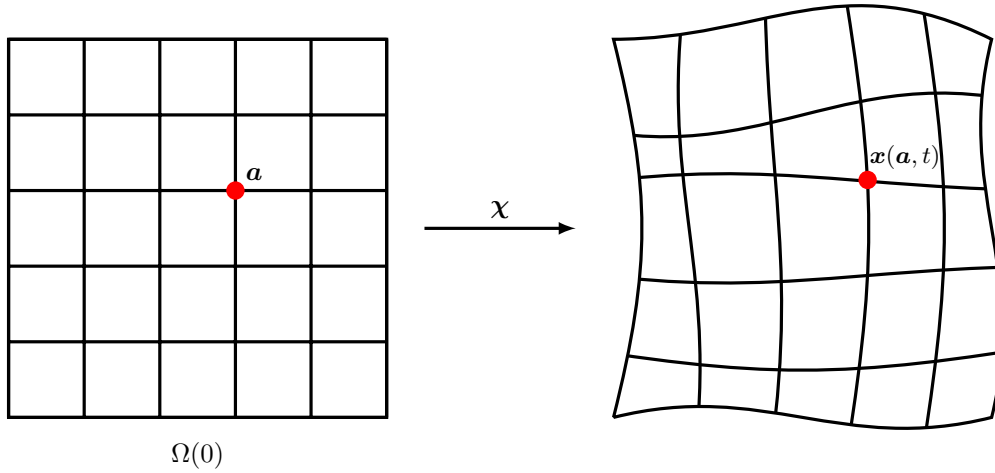
Figure 17.1: Fluid flow maps a parcel initially at $\boldsymbol{x}(\boldsymbol{a}, 0) = \boldsymbol{a}$ to $\boldsymbol{\chi}(\boldsymbol{a}, t) = \boldsymbol{x}(\boldsymbol{a}, t)$.

have to do a transformation of variables. That is if we define the fluid velocity at a point, $\boldsymbol{x}$, in the spatial description as $\boldsymbol{u}(\boldsymbol{x}, t) = \partial_t \boldsymbol{x}(t)$ then the velocity of a material point labelled $\boldsymbol{a}$ is

$$\left. \frac{\partial \boldsymbol{x}}{\partial t} \right|_{\boldsymbol{a}} = \frac{\partial}{\partial t} \boldsymbol{\chi}(\boldsymbol{a}, t) = \boldsymbol{u}(\boldsymbol{x}(\boldsymbol{a}, t), t)$$

where $\boldsymbol{x}(\boldsymbol{a}, 0) = \boldsymbol{a}$. Notice the double time dependence as both the position of $\boldsymbol{a}$ and the map from the spatial to the material description depend on time.

17.3   Material Derivative

Suppose we have a function, f, which depends on the Lagrangian variable, $\boldsymbol{a}$, and time, then we can describe the rate of time of f as $\partial_t f(\boldsymbol{a}, t)$. However if f is expressed in terms of the Eulerian variable, $\boldsymbol{x} = \boldsymbol{\chi}(\boldsymbol{a}, t)$ then we have to account for this time dependence as well and we have

$$\left. \frac{\partial}{\partial t} f(\boldsymbol{x}(\boldsymbol{a}, t), t) \right|_{\boldsymbol{a}} = \left. \frac{\partial f}{\partial t} \right|_{\boldsymbol{x}} + \left. \frac{\partial f}{\partial x_i} \right|_t \left. \frac{\partial x_i}{\partial t} \right|_{\boldsymbol{a}} = \left(\frac{\partial}{\partial t} + \boldsymbol{u} \cdot \nabla \right) f = \frac{\mathrm{D} f}{\mathrm{D} t}$$

This defines the **material derivative**

$$\frac{\mathrm{D}}{\mathrm{D} t} = \frac{\partial}{\partial t} + \boldsymbol{u} \cdot \nabla.$$

We can think of this as a time derivative which follows the fluid as it flows. For example the acceleration of the fluid written in terms of the spatial representation of the velocity, $\boldsymbol{u}(\boldsymbol{x}, t)$, is given by

$$\frac{\mathrm{D} \boldsymbol{u}}{\mathrm{D} t} = \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{u}.$$

One important measure of the fluid deformation is the Jacobian of the map $\boldsymbol{x} = \boldsymbol{\chi}(\boldsymbol{a}, t)$ since the Jacobian gives us the scale factor for volumes. The Jacobian matrix is[10]

$$\mathcal{J} = \left(\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{a}} \right)^{\mathsf{T}},$$

or in index notation

$$\mathcal{J}_{ij} = \frac{\partial x_i}{\partial a_j}.$$

Taking $t = 0$ to be before the material is deformed we have $\mathcal{J}(\boldsymbol{x}, 0) = I$ and for a continuous transformation the Jacobian must also be continuously connected to the identity. The Jacobian determinant,

---

[10]sometimes the Jacobian is defined without the transpose, this has no effect on the important features of the Jacobian since $\det \mathcal{J} = \det \mathcal{J}^{\mathsf{T}}$.

$J = \det \mathcal{J}$, gives the volume scale factor for a volume element, $\mathrm{d}V$. One quantity of interest is how fast this changes,

$$\frac{\mathrm{d}J}{\mathrm{d}t} = (\nabla \cdot \boldsymbol{u})J.$$

If $J(\boldsymbol{x}, t) = 1$ for all $\boldsymbol{x}$ and t then we say that the fluid flow is **volum preserving** or that the fluid is **incompressible**. It turns out that incompressible fluids are much simpler and still a relatively good description of many every day fluids in normal circumstances so it is common to assume a fluid is incompressible. In this case we have that $\mathrm{d}J/\mathrm{d}t = 0$ and so for an incompressible fluid, since $J = 1$, we have $\nabla \cdot \boldsymbol{u} = 0$.

18   Conservation Laws

18.1   Conservation of Mass

The fluid density, $\rho(\boldsymbol{x}, t)$ is defined as the mass per unit volume for a volume element, $\mathrm{d}V$, centred at $\boldsymbol{x}$, so the mass of this infinitesimal fluid parcel is $\rho(\boldsymbol{x}, t)\,\mathrm{d}V$. Consider a volume of fluid, $\Omega(t)$, as the material volume deforms and flows the mass of the fluid is determined by integrating over this volume:

$$M(t) = \int_{\Omega(t)} \rho(\boldsymbol{x}, t)\,\mathrm{d}V\,.$$

If there is no mechanism for creation or destruction of mass then the mass must be conserved so that at all times, t, we have $M(t) = M(0)$, or equivalently

$$\frac{\mathrm{d}M}{\mathrm{d}t} = 0.$$

We can write this in Eulerian form as

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(t)} \rho(\boldsymbol{x}, t)\,\mathrm{d}V_x = 0$$

where $\mathrm{d}V_x = \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3$. We can write this integral equation in a differential form by considering it represented in material coordinates:

$$
\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(t)} \rho(\boldsymbol{x}, t)\,\mathrm{d}V_x \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(0)} \rho(\boldsymbol{x}(\boldsymbol{a}, t), t) J\,\mathrm{d}V_a \\
&= \int_{\Omega(0)} \left[\left(\frac{\mathrm{D}t}{\mathrm{D}\rho}(\boldsymbol{x}(\boldsymbol{a}, t), t) \right) J + \rho(\boldsymbol{x}(\boldsymbol{a}, t)) \frac{\mathrm{d}J}{\mathrm{d}t} \right] \mathrm{d}V_a \\
&= \int_{\Omega(0)} \left[\left(\frac{\mathrm{D}t}{\mathrm{D}\rho}(\boldsymbol{x}(\boldsymbol{a}, t), t) \right) J + \rho(\boldsymbol{x}(\boldsymbol{a}, t))(\nabla \cdot \boldsymbol{u}) J \right] \mathrm{d}V_a \\
&= \int_{\Omega(0)} \left[\frac{\mathrm{D}t}{\mathrm{D}\rho}(\boldsymbol{x}(\boldsymbol{a}, t), t) + \rho(\boldsymbol{x}(\boldsymbol{a}, t))(\nabla \cdot \boldsymbol{u}) \right] J\,\mathrm{d}V_a \\
&= \int_{\Omega(t)} \left[\frac{\mathrm{D}t}{\mathrm{D}\rho}(\boldsymbol{x}, t) + \rho(\boldsymbol{x}, t)(\nabla \cdot \boldsymbol{u}) \right] \mathrm{d}V_x\,.
\end{aligned}
$$

Since this must hold for all volumes, $\Omega(t)$, we must have

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho(\nabla \cdot \boldsymbol{u}) = 0.$$

If the fluid is incompressible then $\nabla \cdot \boldsymbol{u} = 0$ and we have

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = 0.$$

In other words for an incompressible fluid the density at any material point is constant.

18.2   Conservation of Momentum

The momentum of the fluid contained in the volume $\Omega(t)$ in terms of Eulerian variables is

$$\boldsymbol{p}(t) = \int_{\Omega(t)} \rho(\boldsymbol{x}, t)\boldsymbol{u}(\boldsymbol{x}, t)\, \mathrm{d}V_x\,.$$

We can express the rate of change of momentum as

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(t)} \rho(\boldsymbol{x}, t)\boldsymbol{u}(\boldsymbol{x}, t)\, \mathrm{d}V_x \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(0)} \rho(\boldsymbol{x}(\boldsymbol{a}, t), t)\boldsymbol{u}(\boldsymbol{x}(\boldsymbol{a}, t), t)J\, \mathrm{d}V_a \\
&= \int_{\Omega(0)} \left[\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}\boldsymbol{u}J + \rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}J + \frac{\mathrm{d}J}{\mathrm{d}t} \right] \mathrm{d}V_a \\
&= \int_{\Omega(0)} \left[\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}\boldsymbol{u} + \rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + (\nabla \cdot \boldsymbol{u}) \right] J\, \mathrm{d}V_a \\
&= \int_{\Omega(t)} \rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}J\, \mathrm{d}V_a \\
&= \int_{\Omega(t)} \rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}\, \mathrm{d}V_x
\end{aligned}
$$

Here we again assume an incompressible fluid so

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = 0, \qquad \text{and} \qquad \nabla \cdot \boldsymbol{u} = 0.$$

By Newton's second law the rate of change of momentum in the material volume $\Omega(t)$ is equal to the net external force acting on the volume.

The external forces come in two kinds. External forces such as gravity which we consider as force per unit volume, $\boldsymbol{f}$, and surface forces, such as viscous or elastic stress, which we denote as force per unit area, $\boldsymbol{t}$, for traction. The surface of the fluid volume at a point is described by the outward pointing surface normal, $\boldsymbol{n}$, and the surface traction is given by $\boldsymbol{t} = \boldsymbol{n} \cdot \sigma$ where σ is the stress tensor, which we will discuss more in the next section. Balancing forces we must have

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t) = \int_{\Omega(t)} \rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}\, \mathrm{d}V_x = \int_{\Omega(t)} \boldsymbol{f}\, \mathrm{d}V_x + \int_{\partial\Omega(t)} \boldsymbol{t}\, \mathrm{d}S_x$$

where $\partial\Omega(t)$ is the boundary of the volume $\Omega(t)$ and $\mathrm{d}S_x$ is an infinitesimal surface element. Using the divergence theorem the last integral can become a volume integral and we have

$$\int_{\Omega(t)} \rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}\, \mathrm{d}V_x = \int_{\Omega(t)} \boldsymbol{f}\, \mathrm{d}V_x + \int_{\Omega(t)} \nabla \cdot \sigma\, \mathrm{d}V_x\,.$$

Note that the stress tensor is symmetric so it doesn't matter if we take the divergence to be $\nabla \cdot \sigma = \partial_i\sigma_{ij}$ or $\nabla \cdot \sigma = \partial_j\sigma_{ij}$. This equation holds for all volumes and so we must have

$$\rho\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \boldsymbol{f} + \nabla \cdot \sigma. \tag{18.1}$$

This is the Cauchy momentum equation. We saw this equation in elasticity as well but since we only considered systems in equilibrium the acceleration was zero.

19   Navier–Stokes Equations

The most important object within continuum mechanics is undoubtedly the stress tensor, σ. When considering an elastic deformation of a solid we define the stress tensor to depend locally on the material deformation. In fluid mechanics we define the stress slightly differently. The defining difference between a fluid and a solid is that a fluid can deform an arbitrary amount. This means that the size of deformation

does not give us important information as a fluid can have an arbitrarily large deformation. Instead we define the stress to be locally dependent on the *rate* of material deformation. In the most general case the stress tensor, σ, in fluid mechanics depends on the rate of deformation in a complex, non-linear, and history dependent way defined by a system of partial differential equations.

We can consider the stress tensor to give the force per unit area on each face of a cubic control volume in each direction, in a similar way to the elastic case. The traction is then given by $\boldsymbol{t} = \hat{\boldsymbol{n}} \cdot \sigma$ where $\hat{\boldsymbol{n}}$ is the surface normal. So the component σ_{ij} represents the traction in the jth surface perpendicular to the ith direction. For example the traction on the surface with surface normal $\boldsymbol{e_1}$ is

$$\boldsymbol{t} = \boldsymbol{e_1} \cdot \sigma = \sigma_{11}\boldsymbol{e_1} + \sigma_{12}\boldsymbol{e_2} + \sigma_{13}\boldsymbol{e_3}.$$

The net force, $\boldsymbol{F}$, acting on an immersed body with boundary S is given by

$$\boldsymbol{F} = \int_S \hat{\boldsymbol{n}} \cdot \sigma \, \mathrm{d}S \,.$$

Similarly the torque, $\boldsymbol{L}$, about the point $\boldsymbol{x_0}$ on the same object is given by

$$\boldsymbol{L} = \int_S (\boldsymbol{x} - \boldsymbol{x}_0) \times (\hat{\boldsymbol{n}} \cdot \sigma) \, \mathrm{d}S \,.$$

We will only consider cases where there are no external torques which may lead to internal angular momentum. The result of this is that the stress tensor will always be symmetric, $\sigma = \sigma^{\mathsf{T}}$ or $\sigma_{ij} = \sigma_{ji}$. This can be shown by considering the torque on a volume element and imposing the conservation of angular momentum. Since the stress tensor is symmetric we only need to give 6 out of 9 components to completely specify the stress tensor.

So far we have made no assumptions about the type of fluid, simply about the existence (or lack of) external forces and torques. This means that our analysis so far holds for all fluids. We can write the stress tensor as

$$\sigma = -pI + \tau$$

where I is the identity matrix, p is the pressure and τ is the **deviatoric stress tensor**. We can define p to include the normal thermodynamic pressure but also the isotropic components of the stress such that τ is traceless. The deviatoric stress tensor contains information on the viscous stress and other stresses. For many (Newtonian) fluids of interest to us the deviatoric stress tensor can be written in terms of the local strain rate.

Consider a velocity field depending on the spatial coordinate $\boldsymbol{x}$, $\boldsymbol{u}(\boldsymbol{x}, t)$. This can be Taylor expanded about the point $\boldsymbol{x}$ as

$$\boldsymbol{u}(\boldsymbol{x} + \mathrm{d}\boldsymbol{x}, t) = \boldsymbol{u}(\boldsymbol{x}, t) + \mathrm{d}\boldsymbol{x} \cdot \nabla \boldsymbol{u}(\boldsymbol{x}, t) + \mathcal{O}\Big(|\mathrm{d}\boldsymbol{x}|^2\Big).$$

Here we have used

$$(\nabla \boldsymbol{u})_{ij} = \frac{\partial u_j}{\partial x_i}.$$

We can decompose $\nabla \boldsymbol{u}$ into symmetric and antisymmetric parts,

$$\dot{\gamma} = \nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathsf{T}}, \qquad \text{and} \qquad \omega = \nabla \boldsymbol{u} - (\nabla \boldsymbol{u})^{\mathsf{T}}$$

which gives us

$$\nabla \boldsymbol{u} = \frac{1}{2}(\dot{\gamma} + \omega).$$

We define the symmetric part, $\dot{\gamma}$, to be the **rate-of-strain tensor** and the antisymmetric part, ω, to be the **vorticity tensor**.

Consider a line element, $\mathrm{d}\boldsymbol{x}$, which extends from the point $\boldsymbol{x}$ and evolves in a linear flow field. We can then write

$$\frac{\mathrm{d}(\mathrm{d}\boldsymbol{x})}{\mathrm{d}t} = \boldsymbol{u}(\boldsymbol{x} + \mathrm{d}\boldsymbol{x}, t) - \boldsymbol{u}(\boldsymbol{x}, t) = \mathrm{d}\boldsymbol{x} \cdot \nabla \boldsymbol{u} = \mathrm{d}\boldsymbol{x} \cdot \frac{1}{2}(\dot{\gamma} + \omega).$$

We can consider the response of $\mathrm{d}\boldsymbol{x}$ due to the symmetric and antisymmetric parts individually. The response due to $\dot{\gamma}$ is best characterised by its eigenvalues and eigenvectors. Since $\dot{\gamma}$ is symmetric its

eigenvectors, d_i, are orthogonal. The basis formed from these eigenvectors is the principle axes of $\dot{\gamma}$. The associated eigenvalues, λ_i, are called the principle rates-of-strain. We can write the rate-of-strain tensor as

$$\dot{\gamma} = \sum_i 2\lambda_i \boldsymbol{d_i d_i}.$$

Here $\boldsymbol{d_i d_i}$ is the dyadic product of the vector $\boldsymbol{d_i}$ with itself. This is defined as

$$\boldsymbol{ab} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix}.$$

If $\nabla \boldsymbol{u}$ is entirely symmetric we can then write

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{d}\boldsymbol{x} = \frac{1}{2}\mathbf{d}\boldsymbol{x} \cdot \dot{\gamma} = \sum_i \lambda_i (\mathbf{d}\boldsymbol{x} \cdot \boldsymbol{d_i})\boldsymbol{d_i}.$$

This shows that a spherical control volume will deform into an ellipsoid with its axes aligned with the principle axes of $\dot{\gamma}$ and the lengths given by the principle strain rates.

It can also be shown that if $\nabla \boldsymbol{u}$ is entirely antisymmetric then

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}x} = \frac{1}{2}\mathbf{d}\boldsymbol{x} \cdot \omega = \frac{1}{2}(\nabla \times \boldsymbol{u}) \times \mathbf{d}\boldsymbol{x}.$$

A **Newtonian fluid** is defined to be a fluid where the deviatoric stress, τ, is linear in the rate of strain, $\dot{\gamma}$, and isotropic. The most general form of the deviatoric stress tensor for a Newtonian fluid is a linear combination of $\dot{\gamma}$ and $(\nabla \cdot \boldsymbol{u})I$. We can write this in such a way that we get two terms, one of which is traceless and the other is isotropic:

$$\tau = \mu \left(\dot{\gamma} - \frac{2}{3}(\nabla \cdot \boldsymbol{u})I \right) + \mu'(\nabla \cdot \boldsymbol{u})I.$$

We define μ as the **fluid viscosity** and μ' as the **dilational viscosity**. For an incompressible fluid we have $\nabla \cdot \boldsymbol{u} = 0$ and so

$$\tau = \mu \dot{\gamma}.$$

The stress is then given simply by

$$\sigma = -pI + \tau = -pI + \mu\dot{\gamma}.$$

The Cauchy momentum equation (equation 18.1) then becomes

$$\rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = \nabla \cdot \sigma + \boldsymbol{f} = -\nabla p + \mu\nabla^2\boldsymbol{u} + \boldsymbol{f}.$$

This, and the incompressibility condition, $\nabla \cdot \boldsymbol{u} = 0$, are called the **incompressible Navier–Stokes equations**. These are incredibly important equations as any incompressible fluid flow can be modelled as a solution to these equations simply by choosing boundary and initial conditions.

Each term of these equations tells us some important things about the flow:

$$\underbrace{\rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}}_{\text{Inertial Term}} = \rho\frac{\partial \boldsymbol{u}}{\partial t} + \underbrace{\rho\boldsymbol{u} \cdot \nabla \boldsymbol{u}}_{\text{Convection Term}} = -\nabla p + \underbrace{\mu\nabla^2\boldsymbol{u}}_{\text{Diffusive Term}} + \boldsymbol{f}.$$

The inertial term corresponds to flow in the fluid due to inertia of the fluid, the denser the fluid is the more important this term is. The diffusive term relates to energy dissipation in the fluid. The more viscous the fluid is the more important this term is. For a more viscous fluid the diffusion is slower as $\left|\nabla^2\boldsymbol{u}\right|$ must be smaller to for this term to remain constant. Notice that the diffusive term is proportional to $\nabla^2\boldsymbol{u}$, which should be no surprise from the diffusion equation.

20   Reynold's Numbers

Given a particular flow we can choose a characteristic speed, U, and length scale, L. For example for flow in a pipe the characteristic velocity may be the speed in the centre of the pipe and the length scale might be the radius of the pipe. For flow around a sphere the velocity may be the upstream velocity (the velocity of the fluid a long time before it reaches the sphere) and the length scale the radius of the sphere. Once we have chosen these scales we can construct a characteristic time scale, L/U. For pipe flow this corresponds to the time taken for fluid in the centre of the pipe to flow one pipe radius down the pipe. For flow around a sphere this corresponds to the time taken to flow one sphere radius if the sphere is not there. Using the viscosity, μ, we can also define a typical pressure scale $\mu U/L$, and a typical force scale $\mu U/L^2$.

Recall that the stress tensor for a Newtonian fluid is given by

$$\sigma = -pI + \mu\dot{\gamma} = -\pi I + \mu[\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathsf{T}}].$$

We can write this in a dimensionless way by defining the dimensionless quantities

$$\boldsymbol{x}^* = \frac{\boldsymbol{x}}{L}, \qquad \boldsymbol{u}^* = \frac{\boldsymbol{u}}{U}, \qquad t^* = \frac{Ut}{L},$$
$$p^* = \frac{Lp}{\mu U}, \qquad \text{and} \qquad \boldsymbol{f}^* = \frac{L^2\boldsymbol{f}}{\mu U}.$$

The incompressible Navier–Stokes equations written in terms of these dimensionless variables are then

$$\mathrm{Re}\frac{\mathrm{D}\boldsymbol{u}^*}{\mathrm{D}t^*} = \mathrm{Re}\left(\frac{\partial \boldsymbol{u}^*}{\partial t} + \boldsymbol{u}^* \cdot \boldsymbol{u}^*\right) = -\nabla^* p^* + \nabla^{*2}\boldsymbol{u}^* + \boldsymbol{f}^*$$

where $(\nabla^*)_i = \partial_{x_i^*}$ and

$$\mathrm{Re} = \frac{\rho UL}{\mu}$$

is a dimensionless quantity called the **Reynolds number**. This quantity characterises the flow in the system. We can think of it as a ratio,

$$\mathrm{Re} = \frac{\text{inertial effects}}{\text{viscous effects}}.$$

If Re is large then the inertial effects are important and we have nice laminar flow. If Re is small then viscous effects dominate and we have turbulent flow. What counts as a small or large value for Re depends on the situation and exactly which length scale and velocity scale we pick but in general Re > 2000 is considered large[11] and Re < 2000 is considered small. For Re ∼ 2000 we are in a transition state between laminar and turbulent flow.

Laminar flow is considerably simper than turbulent flow since for small Re we can neglect the left hand side and we get the following set of equations, called the **Stokes equations**

$$-\nabla p + \mu\nabla^2\boldsymbol{u} + \boldsymbol{f} = \boldsymbol{0},$$
$$\nabla \cdot \boldsymbol{u} = 0.$$

We will focus on the low Reynold's number regime as this is where analytic solutions can be found.

21   Flow Due to a Point Force

Consider an infinite volume of fluid which experiences a point force given by $\boldsymbol{f}\delta(\boldsymbol{r})$ where δ is the Dirac delta distribution. Suppose also that this force is small enough that we are in the low Reynold's number regime and therefore can use the Stokes equations which in component form are

$$\partial_i p + \mu\nabla^2 u_i + f_i\delta(\boldsymbol{r}) = 0, \qquad \text{and} \qquad \partial_i u_i = 0.$$

---

[11]this is the typically accepted engineering convention

We can expand the velocity field, $\boldsymbol{u}$, in a Fourier series:

$$\boldsymbol{u}(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int \hat{\boldsymbol{u}}(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k\,, \qquad \text{and} \qquad \hat{\boldsymbol{u}}(\boldsymbol{k}) = \int \boldsymbol{u}(\boldsymbol{r}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 r\,.$$

We can also do the same for the pressure:

$$p(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int \hat{p}(\boldsymbol{r}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k\,, \qquad \text{and} \qquad \hat{p}(\boldsymbol{k}) = \int p(\boldsymbol{r}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 r\,.$$

And again for the force term using the Fourier transform of the Dirac delta:

$$\boldsymbol{f}\delta(\boldsymbol{r}) = \boldsymbol{f}\frac{1}{(2\pi)^3} \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k\,.$$

Substituting these into the Stokes equation and using

$$\mathrm{FT}[\partial_a g] = ik_a \hat{g} \implies \mathrm{FT}[\partial_a^2 g] = -k_a^2 \hat{g}$$

we get

$$0 = \frac{1}{(2\pi)^3} \int [-ik_a \hat{p} - \mu \hat{u}_a + f_a] e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k$$

and

$$0 = \frac{1}{(2\pi)^3} \int k_a \hat{u}_a e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k\,.$$

These must hold for any particular position, $\boldsymbol{r}$, and so the integrands must be identically zero:

$$0 = -ik_a \hat{p} - \mu \hat{u}_a + f_a$$
$$0 = k_a \hat{u}_a.$$

We can eliminate the pressure by multiplying the first equation by k_a and summing over a to get

$$0 = -ik^2 \hat{p} - \mu k^2 \underbrace{\hat{u}_a k_a}_{=0} + f_a k_a \implies \hat{p} = \frac{f_a k_a}{ik^2}.$$

Substituting this in for $\hat{p}$ we have

$$0 = -ik_a \frac{k_b f_b}{ik^2} - \mu k^2 \hat{u}_a + f_a.$$

Rearranging this gives us

$$\hat{u}_a = \frac{1}{\mu k^2}[f_a - \frac{k_a k_b}{k^2} f_b] \frac{1}{\mu k^2} \left[\delta_{ab} - \frac{k_a k_b}{k^2} \right] f_b.$$

We now have the solution in k-space which is simply the Fourier transform of the solution in r-space:

$$u_a(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int \frac{1}{\mu k^2} \left[\delta_{ab} - \frac{k_a k_b}{k^2} \right] f_b e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k\,.$$

To compute this we need the following integrals:

$$I_1 = \frac{1}{(2\pi)^3} \int \frac{1}{k^2} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k\,, \qquad \text{and} \qquad I_2 = \frac{1}{(2\pi)^3} \int \frac{k_a k_b}{k^4} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^3 k\,.$$

We have already computed these in section 16 so we simply state the results here:

$$I_1 = \frac{1}{4\pi r}, \qquad \text{and} \qquad I_2 = \frac{1}{8\pi r} \left[\delta_{ab} + \frac{x_a x_b}{r^2} \right].$$

Hence the solution in real space is

$$u_a(\boldsymbol{r}) = \frac{1}{8\pi \mu r} \left[\delta_{ab} + \frac{x_a x_b}{r^2} \right] f_b.$$

Compare this to the solution for the similar elasticity problem in section 16:

$$u_a(\boldsymbol{r}) = \frac{1}{8\pi r E} \frac{1+\nu}{1-\nu} \left[(3 - 4\nu)\delta_{ab} + \frac{x_a x_b}{r^2} \right] f_b.$$

Notice that the form of these two solutions is very similar even though the first gives the velocity of fluid flow and the second gives the displacement of an elastic medium. These solutions go as $\sim 1/r$ far from the point of application. This decays fairly slowly so even at large distances there is often a non-negligible effect from disturbances. This is one of the things that makes fluid dynamics so complicated.

21.1   Stokes Force

Consider a sphere moving through an otherwise stationary fluid (sometimes referred to as a **quiescent** fluid). This sphere experiences a drag force which it can be shown is given by

$$\boldsymbol{F} = -6\pi\mu a \boldsymbol{U}$$

where $\boldsymbol{U}$ is the velocity of the sphere, a is the radius of the sphere, and μ is the viscosity of the fluid. This force is called the Stokes force. If instead the fluid is flowing at constant velocity $\boldsymbol{W}$ in some frame and in the same frame the sphere moves at velocity $\boldsymbol{U}$ then the formula still holds if we adjust for the relative velocity of the sphere and fluid:

$$\boldsymbol{F} = -6\pi\mu a (\boldsymbol{U} - \boldsymbol{W}).$$

By Newton's third law in these scenarios there must be a force, $-\boldsymbol{F}$, exerted on the fluid by the sphere.

Appendix

A   Column Matrices and Vectors

Often we write things of the form $A\boldsymbol{x} = \boldsymbol{y}$ where $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Strictly there is no defined multiplication between matrices and vectors. Instead we consider $\boldsymbol{x}$ and $\boldsymbol{y}$ as elements of $\mathcal{M}_{n \times 1}(\mathbb{R})$, that is as $n \times 1$ column matrices. We then perform the usual matrix multiplication.

The reason that this is possible is that $\mathcal{M}_{n \times 1}(\mathbb{R})$ is a real vector space of dimension $1n = n$ (in general $\mathcal{M}_{n \times m}(\mathbb{R})$ is a real vector space of dimension nm). In this vector space vector addition is given by matrix addition and scalar multiplication is just scalar multiplication of a scalar and matrix. It can be shown that any two real vector spaces of the same dimension are in fact isomorphic. What this means is that the two sets have all the same important structural features, the only thing that changes between them is what an element looks like. An isomorphism between the two sets is trivially given by

$$\varphi \colon \mathcal{M}_{n \times 1}(\mathbb{R}) \to \mathbb{R}^n$$
$$X = (x_{i1}) \mapsto \boldsymbol{x} = x_{i1}\boldsymbol{e_i} = x_i\boldsymbol{e_i}$$

so $\mathcal{M}_{n \times 1}(\mathbb{R}) \cong \mathbb{R}^n$. Because of this isomorphism we don't usually bother to distinguish between $\mathcal{M}_{n \times 1}(\mathbb{R})$ and $\mathbb{R}^n$.

B   Groups

> **Definition 36: Group**
>
> A group, $(G, \cdot)$, is a set, G, and a binary operation, $\cdot \colon G \times G \to G$ which has the following properties:
> - Associativity – For $g_1, g_2, g_3 \in G$ we have $g_1(g_2 g_3) = (g_1 g_2)g_3$.
> - Identity – There exists $e \in G$ such that $eg = ge = g$ for all $g \in G$.
> - Inverse – For each $g \in G$ there exists $g^{-1} \in G$ such that $gg^{-1} = g^{-1}g = e$.
> - Closure – For $g_1, g_2 \in G$ $g_1 g_2 \in G$.

A group is the natural structure to use when talking about symmetry as symmetries (that is transformations that leave some property invariant) naturally form a group structure. The reason for this is quite simple. If we have a set of symmetries then clearly doing nothing is a symmetry and therefore in this set, this corresponds to the identity. Any symmetry we apply can be undone to get back to the original state, this corresponds to the inverse. Any combination of symmetries is another symmetry, this corresponds to closure. Finally associativity allows us to combine symmetries and then apply them or apply them one at a time and either way we get the same result.

We have mentioned several groups in the main text. These are matrix groups, meaning the elements of the groups are matrices and the group operation is matrix multiplication. This is a natural way to talk about transformations. The groups we have mentioned are $\mathrm{GL}(n)$, $\mathrm{O}(n)$, and $\mathrm{SO}(n)$. These are each subgroups, meaning subsets that are also groups, of the previous group in the list. In the text we stated that $\mathrm{GL}(n)$ was 'the most general' matrix group in n dimensions. This is because if we relax the condition and allow matrices with zero determinant then we no longer necessarily have an inverse meaning that $\mathcal{M}_{n \times n}(\mathbb{R})$ is not necessarily a group. On the other hand $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are both subgroups of $\mathrm{GL}(n)$ and therefore are more restricted in some way.

C   Lie Algebras

A matrix Lie group is a group $G \subseteq \mathrm{GL}(n, \mathbb{C})$ which is closed in the sense that if $A_n \in G$ is a sequence of matrices and tends to some matrix A then either A in G or $A \notin \mathrm{GL}(n, \mathbb{C})$. That is a Lie group is a subgroup of invertible matrices such that any sequence either tends to a matrix in G or to a matrix that is non-invertible. Lie groups are continuous groups.

If $G \subseteq \mathrm{GL}(n, \mathbb{C})$ is a Lie group then its Lie algebra, $\mathfrak{g}$, is the set of all matrices $X \in \mathcal{M}_{n \times n}(\mathbb{C})$ such that $e^{tX} \in G$ for all $t \in \mathbb{R}$.

If $\mathfrak{g}$ is a Lie algebra then it is a vector space and it is closed under the commutator. That is if $X, Y \in \mathfrak{g}$ then $[X, Y] \in \mathfrak{g}$.

What we said above is all true for matrix Lie groups but there are Lie groups that aren't made of matrices. For these we need a more general definition. A Lie algebra is a vector space $\mathfrak{g}$ over some field, $\mathbb{F}$, along with a binary operation, $[\cdot, \cdot] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the Lie bracket, which satisfies the following axioms:

- Bilinearity – That is $[\cdot, \cdot]$ is linear in its first and second argument so

$$[ax + by, z] = a[x, y] + b[y, z], \qquad \text{and} \qquad [z, ax + by] = a[z, x] + b[z, y]$$

  for all $x, y, z \in \mathfrak{g}$ and $a, b \in \mathbb{F}$.

- Alternativity – For all $x \in \mathfrak{g}$

$$[x, x] = 0.$$

- The Jacobi identity – For all $x, y, z \in \mathfrak{g}$

$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0.$$

These combined imply a fourth property, anti-commutativity, where for all $x, y \in \mathfrak{g}$

$$[x, y] = -[y, x].$$

Alternatively if the fields characteristic is not 2 then anti-commutativity implies alternativity. Examples of Lie brackets includes the commutator and the cross product.

A Lie algebra as defined this way will then generate a Lie group, G, which is all the elements e^{tx} for $x \in \mathfrak{g}$ and $t \in \mathbb{R}$.

Lie groups have lots of nice properties like continuity and smoothness which make them very useful in physics. A lot can be learned from a lie group by studying its Lie algebra so Lie algebras are also very common. For example $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{GL}(n, \mathbb{C})$, $\mathrm{O}(n)$, and $\mathrm{SO}(n)$ are all Lie groups with corresponding Lie algebras $\mathfrak{gl}(n, \mathbb{R})$, $\mathfrak{gl}(n, \mathbb{C})$, $\mathfrak{o}(n)$, and $\mathfrak{so}(n)$.

D   Tensors, a Mathematicians Definition

We defined tensors as objects that transform in a certain way under a rotation. This definition is often given as

> A tensor is an object that transforms like a tensor.

Clearly this is not a helpful description unless you already know how a tensor transforms. There is however, a more precise mathematical definition which can be shown to be equivalent. This definition, at least as much of it as we need for the way tensors are used in this course,[12] is as follows.

Definition 37: Tensor

A rank n tensor, T, on a vector space, V, with associated field of scalars $\mathbb{F}$ is a multi-linear function

$$T \colon V^n \to \mathbb{F}.$$

By multi-linear function what we mean is

$$T(v_1, v_2, \ldots, v_i + cw, \ldots, v_n) = T(v_1, v_2, \ldots, v_i, \ldots, v_n) + cT(v_1, v_2, \ldots, w, \ldots, v_n)$$

for any $i \in \{1, 2, \ldots, n\}$. Note that $v_i, w \in V$ are vectors but we will use a notation in this section where vectors get no special treatment when it comes to typesetting.

---

[12]The full definition of a tensor allows for it to act on the dual space as well and then we distinguish between co- and contravariant indices

So how does this gel with the transformation definition? First we have to define what we mean by the components of a tensor. To do this we need a basis, $\{e_i\}$. We will denote by v_p^i the ith component of the vector v_p. By repeated application of linearity we have

$$T(v_1, \ldots, v_n) = v_1^{i_1} \cdots v_n^{i_n} T(e_{i_1}, \ldots, e_{i_n}) = v_1^{i_1} \cdots v_n^{i_n} T_{i_1 \cdots i_n}.$$

So when we talk of components of a tensor what we actually mean is the tensor evaluated on the basis vectors:

$$T_{i_1 \cdots i_n} = T(e_{i_1}, \ldots, e_{i_n}).$$

Under a rotation, $L \in \mathrm{SO}(d)$ the basis vectors transform as $e_i' = l_{ij} e_j$ to give a new basis, $\mathcal{B}' = \{e_j'\}$. So we have

$$
\begin{aligned}
T_{i_1 \cdots i_n}' &= T(e_{i_1}', \ldots, e_{i_n}') \\
&= T(l_{i_1 \alpha_1} e_{\alpha_1}, \ldots, l_{i_n \alpha_n} e_{\alpha_n}) \\
&= l_{i_1 \alpha_1} \cdots l_{i_n \alpha_n} T(e_{\alpha_1}, \ldots, e_{\alpha_n}) \\
&= l_{i_1 \alpha_1} \cdots l_{i_n \alpha_n} T_{\alpha_1 \cdots \alpha_n}.
\end{aligned}
$$

We see that the components in this definition do indeed transform as tensors. This and the quotient theorem prove that the two definitions of tensors are indeed equivalent.