MA4J1 Continuum Mechanics Notes

Thomas Hudson

12 March, 2025

Table of contents

1 Preface

These notes are a guide to the main content of the fourth-year Maths module MA4J1 Continuum Mechanics taught at Warwick. They are intended to be essentially self–contained, but are based upon (my interpretation of) the content of chapters drawn from A First Course in Continuum Mechanics by Oscar Gonzalez and Andrew Stuart. This book is a good place to read more about the topics covered and to find various exercises to test your understanding of the content, but it is definitely not the only good book covering many of the topics we consider here. In particular, for those looking for additional reading, various additional books are recommended on the module Moodle page.

1.1 Aims and structure

The central aim of the module is to present a mathematical framework within which various continuum models of solids, liquids and gases can be described and derived, and to provide some examples of the Partial Differential Equation (PDE) models which result.

As background, we make use of various concepts from Linear Algebra, Analysis and Vector Calculus. In Chapter 2 and Chapter 3, we therefore recap various concepts and discuss some operations on these objects which may be new to you, and at the very least, may take a different perspective. Since some of these concepts will be taken as understood, it will be up to you to ensure that you are comfortable with any concepts which are not covered in detail in lectures. In particular, one of the techniques we use as a baseline throughout the module is converting back and forth between tensor notation and component notation in a particular coordinate system. You should ensure you get comfortable with doing this for yourself.

Chapter 4 introduces some of the important physical concepts central to this module, including mass density, force and stress fields. In particular, developing an understanding of *stress* is one of the keys to unlocking this module. Chapter 5 introduces the study of kinematics, or the study of motion of continuous bodies. Here, we introduce the concept of *strain*, which is another central focus for the module. In Chapter 6, we combine these concepts to derive balance laws, which are the Partial Differential Equations which govern a continuum body's motion. As we will see, these balance laws can be formulated

in *Eulerian* or *Lagrangian* form. The former formulation is generally used for the study of fluids, while the latter is generally used for the study of solid materials.

There will doubtless be errors and typos in these notes. If you spot anything, please feel free to let me know via email.

Dr Thomas Hudson, January 2025

2 Tensor Algebra

We begin by equipping ourselves with the mathematical tools required to derive mathematical models for the mechanical behaviour of objects in the world around us. To that end, this chapter will provide an overview of various concepts related to the manipulation of scalars and vectors, and will introduce the over-arching concept of *tensors*, which represent various physical quantities of interest which we will study later. The concept of a tensor generalises various objects you are already familiar with: scalars are *zeroth-order* tensors, and vectors are *first-order* tensors.

Aims: By the end of this chapter, you should be able to:

- Define the notion of a Cartesian coordinate frame;
- Manipulate and recall the definition of first-order, second-order and fourth-order tensors;
- Use index notation and the summation convention to express various operations on tensors;
- Define the Kronecker delta and Levi–Civita symbols and be able to quote various related identities;
- Give definitions for various special classes of second-order tensor;
- Recall the definition of the trace, determinant and exponential of second-order tensors;
- Quote and prove results on various second-order tensor decompositions; and

Basic notation. Here, and throughout these notes, 3-dimensional Euclidean space will be the model for the space within which objects are found. We denote this space $\mathbb{E}^3$, and real numbers are denoted $\mathbb{R}$. Elements of $\mathbb{R}$ are referred to as scalars, and are denoted by italic letters, e.g. $t \in \mathbb{R}$. Elements of $\mathbb{E}^3$ are referred to as points, and are denoted by bold-face italic letters, e.g. $\boldsymbol{x} \in \mathbb{E}^3$.

2.1 Motivation

Before launching into a discussion of tensors, we briefly motivate their introduction, and what physical concepts we wish to explain by using these objects.

Vectors are an important tool which we use to represent various physical concepts, such as displacement, velocity, acceleration and forces. We often represent vectors with respect to some fixed basis, so that we can describe them in coordinates. When considering physical modelling, we often want to think about how different observers may represent the same concept, such as the velocity of the particle, and this leads to the notion of a coordinate transformation. Coordinates are subjective representations of something which we believe is objective, like the velocity of a body, and so we will recap some details of how to do this here.

In addition to considering vectors and how they transform as individual objects, we will also be concerned with how collections of vectors change. For example, consider a small cube within a material: The edges of this cube can be represented by 3 basis vectors. At later time, if the material deforms, these vectors will change, and, assuming the transformation between these two collections is approximately linear, we can collect the 3 new vectors together into an object called the strain tensor. The strain tensor is a second-order tensor: this is just a slightly different way of looking at linear transformations on vectors (or their matrices), which will already be familiar.

At a yet higher level of abstraction, just as second-order tensors represent transformations of vectors, we can consider how second-order tensors are mapped to other second-order tensors, which leads us to the concept of a fourth-order tensor. An important fourth-order tensor is the elasticity tensor which transforms the strain second-order tensor into the stress second-order tensor, and we will discuss this in detail later in the module.

2.2 Vectors

In this module, a *vector* is a quantity with both magnitude and direction in three-dimensional space, denoted by bold lower-case letters, $\boldsymbol{v}$. Note immediately that we are using the same notation for both points and vectors, which could be cause for confusion, but we view these as distinct objects (and see later for further discussion).

Vectors will represent important physical concepts such as forces, momentum, velocity, and acceleration. It is expected that many of these concepts (applied to point particles, rather than continuum bodies) will be familiar. Similarly, you should also already be familiar with various operations on three-dimensional vectors from Linear Algebra, including the dot and cross products.

Throughout this module, any reference to the magnitude (or norm) of a vector will always mean the Euclidean norm, which will be denoted by vertical bars $|\boldsymbol{v}|$. Recall that

the magnitude of a vector is the length of the vector, and is therefore always a positive quantity, i.e. $|\boldsymbol{v}| \geq 0$. The zero vector, $\mathbf{0}$, is the unique vector with zero magnitude, and has no particular direction attached to it. Any vector $\boldsymbol{v}$ with magnitude (or length) 1, i.e. $|\boldsymbol{v}| = 1$, will be called a *unit vector*.

2.2.1 Vector algebra

A standard geometric way to think about (and draw) vectors is to use arrows, with the length of the arrow indicating a vector's magnitude, and the direction of the arrow matching the direction of the vector. Recall that a displacement vector connects spatial points, vector addition joins vectors end-to-end to generate a new vector, and scalar multiplication alters the length and possibly inverts the direction of vectors. Vectors will be viewed as being equal if they have the same magnitude and direction, regardless of their position in space.

We denote the set of all vectors $\mathcal{V}$, and note that with the notions of the addition and scalar multiplication referred to above, $\mathcal{V}$ has the structure of a real vector space, since

$$\boldsymbol{u} + \boldsymbol{v} \in \mathcal{V} \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}, \quad \text{and} \quad \alpha\boldsymbol{v} \in \mathcal{V} \quad \text{for all } \alpha \in \mathbb{R},\ \boldsymbol{v} \in \mathcal{V}.$$

2.2.2 The scalar and vector products

The *scalar product* or *dot product* of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined geometrically as

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\boldsymbol{v}| \cos\theta,$$

where $\theta \in [0, \pi]$ is the angle between the tips of $\boldsymbol{u}$ and $\boldsymbol{v}$, if taken to emanate from the same point. Two vectors are *orthogonal* or *perpendicular* if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. Recall that $\boldsymbol{u} \cdot \boldsymbol{u} = |\boldsymbol{u}|^2$ (this follows directly from the definition above).

The *vector product* or *cross product* of two 3-dimensional vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined as

$$\boldsymbol{u} \times \boldsymbol{v} = \big(|\boldsymbol{u}||\boldsymbol{v}| \sin\theta\big)\boldsymbol{e},$$

where $\theta \in [0, \pi]$ is again the angle between the tips of $\boldsymbol{u}$ and $\boldsymbol{v}$, and $\boldsymbol{e}$ is the unit vector perpendicular to the plane containing $\boldsymbol{u}$ and $\boldsymbol{v}$, oriented in a right-hand fashion. If $\boldsymbol{u} \times \boldsymbol{v} = \mathbf{0}$, then $\boldsymbol{u}$ and $\boldsymbol{v}$ are *parallel*. Recall that the magnitude $|\boldsymbol{u} \times \boldsymbol{v}|$ is the area of the parallelogram with two sides given by $\boldsymbol{u}$ and $\boldsymbol{v}$ emanating from the same point.

2.2.3 Projections, bases and coordinate frames

If $\boldsymbol{e}$ is a unit vector, then any vector $\boldsymbol{v}$ can be decomposed into a component which is parallel to $\boldsymbol{e}$, $\boldsymbol{v}_{\boldsymbol{e}}$, and a component which is perpendicular to $\boldsymbol{e}$, $\boldsymbol{v}_{\boldsymbol{e}}^{\perp}$:

$$\boldsymbol{v} = \boldsymbol{v}_{\boldsymbol{e}} + \boldsymbol{v}_{\boldsymbol{e}}^{\perp}, \tag{2.1}$$

where $\boldsymbol{v}_{\boldsymbol{e}} = (\boldsymbol{v} \cdot \boldsymbol{e})\boldsymbol{e}$ and $\boldsymbol{v}_{\boldsymbol{e}}^{\perp} = \boldsymbol{v} - \boldsymbol{v}_{\boldsymbol{e}}$.

Exercise: Verify the properties of this decomposition. Is it unique?

A *right-handed orthonormal basis* for $\mathcal{V}$ means three perpendicular unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ such that

$$\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3, \quad \boldsymbol{e}_2 \times \boldsymbol{e}_3 = \boldsymbol{e}_1, \quad \text{and} \quad \boldsymbol{e}_3 \times \boldsymbol{e}_1 = \boldsymbol{e}_2.$$

Note that the notion of *right-handedness* requires that $(\boldsymbol{e}_1 \times \boldsymbol{e}_2) \cdot \boldsymbol{e}_3 = 1$.

Exercise: Show that three mutually perpendicular unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ always satisfy $|(\boldsymbol{e}_1 \times \boldsymbol{e}_2) \cdot \boldsymbol{e}_3| = 1$.

By repeatedly applying the orthogonal decomposition Equation 2.1, any vector $\boldsymbol{v}$ can be uniquely written in *components* with respect to the basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ as

$$\boldsymbol{v} = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3 \quad \text{with} \quad v_i = \boldsymbol{v} \cdot \boldsymbol{e}_i \in \mathbb{R}.$$

For brevity here we have written i to mean a generic subscript which ranges from 1 to 3.

Recall that we often arrange the components of a vector into a 3×1 *column matrix*, here denoted

$$[\boldsymbol{v}] = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

Recall that the transpose of this matrix is a 1×3 matrix

$$[\boldsymbol{v}]^T = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}.$$

We call $[\boldsymbol{v}]$ the matrix *representation* of $\boldsymbol{v}$ with respect to the given basis.

2.2.4 Interpreting vectors physically

Above, we have been careful to distinguish between vectors $\boldsymbol{v} \in \mathcal{V}$, which represent arrows in space, and their representations $[\boldsymbol{v}] \in \mathbb{R}^3$, which are triplets of numbers. This is because our representation of $\boldsymbol{v}$ depends upon the choice of basis. Physically, this corresponds to viewing the same vectors from different positions and orientations in space, and corresponds to the idea that if two observers view the same vector from different perspectives, they will describe the vector in different ways. Later in this section

we will consider the same vector relative to different bases, and this distinction will be important.

As you are likely already aware, we can also represent spatial points $\boldsymbol{x} \in \mathbb{E}^3$ through their displacement from some fixed origin. To encode this idea, we define the following notion: A Cartesian *coordinate frame* for $\mathbb{E}^3$ means a reference point $\boldsymbol{o} \in \mathbb{E}^3$ called the *origin* together with a right-handed orthonormal basis $\{\boldsymbol{e}_i\}$ for the associated vector space $\mathcal{V}$. Relative to this coordinate frame, any point $\boldsymbol{x} \in \mathbb{E}^3$ has *coordinates* x_i where

$$x_i = (\boldsymbol{x} - \boldsymbol{o}) \cdot \boldsymbol{e}_i.$$

As such, we can write the displacement vector $\boldsymbol{x} - \boldsymbol{o}$ uniquely as

$$\boldsymbol{x} - \boldsymbol{o} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3.$$

Later, we will see that it is an important physical principle that numerous quantities of interest are independent of the position of the observer, and therefore the origin of our coordinate system is not very important. From now on, we therefore avoid making explicit reference to the origin, and identify points $\boldsymbol{x} \in \mathbb{E}^3$ with their position vectors $\boldsymbol{x} - \boldsymbol{o} \in \mathcal{V}$. Note that this explains the ambiguity in our notation up to this point!

2.3 Index notation

In this section, we fix a coordinate frame, and consider vectors in terms of their components relative to this frame.

2.3.1 The summation convention

Many operations acting on vectors (and other tensors) expressed in coordinates involve sums. For example, given a coordinate frame with basis $\{\boldsymbol{e}_i\}$ and vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$ with components u_i and v_i, the scalar product is

$$\boldsymbol{u} \cdot \boldsymbol{v} = \left(\sum_{i=1}^{3} u_i\boldsymbol{e}_i\right) \cdot \left(\sum_{j=1}^{3} v_i\boldsymbol{e}_i\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} u_iv_j\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \sum_{i=1}^{3} u_iv_i,$$

where we have used the properties of an orthonormal basis to conclude that

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Since sums over indices occur so frequently, it is usual to adopt the convention that *any repeated index letter in a term is summed over*. For example, we write

$$u_i\boldsymbol{e}_i \quad \text{in place of} \quad \sum_{i=1}^{3} u_i\boldsymbol{e}_i \quad \text{and} \quad u_iv_i \quad \text{in place of} \quad \sum_{i=1}^{3} u_iv_i.$$

Throughout the module, the summation convention will always be assumed to be in force unless otherwise stated. Further examples of the convention in action include:

$$a_3 b_k v_k = a_3(b_1 v_1 + b_2 v_2 + b_3 v_3)$$
$$u_i v_i + v_i v_i = (u_1 v_1 + u_2 v_2 + u_3 v_3) + (v_1^2 + v_2^2 + v_3^2)$$
$$x_i x_j y_j y_i = x_i y_i x_j y_j = (x_1 y_1 + x_2 y_2 + x_3 y_3)(x_1 y_1 + x_2 y_2 + x_3 y_3).$$

A repeated index is called a *dummy index*. This reflects the fact that the actual letter used is irrelevant to the result (just as the symbol used in the argument of a integral is irrelevant to the result):

$$u_i v_i = u_1 v_1 + u_2 v_2 + u_3 v_3 = u_z v_z.$$

Note that the convention only applies to pairs of indices: no summation is implied in the expressions a_i, $a_i b_i b_i$.

Any index which is not prescribed a numerical value but is not a dummy index is called a *free index*. For example, in the equation

$$w_i = u_i v_j v_j, \tag{2.2}$$

the index i is a free index, and j is a dummy index. Since free indices can take on the value 1, 2 or 3, we can use free indices to abbreviate groups of similar equations. For example Equation 2.2 is shorthand for the three equations

$$w_1 = u_1(v_1 v_1 + v_2 v_2 + v_3 v_3),$$
$$w_2 = u_2(v_1 v_1 + v_2 v_2 + v_3 v_3),$$
$$w_3 = u_3(v_1 v_1 + v_2 v_2 + v_3 v_3),$$

which saves a lot of writing. Similarly, equations with two free indices condense 9 different equations. Note that the free indices on both sides of an equality should always be consistent. The following examples are all ambiguous!

$$a_i = b_j \quad u_i v_j = w_i y_j y_j \quad e_i e_j = a_i a_k b_k b_j + c_p d_l d_l c_q.$$

2.3.2 The Kronecker delta and Levi–Civita symbols

A frame $\{\boldsymbol{e}_i\}$ has an associated *Kronecker delta* symbol δ_{ij}, defined to be

$$\delta_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j, \end{cases}$$

and a *Levi–Civita symbol* or *permutation symbol* ϵ_{ijk}, defined to be

$$\epsilon_{ijk} = (\boldsymbol{e}_i \times \boldsymbol{e}_j) \cdot \boldsymbol{e}_k = \begin{cases} +1 & ijk = 123, 231, \text{ or } 312, \\ -1 & ijk = 321, 213, \text{ or } 132, \\ 0 & \text{otherwise (i.e. an index is repeated).} \end{cases}$$

Notice that the numerical values of these symbols are the same, independently of the particular right-handed orthonormal frame we choose. The Kronecker delta is invariant under swapping of the indices, while the Levi–Civita symbol changes sign if two indices are transposed, but it invariant under cyclic permutation of the indices. In other words, we have the following identities:

$$\delta_{ij} = \delta_{ji}, \quad \text{and} \quad \epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji}.$$

Note that this is our first significant example of free indices in action!

2.3.3 Useful identities

Since the Kronecker delta and Levi–Civita symbol are defined in terms of the frame vectors $\{\boldsymbol{e}_i\}$, various identities follow. These include:

$$\boldsymbol{e}_i = \delta_{ij}\boldsymbol{e}_j, \quad \boldsymbol{e}_i \times \boldsymbol{e}_j = \epsilon_{ijk}\boldsymbol{e}_k, \quad \text{and} \quad \boldsymbol{e}_i = \tfrac{1}{2}\epsilon_{ijk}\boldsymbol{e}_j \times \boldsymbol{e}_k. \tag{2.3}$$

Exercise: Verify these expressions from the definitions.

Moreover, we can use the Kronecker delta and Levi–Civita symbol and the frame identities Equation 2.3 to express various vector operations. In particular, if $\boldsymbol{a} = a_i\boldsymbol{e}_i$, $\boldsymbol{b} = b_j\boldsymbol{e}_j$ and $\boldsymbol{c} = c_k\boldsymbol{e}_k$, then

$$\begin{aligned}
\boldsymbol{a}\cdot\boldsymbol{b} &= a_i b_j(\boldsymbol{e}_i\cdot\boldsymbol{e}_j)\\
&= a_i b_j \delta_{ij}\\
&= a_i b_i,\\
\boldsymbol{a}\times\boldsymbol{b} &= a_i b_j(\boldsymbol{e}_i\times\boldsymbol{e}_j)\\
&= a_i b_j \epsilon_{ijk}\boldsymbol{e}_k,\\
(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c} &= (a_i b_j \epsilon_{ijk}\boldsymbol{e}_k)\cdot(c_m\boldsymbol{e}_m)\\
&= \epsilon_{ijk}a_i b_j c_m \delta_{km}\\
&= \epsilon_{ijk}a_i b_j c_k,\\
\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) &= (a_q\boldsymbol{e}_q)\times(b_i c_j \epsilon_{ijk}\boldsymbol{e}_k)\\
&= \epsilon_{ijk}a_q b_i c_j \boldsymbol{e}_q\times\boldsymbol{e}_k\\
&= \epsilon_{qkp}\epsilon_{ijk}a_q b_i c_j \boldsymbol{e}_p\\
&= \epsilon_{pqk}\epsilon_{ijk}a_q b_i c_j \boldsymbol{e}_p.
\end{aligned}$$

We note that the permutation invariance of the scalar triple product follows from the permutation invariance of ϵ_{ijk}, and the latter identity will be used in a moment to verify an important connection between the Levi–Civita symbol and the Kronecker delta.

Note that our geometrical understanding of the scalar product and scalar triple product is expressed in terms of lengths and angles, and so does not depend upon the coordinate system used to evaluate these quantities. This means they are *frame-independent*, and so do not depend upon the coordinate system in which we observe them.

In addition, we have the following identities connecting the Levi–Civita symbol and the Kronecker delta. These identities are often useful for reducing complex expression written in index notation.

Proposition 2.1 (Epsilon-delta identities)**.** *Let ϵ_{ijk} be the Levi–Civita symbol and δ_{ij} the Kronecker delta. Then*

$$\epsilon_{abq}\epsilon_{cdq} = \delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc} \quad \textit{and} \quad \epsilon_{apq}\epsilon_{bpq} = 2\delta_{ab}.$$

Exercise: Prove this result.

As an application of this result, we can use the expression for the vector triple product above to show that

$$\begin{aligned}\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) &= \epsilon_{pqk}\epsilon_{ijk} a_q b_i c_j \boldsymbol{e}_p \\ &= (\delta_{pi}\delta_{qj} - \delta_{pj}\delta_{qi}) a_q b_i c_j \boldsymbol{e}_p \\ &= (a_q c_q) b_p \boldsymbol{e}_p - (a_q b_q) c_p \boldsymbol{e}_p \\ &= (\boldsymbol{a}\cdot\boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{c}.\end{aligned}$$

2.4 Second-order tensors

Many physical quantities, including forces, velocities and accelerations are well-represented by vectors, or first-order tensors, which have 3 components in a given Cartesian coordinate frame. However, other important quantities in our study, including the concepts of *stress* and *strain* are represented not by vectors, but by linear transformations between vectors. The need to describe these quantities leads to the notion of a second-order tensor with nine components in a given coordinate frame.

2.4.1 Definition and algebra

A *second-order tensor* $\boldsymbol{T}$ on the vector space $\mathcal{V}$ is a linear mapping $\boldsymbol{T} : \mathcal{V} \to \mathcal{V}$, i.e.

$$\boldsymbol{T}(\alpha\boldsymbol{u} + \beta\boldsymbol{v}) = \alpha\boldsymbol{T}(\boldsymbol{u}) + \beta\boldsymbol{T}(\boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{V} \text{ and } \alpha, \beta \in \mathbb{R}.$$

The set of all second-order tensors is denoted $\mathcal{V}^2$. Just as with vectors, we define the *zero tensor* $\boldsymbol{O}$ which satisfies $\boldsymbol{O}(\boldsymbol{v}) = \boldsymbol{0}$ for all $\boldsymbol{v} \in \mathcal{V}$, and the identity tensor $\boldsymbol{I}$, which satisfies $\boldsymbol{I}(\boldsymbol{v}) = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$. Second-order tensors $\boldsymbol{S}$ and $\boldsymbol{T}$ are equal if and only if $\boldsymbol{S}(\boldsymbol{v}) = \boldsymbol{T}(\boldsymbol{v})$ for all $\boldsymbol{v} \in \mathcal{V}$. Since second-order tensors act linearly, it is common to follow a mathematical convention and write $\boldsymbol{T}\boldsymbol{v}$ in place of $\boldsymbol{T}(\boldsymbol{v})$, and we will follow this rule from now on.

Examples:

1. Projecting a vector onto the subspace generated by a given unit vector $\boldsymbol{e}$ can be expressed in terms of a second-order *projection* tensor $\boldsymbol{P}$ with

$$\boldsymbol{P}\boldsymbol{v} = (\boldsymbol{v} \cdot \boldsymbol{e})\boldsymbol{e}. \tag{2.4}$$

2. Reflecting a vector in the plane with unit normal $\boldsymbol{e}$ can be expressed in terms of a second-order tensor $\boldsymbol{T}$ with

$$\boldsymbol{T}\boldsymbol{v} = \boldsymbol{v} - 2(\boldsymbol{v} \cdot \boldsymbol{e})\boldsymbol{e}. \tag{2.5}$$

Using the definition linear algebra, it is easy to check that the set of second-order tensors is indeed a vector space, defining $(\boldsymbol{S} + \boldsymbol{T})\boldsymbol{v} = \boldsymbol{S}\boldsymbol{v} + \boldsymbol{T}\boldsymbol{v}$ and $(\alpha\boldsymbol{T})\boldsymbol{v} = \alpha(\boldsymbol{T}\boldsymbol{v})$. Moreover, it has additional algebraic structure as we can multiply tensors through composition, and we define $\boldsymbol{S}\boldsymbol{T} \in \mathcal{V}^2$ via $(\boldsymbol{S}\boldsymbol{T})\boldsymbol{v} = \boldsymbol{S}(\boldsymbol{T}\boldsymbol{v})$ for any $\boldsymbol{S}, \boldsymbol{T} \in \mathcal{V}^2$ and arbitrary $\boldsymbol{v}$.

2.4.2 Coordinate representation

In the same way that vectors have components in a coordinate frame, so too do second-order tensors. Given a second-order tensors $\boldsymbol{S}$, the *components* S_{ij} of $\boldsymbol{S}$ in a Cartesian coordinate frame $\{\boldsymbol{e}_i\}$ are the nine numbers defined to be

$$S_{ij} = \boldsymbol{e}_i \cdot (\boldsymbol{S}\boldsymbol{e}_j).$$

We claim that these components completely describe $\boldsymbol{S} : \mathcal{V} \to \mathcal{V}$. Given vectors $\boldsymbol{v} = v_i\boldsymbol{e}_i$ and $\boldsymbol{u} = u_j\boldsymbol{e}_j$ with $\boldsymbol{v} = \boldsymbol{S}\boldsymbol{u}$, we have

$$v_i = \boldsymbol{e}_i \cdot (\boldsymbol{S}\boldsymbol{u}) = (\boldsymbol{e}_i \cdot \boldsymbol{S}\boldsymbol{e}_j)u_j = S_{ij}u_j.$$

Thus the components of $\boldsymbol{S}$ are the coefficients in the linear relationship between the components of $\boldsymbol{v}$ and $\boldsymbol{u}$, as expressed in the particular coordinate frame we consider. As you know, it is common to represent these component as a 3×3 matrix,

$$[\boldsymbol{S}] = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

As usual, the transpose of this matrix is denoted $[\boldsymbol{S}]^T$, which has components $[\boldsymbol{S}]^T_{ij} = [\boldsymbol{S}]_{ji} = S_{ji}$.

Examples: Choosing $\boldsymbol{e} = \boldsymbol{e}_2$ in the projection and reflection tensor examples given in Equation 2.4 and Equation 2.5, we have that

$$[\boldsymbol{P}] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad [\boldsymbol{T}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2.4.3 The tensor product of vectors

The *tensor product* (sometimes called the *dyadic product*) of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is the second-order tensor denoted $\boldsymbol{a} \otimes \boldsymbol{b}$ defined by

$$(\boldsymbol{a} \otimes \boldsymbol{b})\boldsymbol{v} = (\boldsymbol{b} \cdot \boldsymbol{v})\boldsymbol{a} \quad \text{for all } \boldsymbol{v} \in \mathcal{V}.$$

In terms of components with respect to a particular Cartesian coordinate frame, it can be checked that

$$[\boldsymbol{a} \otimes \boldsymbol{b}]_{ij} = a_i b_j, \quad \text{i.e.} \quad [\boldsymbol{a} \otimes \boldsymbol{b}] = \begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{pmatrix}.$$

In a given a coordinate frame $\{\boldsymbol{e}_j\}$, the elementary tensor products $\{\boldsymbol{e}_i \otimes \boldsymbol{e}_j\}_{i,j=1}^3$ form a basis for the space of tensors $\mathcal{V}^2$. In particular, for any $\boldsymbol{A} \in \mathcal{V}^2$, we have

$$\boldsymbol{A} = A_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j.$$

It turns out that the basis defined in this way is actually an orthonormal basis in the standard inner product on second-order tensors; we discuss this fact further below.

2.4.4 Change of basis

As we have discussed above, in a given coordinate frame, we may represent a vector $\boldsymbol{v} \in \mathcal{V}$ as a triplet of components $[\boldsymbol{v}] \in \mathbb{R}^3$, and a second-order tensor $\boldsymbol{S} \in \mathcal{V}^2$ as a matrix of components $[\boldsymbol{S}] \in \mathbb{R}^{3\times 3}$. In this section, we discuss how the representations $[\boldsymbol{v}]$ and $[\boldsymbol{S}]$ change with a change of coordinate frame.

To that end, suppose that $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{e}'_i\}$ are two coordinate frame for $\mathbb{E}^3$. By a *change of basis* tensor $\boldsymbol{A}$ mapping from $\{\boldsymbol{e}_i\}$ to $\{\boldsymbol{e}'_i\}$, we mean

$$\boldsymbol{A} = \boldsymbol{e}'_j \otimes \boldsymbol{e}_j = A_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad \text{where} \quad A_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}'_j.$$

We could equally define a change of basis tensor $\boldsymbol{B}$ from $\{\boldsymbol{e}'_i\}$ to $\{\boldsymbol{e}_i\}$, defined to be $\boldsymbol{B} = \boldsymbol{e}_j \otimes \boldsymbol{e}'_j = B_{ij}\boldsymbol{e}'_i \otimes \boldsymbol{e}'_j$ with $B_{ij} = \boldsymbol{e}'_i \cdot \boldsymbol{e}_j$, and we note that since the frames we choose are arbitrary, any properties we deduce for $\boldsymbol{A}$ will also hold for $\boldsymbol{B}$.

The idea behind a change of basis tensor is that it allow us to express the basis vectors of a frame in terms of the basis vectors of another. For example, by orthonormality of the basis $\{\boldsymbol{e}'_i\}$, we have

$$\boldsymbol{e}'_j = (\boldsymbol{e}_i \cdot \boldsymbol{e}'_j)\boldsymbol{e}_i = A_{ij}\boldsymbol{e}_i. \tag{2.6}$$

Similarly, working with respect to the second basis, we have that

$$\boldsymbol{e}_i = (\boldsymbol{e}_i \cdot \boldsymbol{e}'_k)\boldsymbol{e}'_k = A_{ik}\boldsymbol{e}'_k. \tag{2.7}$$

These two identities imply the following result:

Proposition 2.2. *The components A_{ij} of a change of basis tensor $\boldsymbol{A}$ have the properties that*

$$A_{ij}A_{ik} = \delta_{jk} \quad \textit{and} \quad A_{ik}A_{jk} = \delta_{ij},$$

or in matrix notation,

$$[\boldsymbol{A}]^T[\boldsymbol{A}] = [\boldsymbol{I}] \quad \textit{and} \quad [\boldsymbol{A}][\boldsymbol{A}]^T = [\boldsymbol{I}].$$

It follows that $\boldsymbol{A}$ is an orthogonal tensor.

Proof. Substituting Equation 2.6 into the right-hand side of Equation 2.7 gives

$$\boldsymbol{e}_i = A_{ik}A_{jk}\boldsymbol{e}_j,$$

which implies that $A_{ik}A_{jk} = \delta_{jk}$. Substituting Equation 2.7 into Equation 2.6 gives the other result. □

Now consider an arbitrary vector $\boldsymbol{v} \in \mathcal{V}$. Let $[\boldsymbol{v}]_i = v_i$ be its component representation with respect to the basis $\{\boldsymbol{e}_i\}$, and $[\boldsymbol{v}]'_i = v'_i$ be its component representation with respect to the basis $\{\boldsymbol{e}'_i\}$. These components can be related using the basis change tensor $\boldsymbol{A}$:

$$\boldsymbol{v} = v_i\boldsymbol{e}_i = \boldsymbol{v}'_j\boldsymbol{e}'_j = v'_jA_{ij}\boldsymbol{e}_i, \quad \text{so that} \quad v_i = A_{ij}v'_j.$$

We can similarly deduce that $v'_k = A_{ik}v_i$, and these relations can be written in matrix notation as

$$[\boldsymbol{v}]' = [\boldsymbol{A}][\boldsymbol{v}] \quad \text{and} \quad [\boldsymbol{v}] = [\boldsymbol{A}]^T[\boldsymbol{v}]'.$$

For second-order tensors $\boldsymbol{S}$ with component representations $[\boldsymbol{S}]$ and $[\boldsymbol{S}]'$ with respect to the bases consider above, we have

$$[\boldsymbol{S}] = [\boldsymbol{A}]^T[\boldsymbol{S}]'[\boldsymbol{A}] \quad \text{and} \quad [\boldsymbol{S}]' = [\boldsymbol{A}][\boldsymbol{S}][\boldsymbol{A}]^T. \tag{2.8}$$

2.4.5 Traces and determinants

The *trace* is a function defined for second-order tensors, and is a linear mapping $\mathrm{tr} : \mathcal{V}^2 \to \mathbb{R}$ defined by

$$\mathrm{tr}\boldsymbol{S} = \mathrm{tr}[\boldsymbol{S}] = [\boldsymbol{S}]_{ii},$$

where, as previously, $[\boldsymbol{S}]$ denotes a component-matrix representation in an arbitrary coordinate frame. The trace of a tensor is therefore the same as the definition of the trace of a matrix given in Linear Algebra courses. The following result shows that the definition does not depend upon the particular coordinate frame chosen.

Proposition 2.3. *Suppose $\boldsymbol{S}$ is a second-order tensor with respective matrix representations $[\boldsymbol{S}]$ and $[\boldsymbol{S}]'$ in the frames $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{e}'_i\}$. Then*

$$\mathrm{tr}[\boldsymbol{S}] = \mathrm{tr}[\boldsymbol{S}]',$$

and hence the value of the trace is independent of the coordinate frame in which it is computed.

Proof. Suppose $\boldsymbol{A}$ is the change of basis tensor mapping $\{\boldsymbol{e}_i\}$ to $\{\boldsymbol{e}'_i\}$. Then by applying Equation 2.8 we have $[\boldsymbol{S}]_{ij} = [\boldsymbol{A}]_{ik}[\boldsymbol{S}]'_{kl}[\boldsymbol{A}]_{jl}$, and so, using Proposition 2.2, we have

$$\operatorname{tr}[\boldsymbol{S}] = [\boldsymbol{S}]_{ii} = [\boldsymbol{S}]'_{kl}[\boldsymbol{A}]_{ik}[\boldsymbol{A}]_{il} = [\boldsymbol{S}]'_{kl}\delta_{kl} = [\boldsymbol{S}]'_{kk} = \operatorname{tr}[\boldsymbol{S}]'.$$

□

The *determinant* function on second-order tensors is the mapping $\det : \mathcal{V}^2 \to \mathbb{R}$ defined by

$$\det \boldsymbol{S} = \det[\boldsymbol{S}] = \epsilon_{ijk}[\boldsymbol{S}]_{1i}[\boldsymbol{S}]_{2j}[\boldsymbol{S}]_{3k},$$

where $[\boldsymbol{S}]$ again denotes the component matrix representation of $\boldsymbol{S}$ in an arbitrary coordinate frame. As for the trace, the determinant of a tensor is the same as the determinant of its matrix representation. The following result shows the definition is independent of the coordinate frame.

Proposition 2.4. *Suppose $\boldsymbol{S}$ is a second-order tensor with respective matrix representations $[\boldsymbol{S}]$ and $[\boldsymbol{S}]'$ in the frames $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{e}'_i\}$. Then*

$$\det[\boldsymbol{S}] = \det[\boldsymbol{S}]',$$

and hence the value of the determinant is independent of the coordinate frame in which it is computed.

Exercise: Prove this result.

As you may have seen in a multivariable calculus module, the quantity $|\det \boldsymbol{S}|$ may be interpreted as the volume of the parallelipiped with edges defined by the vectors $\boldsymbol{S}\boldsymbol{e}_1$, $\boldsymbol{S}\boldsymbol{e}_2$ and $\boldsymbol{S}\boldsymbol{e}_3$, all emanating from the same point. The determinant also arises in the classification of certain types of tensors; for example, a second-order tensor $\boldsymbol{S}$ is invertible if and only if $\det \boldsymbol{S} \neq 0$. Moreover, since $\det(\boldsymbol{A}\boldsymbol{B}) = \det \boldsymbol{A} \det \boldsymbol{B}$, every orthogonal tensor $\boldsymbol{Q}$ has the property that $|\det \boldsymbol{Q}| = 1$, and rotation tensors are exactly those which satisfy $\det \boldsymbol{Q} = 1$.

2.4.6 Transpose and symmetric tensors

To any tensor $\boldsymbol{S} \in \mathcal{V}^2$, we associate a *transpose* $\boldsymbol{S}^T \in \mathcal{V}^2$, which is the unique tensors with the property that

$$(\boldsymbol{S}\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (\boldsymbol{S}^T\boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}.$$

We say that a tensor $\boldsymbol{S}$ is *symmetric* if $\boldsymbol{S} = \boldsymbol{S}^T$ and *skew-symmetric* if $\boldsymbol{S}^T = -\boldsymbol{S}$. Note that the components of the transpose are

$$[\boldsymbol{S}^T]_{ij} = S_{ji}.$$

2.4.7 Scalar product and norm on tensors

Now, just as for vectors, we can define a *scalar product* for second-order tensors, denoted

$$\boldsymbol{S} : \boldsymbol{D} = \mathrm{tr}(\boldsymbol{S}^T \boldsymbol{D}). \tag{2.9}$$

In addition, whenever we have a scalar product, we also have a norm:

$$|\boldsymbol{S}| := \sqrt{\boldsymbol{S} : \boldsymbol{S}} = \sqrt{\mathrm{tr}(\boldsymbol{S}^T \boldsymbol{S})}. \tag{2.10}$$

This norm is equivalent to the Frobenius norm which you may have seen defined for matrices.

Since the scalar product and norm are defined using the trace, it is immediate that these definitions are independent of the coordinate frame in which they are computed. In practice, we can use the following result to compute them using the tensor components.

Proposition 2.5 (Scalar product and norm of tensors)**.** *Let $\{\boldsymbol{e}_i\}$ be a Cartesian coordinate frame, and let $\boldsymbol{S}, \boldsymbol{D} \in \mathcal{V}^2$. Then*

$$\boldsymbol{S} : \boldsymbol{D} = S_{ij} D_{ij}, \quad \textit{and} \quad |\boldsymbol{S}| = \sqrt{S_{ij} S_{ij}}$$

Exercise: Prove this result.

We observe from the result of Proposition 2.5 that the Frobenius norm on tensors is similar to the Euclidean norm on vectors: both are the square root of the sum of the squares of the components.

2.4.8 Special classes of tensor

A tensor $\boldsymbol{S} \in \mathcal{V}^2$ is said to be *positive-definite* if it satisfies

$$\boldsymbol{v} \cdot (\boldsymbol{S}\boldsymbol{v}) > 0 \quad \text{for all } \boldsymbol{v} \neq \boldsymbol{0},$$

and is said to be *invertible* if there exists an *inverse* $\boldsymbol{S}^{-1} \in \mathcal{V}^2$ satisfying

$$\boldsymbol{S}\boldsymbol{S}^{-1} = \boldsymbol{S}^{-1}\boldsymbol{S} = \boldsymbol{I}.$$

The operations of inversion and transposition commute, i.e.

$$(\boldsymbol{S}^{-1})^T = (\boldsymbol{S}^T)^{-1},$$

so as shorthand, we denote the resulting tensor $\boldsymbol{S}^{-T}$.

Exercise: Verify that inversion and transposition commute.

A tensor $\boldsymbol{Q} \in \mathcal{V}^2$ is said to be *orthogonal* if its transpose is its inverse, i.e. $\boldsymbol{Q}^T = \boldsymbol{Q}^{-1}$, and hence

$$\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}.$$

An orthogonal tensor is called a *rotation* if, given any right-handed orthonormal frame $\{\boldsymbol{e}_i\}$, the vectors $\{\boldsymbol{Q}\boldsymbol{e}_i\}$ are also a right-handed orthonormal frame. We will see later that an orthogonal tensor is a rotation if and only if it has positive determinant.

In a given coordinate frame, the above ideas correspond to the standard notions for the corresponding component matrices. Note that these concepts do not depend upon the exact choice of the frame, since none of the definitions required any manipulation of components!

Later, it will be important to have various decompositions of tensors available. The first of these is given by the following result.

Proposition 2.6. *Every second-order tensor $\boldsymbol{S} \in \mathcal{V}^2$ can be uniquely written as*

$$\boldsymbol{S} = \boldsymbol{E} + \boldsymbol{W},$$

where $\boldsymbol{E} \in \mathcal{V}^2$ is a symmetric tensor, and $\boldsymbol{W} \in \mathcal{V}^2$ is a skew-symmetric tensor. More specifically, we have

$$\boldsymbol{E} = \tfrac{1}{2}(\boldsymbol{S} + \boldsymbol{S}^T) \quad \textit{and} \quad \boldsymbol{W} = \tfrac{1}{2}(\boldsymbol{S} - \boldsymbol{S}^T).$$

This result motivates the definition of two mappings, $\mathrm{sym} : \mathcal{V}^2 \to \mathcal{V}^2$ and $\mathrm{skew} : \mathcal{V}^2 \to \mathcal{V}^2$, defined as

$$\mathrm{sym}(\boldsymbol{S}) = \tfrac{1}{2}(\boldsymbol{S} + \boldsymbol{S}^T) \quad \text{and} \quad \mathrm{skew}(\boldsymbol{S}) = \tfrac{1}{2}(\boldsymbol{S} - \boldsymbol{S}^T).$$

It is straightforward to show that, in any coordinate frame, a symmetric tensor $\boldsymbol{E}$ has components which satisfy $E_{ij} = E_{ji}$ and a skew-symmetric tensor $\boldsymbol{W}$ has components which satisfy $W_{ij} = -W_{ji}$.

In 3 dimensions, skew-symmetric tensors have an important representation as a cross product, as shown by the following result.

Proposition 2.7. *For any skew-symmetric tensor $\boldsymbol{W} \in \mathcal{V}^2$, there is a unique vector $\boldsymbol{w} \in \mathcal{V}$ such that*

$$\boldsymbol{W}\boldsymbol{v} = \boldsymbol{w} \times \boldsymbol{v} \quad \textit{for all } \boldsymbol{v} \in \mathcal{V}.$$

We write $\boldsymbol{w} = \mathrm{vec}(\boldsymbol{W})$, and refer to $\boldsymbol{w}$ as the axial vector *associated with $\boldsymbol{W}$. In any frame, we have the component identity*

$$w_j = \tfrac{1}{2}\epsilon_{njm}W_{nm}.$$

Conversely, given any vector $\boldsymbol{w} \in \mathcal{V}$, there is a unique skew-symmetric tensor $\boldsymbol{W} \in \mathcal{V}^2$ such that

$$\boldsymbol{w} \times \boldsymbol{v} = \boldsymbol{W}\boldsymbol{v} \quad \textit{for all } \boldsymbol{v} \in \mathcal{V}.$$

We write $\boldsymbol{W} = \operatorname{ten}(\boldsymbol{w})$*, and call* $\boldsymbol{W}$ *the* axial tensor *associated with* $\boldsymbol{w}$*. In any frame, we have the component identity*

$$W_{ik} = \epsilon_{ijk} w_j.$$

By definition, we note that the functions $\operatorname{vec} : \mathcal{V}^2 \to \mathcal{V}$ *and* $\operatorname{ten} : \mathcal{V} \to \mathcal{V}^2$ *are inverses.*

Proof. We prove the two statements in order. Suppose $\boldsymbol{W} \in \mathcal{V}^2$ is given, and we seek $\boldsymbol{v}$ such that $\boldsymbol{W}\boldsymbol{v} = \boldsymbol{w} \times \boldsymbol{v}$. Taking components in an arbitrary frame, this equation reads

$$W_{ik} v_k = \epsilon_{ijk} w_j v_k.$$

Since $\boldsymbol{v}$ (and therefore v_k) is arbitrary, it must hold that

$$W_{ik} = \epsilon_{ijk} w_j. \tag{2.11}$$

Our aim now is to 'solve' for w_j. Since we don't have a simple way to invert the Levi–Civita symbol, the natural choice is to attempt to apply the Levi–Civita symbol and use the identities proved in Proposition 2.1:

$$\begin{aligned} \epsilon_{aik} W_{ik} &= \epsilon_{aik} \epsilon_{ijk} w_j \\ &= -\epsilon_{aik} \epsilon_{jik} w_j \\ &= -2\delta_{aj} w_j \\ &= -2w_a. \end{aligned}$$

Rearranging and renaming indices, we now obtain

$$w_j = \tfrac{1}{2} \epsilon_{njm} W_{nm},$$

as required. These equations prescribe all components of $\boldsymbol{w}$, and so it is clearly unique, since we have obtained it above directly as a mapping from components of W.

To establish the second result, suppose that $\boldsymbol{w}$ is given. We seek a symmetric tensor $\boldsymbol{W}$ such that $\boldsymbol{w} \times \boldsymbol{v} = \boldsymbol{W}\boldsymbol{v}$ for all $\boldsymbol{v}$. Taking components in an arbitrary frame, this requires that

$$\epsilon_{ijk} w_j v_k = W_{ik} v_k.$$

Using the same argument as above, we arrive at Equation 2.11, which is an explicit expression for W_{ik}, and so leads to a unique definition of the tensor $\boldsymbol{W}$, since once again, all components of $\boldsymbol{W}$ are given through the equations above.

To establish the final result, let $\boldsymbol{w} = \operatorname{vec}(\boldsymbol{W})$, take components and apply Proposition 2.1 to obtain

$$\begin{aligned} \big[\operatorname{ten}(\operatorname{vec}(\boldsymbol{W}))\big]_{ik} &= [\operatorname{ten}(\boldsymbol{w})]_{ik} \\ &= \epsilon_{ijk} w_j \\ &= \tfrac{1}{2} \epsilon_{ijk} \epsilon_{njm} W_{nm} \\ &= \tfrac{1}{2} \big(\delta_{in}\delta_{km} - \delta_{im}\delta_{kn}\big) W_{nm} \\ &= \tfrac{1}{2} (W_{ik} - W_{ki}) \\ &= W_{ik} = [\boldsymbol{W}]_{ik}. \end{aligned}$$

Since this identity holds for all components, it follows that $\mathrm{ten}(\mathrm{vec}(\boldsymbol{W})) = \boldsymbol{W}$, and a similar argument shows that $\mathrm{vec}(\mathrm{ten}(\boldsymbol{w})) = \boldsymbol{w}$. These two results show that the functions ten and vec are mutual inverses. □

2.4.9 Eigenvalues, Eigenvectors and Principal Invariants

An *eigenpair* for $\boldsymbol{S} \in \mathcal{V}^2$ means a scalar λ and a unit vector $\boldsymbol{e}$ satisfying

$$\boldsymbol{S}\boldsymbol{e} = \lambda \boldsymbol{e}.$$

Any such λ is called an *eigenvalue* and any such $\boldsymbol{e}$ is called an *eigenvector* of $\boldsymbol{S}$. We recall from Linear Algebra that λ is an eigenvalue of $\boldsymbol{S}$ if and only if it is a root of the characteristic polynomial

$$p(\lambda) = \det(\boldsymbol{S} - \lambda \boldsymbol{I}). \tag{2.12}$$

To any eigenvalue, we associate one or more linearly independent eigenvectors $\boldsymbol{e}$ by solving the equation

$$(\boldsymbol{S} - \lambda \boldsymbol{I})\boldsymbol{e} = \boldsymbol{0}. \tag{2.13}$$

Throughout this module, we will exclusively consider eigenpairs in the context of symmetric tensors, and in this case, we note the following important facts. When $\boldsymbol{S}$ is symmetric, it can be shown that all three roots of the characteristic polynomial $p(\lambda)$ defined in Equation 2.12 are real, and consequently, all solution $\boldsymbol{e}$ to Equation 2.13 also have real components. It therefore holds that any symmetric tensor has three eigenpairs consisting of real eigenvalues and real eigenvectors. We also have the following important result.

Proposition 2.8. *The eignvalues of any second-order symmetric, positive-definite tensor are strictly positive, and any two eigenvectors corresponding to distinct eigenvalues of a symmetric tensor are orthogonal.*

Proof. Let $\boldsymbol{S} \in \mathcal{V}^2$ be symmetric and positive definite, and let $(\lambda, \boldsymbol{e})$ be an eigenpair for $\boldsymbol{S}$. By definition, we have that $\boldsymbol{S}\boldsymbol{e} = \lambda \boldsymbol{e}$, and since $\boldsymbol{e}$ is a unit vector, we have

$$\lambda = \lambda \boldsymbol{e} \cdot \boldsymbol{e} = (\boldsymbol{S}\boldsymbol{e}) \cdot \boldsymbol{e} > 0,$$

so any eigenvalue must be positive, as stated. Suppose that $(\omega, \boldsymbol{e}')$ is a second eigenpair with $\lambda \neq \omega$. Then we have

$$\lambda \boldsymbol{e} \cdot \boldsymbol{e}' = (\boldsymbol{S}\boldsymbol{e}) \cdot \boldsymbol{e}' = \boldsymbol{e} \cdot (\boldsymbol{S}^T \boldsymbol{e}') = \boldsymbol{e} \cdot (\boldsymbol{S}\boldsymbol{e}') = \omega \boldsymbol{e} \cdot \boldsymbol{e}',$$

which entails that $(\lambda - \omega)\boldsymbol{e} \cdot \boldsymbol{e}' = 0$. Since $\lambda \neq \omega$, we must have that $\boldsymbol{e} \cdot \boldsymbol{e}' = 0$, completing the proof. □

We now state the following further result concerning symmetric second-order tensors, which is given without proof here.

Proposition 2.9. *Let $\boldsymbol{S} \in \mathcal{V}^2$ be a symmetric tensor. Then there exists a right-handed orthonormal basis $\{\boldsymbol{e}_i\}$ for $\mathcal{V}$ consisting of eigenvectors of $\boldsymbol{S}$. The corresponding set of eigenvalues $\{\lambda_i\}$ are the same (up to a permutation of the ordering) for any such basis, and form a complete set of eigenvalues of $\boldsymbol{S}$. Moreover, we may write*

$$\boldsymbol{S} = \sum_{i=1}^{3} \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i,$$

and with respect to the basis $\{\boldsymbol{e}_i\}$, $\boldsymbol{S}$ has component matrix representation

$$[\boldsymbol{S}] = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

We define the *principal invariants* of a second-order tensor $\boldsymbol{S}$ as the three scalars

$$\begin{aligned} I_1(\boldsymbol{S}) &= \mathrm{tr}\boldsymbol{S}, \\ I_2(\boldsymbol{S}) &= \tfrac{1}{2}[(\mathrm{tr}\boldsymbol{S})^2 - \mathrm{tr}(\boldsymbol{S}^2)], \\ I_3(\boldsymbol{S}) &= \det \boldsymbol{S}. \end{aligned}$$

We note that, due to the fact that both the trace and the determinant are independent of the coordinate frame in which we consider components of the tensor $\boldsymbol{S}$, these invariants are independent of the frame too (justifying their name). For any $\alpha \in \mathbb{R}$, these invariants satisfy the relation

$$p(\alpha) = \det(\boldsymbol{S} - \alpha \boldsymbol{I}) = -\alpha^3 + I_1(\boldsymbol{S})\alpha^2 - I_2(\boldsymbol{S})\alpha + I_3(\boldsymbol{S}),$$

which can be verified by working in an arbitrary coordinate frame, and if $\boldsymbol{S}$ is symmetric with eigenvalues λ_i, then

$$\begin{aligned} I_1(\boldsymbol{S}) &= \lambda_1 + \lambda_2 + \lambda_3, \\ I_2(\boldsymbol{S}) &= \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1, \\ I_3(\boldsymbol{S}) &= \lambda_1\lambda_2\lambda_3. \end{aligned}$$

For brevity, we will sometimes use the symbol $\mathcal{I}_{\boldsymbol{S}}$ to denote the triplet of invariants $(I_1(\boldsymbol{S}), I_2(\boldsymbol{S}), I_3(\boldsymbol{S}))$. The principal invariants can be used to state the following important result in Linear Algebra (which is stated without proof here).

Proposition 2.10. *Any second-order tensor $\boldsymbol{S} \in \mathcal{V}^2$ satisfies*

$$\boldsymbol{S}^3 - I_1(\boldsymbol{S})\boldsymbol{S}^2 + I_2(\boldsymbol{S})\boldsymbol{S} - I_3(\boldsymbol{S})\boldsymbol{I} = \boldsymbol{O},$$

or in other words, $p(\boldsymbol{S}) = \boldsymbol{O}$ if p is the characteristic polynomial of the tensor $\boldsymbol{S}$.

2.4.10 Further special decompositions

Later, we will require use of several important tensor decompositions, in addition to those proved in Proposition 2.6 and Proposition 2.9. These are stated here.

Proposition 2.11. *Suppose that $\boldsymbol{S} \in \mathcal{V}^2$ is symmetric and positive definite. Then, there exists a unique symmetric postive definite tensor $\boldsymbol{U}$ such that $\boldsymbol{U}^2 = \boldsymbol{S}$. In particular, if $(\lambda_i, \boldsymbol{e}_i)$ are the eigenpairs for $\boldsymbol{S}$, then we have*

$$\boldsymbol{U} = \sum_{i=1}^{3} \sqrt{\lambda_i} \boldsymbol{e}_i \otimes \boldsymbol{e}_i,$$

and it is usual to write $\boldsymbol{U} = \sqrt{\boldsymbol{S}}$.

The existence of square-root tensors allows us to state the following important representation theorem for second-order tensors.

Proposition 2.12. *Suppose that $\boldsymbol{F} \in \mathcal{V}^2$ satisfies $\det \boldsymbol{F} > 0$. Then there exist* right *and* left polar decompositions *of $\boldsymbol{F}$, which take the respective forms*

$$\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} \quad \textit{and} \quad \boldsymbol{F} = \boldsymbol{V}\boldsymbol{R},$$

where $\boldsymbol{U} = \sqrt{\boldsymbol{F}^T\boldsymbol{F}}$ and $\boldsymbol{V} = \sqrt{\boldsymbol{F}\boldsymbol{F}^T}$ are symmetric positive definite tensors, and $\boldsymbol{R}$ is a rotation tensor.

Proof. We note that $\boldsymbol{F}$ is invertible since it has positive determinant, and so $\boldsymbol{F}\boldsymbol{v} \neq \boldsymbol{0}$ for any $\boldsymbol{v} \neq \boldsymbol{0}$. Similarly, we have that $\boldsymbol{F}^T\boldsymbol{v} \neq \boldsymbol{0}$ for any $\boldsymbol{v} \neq \boldsymbol{0}$. From these observations we deduce that for any $\boldsymbol{v} \neq \boldsymbol{0}$,

$$\begin{aligned} \boldsymbol{v} \cdot (\boldsymbol{F}^T\boldsymbol{F}\boldsymbol{v}) &= (\boldsymbol{F}\boldsymbol{v}) \cdot (\boldsymbol{F}\boldsymbol{v}) > 0, \\ \boldsymbol{v} \cdot (\boldsymbol{F}\boldsymbol{F}^T\boldsymbol{v}) &= (\boldsymbol{F}^T\boldsymbol{v}) \cdot (\boldsymbol{F}^T\boldsymbol{v}) > 0. \end{aligned}$$

It follows that $\boldsymbol{F}^T\boldsymbol{F}$ and $\boldsymbol{F}\boldsymbol{F}^T$ are positive definite and are clearly symmetric, so we can apply Proposition 2.11 to ensure that $\boldsymbol{U}$ and $\boldsymbol{V}$ are well-defined.

Next, consider the tensor defined to be $\boldsymbol{R} = \boldsymbol{F}\boldsymbol{U}^{-1}$. As $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$, and so $\det \boldsymbol{F} = \det \boldsymbol{R} \det \boldsymbol{U}$, we have that

$$\det \boldsymbol{R} = \frac{\det \boldsymbol{F}}{\det \boldsymbol{U}}.$$

This allows us to deduce that $\det \boldsymbol{R} > 0$ as the determinants of both $\boldsymbol{F}$ and $\boldsymbol{U}$ are positive. Moreover,

$$\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{U}^{-T}(\boldsymbol{F}^T\boldsymbol{F})\boldsymbol{U}^{-1} = \boldsymbol{U}^{-1}\boldsymbol{U}^2\boldsymbol{U}^{-1} = \boldsymbol{I}.$$

It follows that $\boldsymbol{R}$ is a rotation as claimed, and we have established the right polar decomposition.

To establish the left polar decomposition, we let $\boldsymbol{Q} = \boldsymbol{V}^{-1}\boldsymbol{F}$, and apply similar arguments to those above to show $\boldsymbol{Q}$ is a rotation. To show that $\boldsymbol{Q} = \boldsymbol{R}$, consider $\boldsymbol{C} = \boldsymbol{Q}^T\boldsymbol{V}\boldsymbol{Q}$, which is symmetric and positive definite. Then $\boldsymbol{R}\boldsymbol{U} = \boldsymbol{Q}\boldsymbol{C}$, and $\boldsymbol{R} = \boldsymbol{Q}\boldsymbol{C}\boldsymbol{U}^{-1}$. Since $\boldsymbol{R}$ is a rotation, $\boldsymbol{R}^{-1} = \boldsymbol{R}^T$, so multiplying $\boldsymbol{R}\boldsymbol{U} = \boldsymbol{Q}\boldsymbol{C}$ on the left by $\boldsymbol{R}^T = \boldsymbol{U}^{-T}\boldsymbol{C}^T\boldsymbol{Q}^T = \boldsymbol{U}^{-1}\boldsymbol{C}\boldsymbol{Q}^T$, we have

$$\boldsymbol{U} = \boldsymbol{U}^{-1}\boldsymbol{C}\boldsymbol{Q}^T\boldsymbol{C} \quad \text{so} \quad \boldsymbol{C}^2 = \boldsymbol{U}^2 = \boldsymbol{F}^T\boldsymbol{F}.$$

From the uniqueness of the tensor square root, we have that $\boldsymbol{C} = \boldsymbol{U}$, and so $\boldsymbol{R} = \boldsymbol{Q}\boldsymbol{C}\boldsymbol{U}^{-1} = \boldsymbol{Q}$. It follows that $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} = \boldsymbol{V}\boldsymbol{R}$ as claimed. □

3 Tensor Calculus

Realistic bodies have properties which vary from place to place within them. If we seek to describe their properties using the language of tensors introduced in the previous chapter, then a way to describe their properties is through tensor-valued functions which assign properties to spatial points in the body. More specifically, this chapter will consider functions $\phi : \mathbb{E}^3 \to \mathbb{R}$, $\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ and $\boldsymbol{S} : \mathbb{E}^3 \to \mathcal{V}^2$ which describe these properties. We will also wish to consider functions which describe changes in the shape of a body, which take the form $\boldsymbol{\varphi} : \mathbb{E}^3 \to \mathbb{E}^3$, $g : \mathcal{V}^2 \to \mathbb{R}$ and $\boldsymbol{A} : \mathcal{V}^2 \to \mathcal{V}^2$. Many of these concepts will already be very familiar to you from earlier calculus modules.

Aims. By the end of this chapter, you should be able to:

- Give a definition of differentiability for scalar, vector, and tensor fields.
- Define the divergence for vector and tensor fields.
- Manipulate the gradient, divergence, curl, and Laplacian of fields in components using the summation convention, and so derive various product formula for these differential operators.
- Recall and apply the Divergence Theorem, the Localisation Theorem and the Mean Value Theorem for multidimensional integrals.
- Understand how to differentiate functions of second-order tensors.

3.1 Basic definitions

3.1.1 Points, tensors and representations

Throughout this chapter, we assume that a single fixed Cartesian coordinate frame $\{\boldsymbol{e}_i\}$ has be specified for Euclidean space $\mathbb{E}^3$, and we identify points $\boldsymbol{x} \in \mathbb{E}^3$, vectors $\boldsymbol{v} \in \mathcal{V}$ and second-order tensors $\boldsymbol{S} \in \mathcal{V}^2$ with their respective matrix representations, writing

$$\boldsymbol{x} = x_i \boldsymbol{e}_i = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3, \quad \boldsymbol{v} = v_i \boldsymbol{e}_i = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3$$

$$\text{and} \quad \boldsymbol{S} = S_{ij} \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

It follows that we have identified $\mathbb{E}^3$ and $\mathcal{V}$ both with $\mathbb{R}^3$, and $\mathcal{V}^2$ with $\mathbb{R}^{3\times3}$. Any function of a spatial point $\boldsymbol{x}$ is a function of the three real variables (x_1, x_2, x_3), and a function of a second-order tensor $\boldsymbol{S}$ is a function of nine real variables, $(S_{11}, S_{12}, \ldots, S_{33})$.

3.1.2 Norms and order symbols

The standard *norms* on the spaces $\mathbb{E}^3$, $\mathcal{V}$ and $\mathcal{V}^2$ are the scalar functions

$$\begin{aligned} |\boldsymbol{x}| &= \sqrt{\boldsymbol{x}\cdot\boldsymbol{x}} \quad \text{for any } \boldsymbol{x}\in\mathbb{E}^3, \\ |\boldsymbol{v}| &= \sqrt{\boldsymbol{v}\cdot\boldsymbol{v}} \quad \text{for any } \boldsymbol{v}\in\mathcal{V}, \\ \text{and} \qquad |\boldsymbol{S}| &= \sqrt{\boldsymbol{S}:\boldsymbol{S}} \quad \text{for any } \boldsymbol{S}\in\mathcal{V}^2. \end{aligned}$$

Just as the norms for vectors measure their size, $|\boldsymbol{S}|$ measures the size of the tensor $\boldsymbol{S}$. The notion of a norm allows us to define limits on the spaces $\mathbb{E}^3$, $\mathcal{V}$ and $\mathcal{V}^2$, along with notions of continuity.

Consider a function $\boldsymbol{f} : U \to W$, where U and W can be any of the spaces $\mathbb{E}^3$, $\mathcal{V}$, $\mathcal{V}^2$ or $\mathbb{R}$. If there are constants $C > 0$ and $r > 0$ such that $|\boldsymbol{f}(\boldsymbol{u})| \leq C|\boldsymbol{u}|^r$ as $\boldsymbol{u} \to \boldsymbol{0}$, then we write

$$\boldsymbol{f}(\boldsymbol{u}) = O(|\boldsymbol{u}|^r) \quad \text{as} \quad \boldsymbol{u}\to\boldsymbol{0}.$$

If there is a constant $r > 0$ such that

$$\frac{|\boldsymbol{f}(\boldsymbol{u})|}{|\boldsymbol{u}|^r} \to 0$$

as $\boldsymbol{u} \to \boldsymbol{0}$, then we write

$$\boldsymbol{f}(\boldsymbol{u}) = o(|\boldsymbol{u}|^r) \quad \text{as} \quad \boldsymbol{u}\to\boldsymbol{0}.$$

The symbols O and o are called the standard *order symbols*, and are used in many contexts.

The order symbols are a way to describe the behaviour of a function close to zero, or by translating, close to any other given value. We can read $O(|\boldsymbol{u}|^r)$ as a function which goes to zero 'at least as fast as' $|\boldsymbol{u}|^r$, while $o(|\boldsymbol{u}|^r)$ should be read as a function which goes to zero 'faster than' $|\boldsymbol{u}|^r$. This means that $O\boldsymbol{u}) = o(|\boldsymbol{u}|^r)$ implies that $\boldsymbol{f}(\boldsymbol{u}) = O(|\boldsymbol{u}|^r)$, but not the converse. When $p > r$, it is however true that $\boldsymbol{f}(\boldsymbol{u}) = O(|\boldsymbol{u}|^p)$ implies $\boldsymbol{f}(\boldsymbol{u}) = o(|\boldsymbol{u}|^r)$.

Given two functions $\boldsymbol{f}, \boldsymbol{g} : U \to W$, we will use the notation

$$\boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{g}(\boldsymbol{u}) + O(|\boldsymbol{u}|^r) \quad \text{as } \boldsymbol{u}\to\boldsymbol{0}$$

to mean that

$$\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{g}(\boldsymbol{u}) = O(|\boldsymbol{u}|^r) \quad \text{as } \boldsymbol{u}\to\boldsymbol{0},$$

and the same for the order symbol o. We sometimes indicate the limit $\boldsymbol{u} \to \boldsymbol{0}$ using the informal terminology 'for $\boldsymbol{u}$ small'.

3.2 Differentiation of tensor fields

In this module, a *field* is a function defined in a region of $\mathbb{E}^3$. A *scalar field* means a function $\phi : \mathbb{E}^3 \to \mathbb{R}$, a *vector field* means a function $\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$, and a *second-order tensor field* means a function $\boldsymbol{S} : \mathbb{E}^3 \to \mathcal{V}^2$. This section will discuss the definitions and properties of various differential operators we may apply to these fields.

3.2.1 Derivatives, gradients

The derivative or gradient of a field is the central notion of differentiation: other operations will be based on these basic definitions.

A scalar field $\phi : \mathbb{E}^3 \to \mathbb{R}$ is said to be differentiable at $\boldsymbol{x} \in \mathbb{E}^3$ if there exists a vector $\nabla\phi(\boldsymbol{x}) \in \mathcal{V}$ such that

$$\phi(\boldsymbol{x} + \boldsymbol{h}) = \phi(\boldsymbol{x}) + \nabla\phi(\boldsymbol{x}) \cdot \boldsymbol{h} + o(|\boldsymbol{h}|),$$

or equivalently, such that

$$\nabla\phi(\boldsymbol{x}) \cdot \boldsymbol{a} = \frac{d}{d\alpha}\phi(\boldsymbol{x} + \alpha\boldsymbol{a})\bigg|_{\alpha=0} \quad \text{for all } \boldsymbol{a} \in \mathcal{V},$$

where $\alpha \in \mathbb{R}$. $\nabla\phi(\boldsymbol{x})$ is called the *derivative* or *gradient* of ϕ at the point $\boldsymbol{x}$.

A vector field $\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ is said to be differentiable at $\boldsymbol{x} \in \mathbb{E}^3$ if there exists a second-order tensor $\nabla\boldsymbol{v}(\boldsymbol{x}) \in \mathcal{V}^2$ such that

$$\boldsymbol{v}(\boldsymbol{x} + \boldsymbol{h}) = \boldsymbol{v}(\boldsymbol{x}) + \nabla\boldsymbol{v}(\boldsymbol{x})\boldsymbol{h} + o(|\boldsymbol{h}|),$$

or equivalently, such that

$$\nabla\boldsymbol{v}(\boldsymbol{x})\boldsymbol{a} = \frac{d}{d\alpha}\boldsymbol{v}(\boldsymbol{x} + \alpha\boldsymbol{a})\bigg|_{\alpha=0} \quad \text{for all } \boldsymbol{a} \in \mathcal{V},$$

where $\alpha \in \mathbb{R}$. $\nabla\boldsymbol{v}(\boldsymbol{x})$ is again called the *derivative* or *gradient* of $\boldsymbol{v}$ at the point $\boldsymbol{x}$.

Proposition 3.1 (Gradients in coordinates)**.** *Suppose that $\phi : \mathbb{E}^3 \to \mathbb{R}$ and $\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ are differentiable at $\boldsymbol{x}$, and let $\{\boldsymbol{e}_i\}$ be an arbitrary Cartesian coordinate frame. Then*

$$\nabla\phi(\boldsymbol{x}) = \frac{\partial\phi}{\partial x_i}(\boldsymbol{x})\boldsymbol{e}_i \quad \textit{and} \quad \nabla\boldsymbol{v}(\boldsymbol{x}) = \frac{\partial v_i}{\partial x_j}(\boldsymbol{x})\boldsymbol{e}_i \otimes \boldsymbol{e}_j,$$

where $\boldsymbol{v}(\boldsymbol{x}) = v_i(\boldsymbol{x})\boldsymbol{e}_i$ and x_i are the coordinates of $\boldsymbol{x}$ in the frame considered.

Proof. Writing ϕ and v_i as functions of the coordinates x_i, we have $\phi(\boldsymbol{x}) = \phi(x_1, x_2, x_3)$ and $v_i(\boldsymbol{x}) = v_i(x_1, x_2, x_3)$. If $\boldsymbol{a} = a_k \boldsymbol{e}_k$, then we can use the chain rule to find

$$\nabla\phi(\boldsymbol{x}) \cdot \boldsymbol{a} = \frac{d}{d\alpha}\phi(\boldsymbol{x} + \alpha\boldsymbol{a})\bigg|_{\alpha=0} = \frac{\partial\phi}{\partial x_i}(\boldsymbol{x})a_i = \frac{\partial\phi}{\partial x_i}(\boldsymbol{x})\boldsymbol{e}_i \cdot a_k\boldsymbol{e}_k.$$

Since this must hold for all $\boldsymbol{a}$, the first result follows. Similarly, we find

$$\frac{d}{d\alpha}v_i(\boldsymbol{x} + \alpha\boldsymbol{a})\bigg|_{\alpha=0} = \frac{\partial v_i}{\partial x_j}(\boldsymbol{x})a_j = \frac{\partial v_i}{\partial x_j}(\boldsymbol{x})\boldsymbol{e}_j \cdot a_k\boldsymbol{e}_k,$$

and so

$$\nabla\boldsymbol{v}(\boldsymbol{x})\boldsymbol{a} = \boldsymbol{e}_i\frac{d}{d\alpha}v_i(\boldsymbol{x} + \alpha\boldsymbol{a})\bigg|_{\alpha=0} = \left(\frac{\partial v_i}{\partial x_j}(\boldsymbol{x})\boldsymbol{e}_i \otimes \boldsymbol{e}_j\right)a_k\boldsymbol{e}_k.$$

Again, since this holds for all $\boldsymbol{a} \in \mathcal{V}$, we obtain the desired result. □

Important note: We will often use the shorthand $\phi_{,i}$ to denote $\frac{\partial\phi}{\partial x_i}$ and $v_{i,j}$ to denote $\frac{\partial v_i}{\partial x_j}$.

3.2.2 Divergence

The *divergence* of a vector field $\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ is the scalar field $\nabla \cdot \boldsymbol{v} : \mathbb{E}^3 \to \mathbb{R}$ defined to be

$$\nabla \cdot \boldsymbol{v} = \operatorname{tr}(\nabla\boldsymbol{v}).$$

If $\boldsymbol{v}$ is the velocity field in a flowing fluid, then $(\nabla \cdot \boldsymbol{v})(\boldsymbol{x})$ can be interpreted as the rate of volume expansion at the point $\boldsymbol{x}$.

Similarly, we can also define a divergence for second-order tensor fields. If $\boldsymbol{S} : \mathbb{E}^3 \to \mathcal{V}^2$, then the vector field $\nabla \cdot \boldsymbol{S} : \mathbb{E}^3 \to \mathcal{V}$ is defined to be the vector field which satisfies

$$(\nabla \cdot \boldsymbol{S}) \cdot \boldsymbol{a} = \nabla \cdot (\boldsymbol{S}^T\boldsymbol{a})$$

for all constant vectors $\boldsymbol{a} \in \mathcal{V}$. The field $\nabla \cdot \boldsymbol{S}$ is again called the *divergence* of $\boldsymbol{S}$.

The following result provides a coordinate expression for these two notions of divergence in any coordinate frame.

Proposition 3.2 (Divergence in coordinates)**.** *Let $\{\boldsymbol{e}_i\}$ be an arbitrary coordinate frame, and let $\boldsymbol{v}(\boldsymbol{x}) = v_i(\boldsymbol{x})\boldsymbol{e}_i$, and $\boldsymbol{S}(\boldsymbol{x}) = S_{ij}(\boldsymbol{x})\boldsymbol{e}_i \otimes \boldsymbol{e}_j$. Then*

$$(\nabla \cdot \boldsymbol{v})(\boldsymbol{x}) = \frac{\partial v_i}{\partial x_i}(\boldsymbol{x}) = v_{i,i}(\boldsymbol{x}) \quad \textit{and} \quad (\nabla \cdot \boldsymbol{S})(\boldsymbol{x}) = \frac{\partial S_{ij}}{\partial x_j}(\boldsymbol{x})\boldsymbol{e}_i = S_{ij,j}\boldsymbol{e}_i,$$

where x_i are the coordinates of $\boldsymbol{x}$ with respect to the frame.

Proof. We prove only the second expression, since the first follows directly from Proposition 3.1 and the component expression for the trace. Fix $\boldsymbol{a} = a_k \boldsymbol{e}_k$ and let $\boldsymbol{q} = \boldsymbol{S}^T \boldsymbol{a}$. In components, we have $\boldsymbol{q} = q_j \boldsymbol{e}_j$ with $q_j = S_{ij} a_i$. Then, using the definition and the first expression, we have

$$(\nabla \cdot \boldsymbol{S}) \cdot \boldsymbol{a} = \nabla \cdot \boldsymbol{q} = q_{i,i} = S_{ij,j} a_i = \left(s_{ij,j} \boldsymbol{e}_i\right) \cdot a_k \boldsymbol{e}_k.$$

Since this holds for any $\boldsymbol{a} \in \mathcal{V}$, we have that $\nabla \cdot \boldsymbol{S} = S_{ij,j} \boldsymbol{e}_i$. □

Proposition 3.3 (Product rules). *Let ϕ, $\boldsymbol{v}$ and $\boldsymbol{S}$ respectively be scalar, vector and second-order tensor fields. Then*

$$\begin{aligned} \nabla(\phi \boldsymbol{v}) &= \boldsymbol{v} \otimes \nabla\phi + \phi \nabla \boldsymbol{v}, \\ \nabla \cdot (\phi \boldsymbol{S}) &= \phi \nabla \cdot \boldsymbol{S} + \boldsymbol{S} \nabla \phi, \\ \nabla \cdot (\boldsymbol{S}^T \boldsymbol{v}) &= (\nabla \cdot \boldsymbol{S}) \cdot \boldsymbol{v} + \boldsymbol{S} : \nabla \boldsymbol{v}. \end{aligned}$$

Exercise: Prove this proposition by using the component expressions given above.

3.2.3 Curl

The *curl* of a vector field $\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ is the vector field $\nabla \times \boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ which satisfies

$$(\nabla \times \boldsymbol{f}) \times \boldsymbol{a} = (\nabla \boldsymbol{v} - \nabla \boldsymbol{v}^T) \boldsymbol{a}$$

for any constant vector $\boldsymbol{a} \in \mathcal{V}$.

If $\boldsymbol{v}$ is the velocity field of a flowing fluid, then the curl provides information on the rate and direction of rotation of this field. In practice, we use the following result to compute the curl in a particular coordinate frame.

Proposition 3.4 (Curl in coordinates). *Let $\{\boldsymbol{e}_i\}$ be an arbitrary Cartesian coordinate frame, and let $\boldsymbol{v}(\boldsymbol{x}) = v_i(\boldsymbol{x}) \boldsymbol{e}_i$. Then*

$$(\nabla \times \boldsymbol{v})(\boldsymbol{x}) = \epsilon_{ijk} v_{i,k}(\boldsymbol{x}) \boldsymbol{e}_j,$$

where x_i are the coordinates of $\boldsymbol{x}$ with respect to the frame. Equivalently, we have

$$\nabla \times \boldsymbol{v} = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}\right) \boldsymbol{e}_1 + \left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}\right) \boldsymbol{e}_2 + \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) \boldsymbol{e}_3.$$

3.2.4 Laplacian

The *Laplacian* of a scalar field $\phi : \mathbb{E}^3 \to \mathbb{R}$ is the scalar field $\nabla^2\phi : \mathbb{E}^3 \to \mathbb{R}$ defined via

$$\nabla^2\phi = \nabla \cdot (\nabla\phi).$$

Likewise, the *Laplacian* of a vector field $\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ is the vector field $\nabla^2\boldsymbol{v} : \mathbb{E}^3 \to \mathcal{V}$ defined by

$$\nabla^2\boldsymbol{v} = \nabla \cdot (\nabla\boldsymbol{v}).$$

We remark that in alternative standard notation the Laplacian is also often denoted Δ, i.e. $\Delta\phi = \nabla^2\phi$. In coordinates, we have the following representation.

Proposition 3.5 (Laplacian in coordinates)**.** *Let $\{\boldsymbol{e}_i\}$ be an arbitrary Cartesian coordinate frame, and let $\boldsymbol{v}(\boldsymbol{x}) = v_i(\boldsymbol{x})\boldsymbol{e}_i$. Then*

$$\nabla^2\phi(\boldsymbol{x}) = \phi_{,ii}(\boldsymbol{x}) \quad \textit{and} \quad \nabla^2\boldsymbol{v}(\boldsymbol{x}) = v_{i,jj}(\boldsymbol{x})\boldsymbol{e}_i,$$

where x_i are the coordinates of $\boldsymbol{x}$ with respect to the frame.

Exercise: Prove this proposition.

3.3 Some important integral theorems

In this section, we provide statements of various important theorems about integrals. We will use these theorems to convert between integral and differential forms of the balance laws we derive later, and to prove various other facts about the physical tensors we study.

3.3.1 Divergence Theorem

Recall that the divergence theorem for vector fields connects the volume integral of the divergence of a field over a region to the total flux across the boundary of the region. Various versions of this theorem can be proved under different assumptions on the properties of the region. The particular class of regions $B \subset \mathbb{E}^3$ we will consider are those which are *regular* in the sense that:

- B consists of a finite number of disjoint, open and bounded components;
- The bounding surface ∂B is piecewise smooth and consists of a finite number of disjoint components; and
- Each component of ∂B is orientable in the sense that it has two clearly distinct sides.

Examples of regions that do fit these assumptions include pretty much all shapes we would like to consider for the purposes of modelling, but it is worth thinking about what sorts of regions do not fit these assumptions!

Proposition 3.6 (Vector divergence theorem)**.** *Suppose B is a regular region in $\mathbb{E}^3$ with a piecewise smooth boundary ∂B, and consider a vector field $\boldsymbol{v} : B \to \mathcal{V}$. Then*

$$\int_{\partial B} \boldsymbol{v} \cdot \boldsymbol{n} \, \mathrm{d}A_{\boldsymbol{x}} = \int_B \nabla \cdot \boldsymbol{v} \, \mathrm{d}V_{\boldsymbol{x}},$$

or in components,

$$\int_{\partial B} v_i n_i \, \mathrm{d}A_{\boldsymbol{x}} = \int_B v_{i,i} \, \mathrm{d}V_{\boldsymbol{x}},$$

where $\boldsymbol{n}$ is the outward-point unit normal field on ∂B. The quantity on the left-hand side of these equations is called the flux *of $\boldsymbol{v}$ across the oriented surface ∂B.*

We note that this theorem allows us to understand what the divergence of a vector field really means. Consider a point $\boldsymbol{y}$, and the ball of radius δ around that point, $B_\delta(\boldsymbol{y})$. If $\boldsymbol{v}$ is smooth, we can use the Divergence theorem and Taylor expand the argument in the volume integral to conclude that

$$\begin{aligned}\int_{\partial B_\delta(\boldsymbol{y})} \boldsymbol{v}(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x}) \, \mathrm{d}A_{\boldsymbol{x}} &= \int_{B_\delta(\boldsymbol{y})} (\nabla \cdot \boldsymbol{v})(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}} \\ &= \int_{B_\delta(\boldsymbol{y})} \big[(\nabla \cdot \boldsymbol{v})(\boldsymbol{y}) + O(\delta)\big] \, \mathrm{d}V_{\boldsymbol{x}} \\ &= \mathrm{vol}\big(B_\delta(\boldsymbol{y})\big)\big[(\nabla \cdot \boldsymbol{v})(\boldsymbol{y}) + O(\delta)\big].\end{aligned}$$

This implies that, in the limit as $\delta \to 0$,

$$(\nabla \cdot \boldsymbol{v})(\boldsymbol{y}) = \lim_{\delta \to 0} \frac{1}{\mathrm{vol}(B_\delta(\boldsymbol{y}))} \int_{\partial B_\delta(\boldsymbol{y})} \boldsymbol{v} \cdot \boldsymbol{n} \, \mathrm{d}A_{\boldsymbol{x}}.$$

If we interpret $\boldsymbol{v}$ as a velocity field of a fluid, then the flux integral is the volume of fluid per unit time leaving $B_\delta(\boldsymbol{y})$. When divided by the volume, we therefore obtain the proportion of fluid leaving $B_\delta(\boldsymbol{y})$ per unit time, and so $(\nabla \cdot \boldsymbol{v})(\boldsymbol{y})$ is a measure of the volume expansion at the point $\boldsymbol{y}$.

We can generalise the divergence theorem to second-order tensors as follows:

Proposition 3.7 (Tensor divergence theorem)**.** *Suppose B is a regular region in $\mathbb{E}^3$ with a piecewise smooth boundary ∂B, and consider a second-order tensor field $\boldsymbol{S} : B \to \mathcal{V}^2$. Then*

$$\int_{\partial B} \boldsymbol{S}\boldsymbol{n} \, \mathrm{d}A_{\boldsymbol{x}} = \int_B \nabla \cdot \boldsymbol{S} \, \mathrm{d}V_{\boldsymbol{x}},$$

or in components,

$$\int_{\partial B} S_{ij} n_j \, \mathrm{d}A_{\boldsymbol{x}} = \int_B S_{ij,j} \, \mathrm{d}V_{\boldsymbol{x}},$$

where $\boldsymbol{n}$ is the outward-point unit normal field on ∂B.

3.3.2 Further useful theorems from Analysis

For some purposes we will need the following two theorems. The first, a localisation theorem, states that if the integral of a scalar field on any subset of its domain is zero, then the field itself must be zero everywhere. The second provides forms of the Mean Value Theorem for multidimensional integrals.

Proposition 3.8 (Localisation theorem)**.** *Consider a continuous function $\phi : B \to \mathbb{R}$ where B is an open subset of $\mathbb{E}^3$, and let Ω denote an arbitrary open subset of B. If*

$$\int_\Omega \phi(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} = 0 \quad \textit{for any } \Omega \subseteq B,$$

then it follows that $\phi(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in B$.

Exercise: Prove this by contradiction (assume that ϕ is non-zero somewhere, and proceed from here).

Proposition 3.9 (Mean Value Theorem for surfaces)**.** *Let B be an open, connected subset of $\mathbb{E}^3$, and let Σ be a compact, connected surface in B. Suppose that $\phi : B \to \mathbb{R}$ is continuous; then there exist points $\boldsymbol{x}' \in B$ and $\boldsymbol{x}'' \in \Sigma$ such that*

$$\phi(\boldsymbol{x}') = \frac{1}{\mathrm{vol}(B)} \int_B \phi(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} \quad \textit{and} \quad \phi(\boldsymbol{x}'') = \frac{1}{\mathrm{area}(\Sigma)} \int_\Sigma \phi(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}}.$$

Exercise: Prove this. The method of proof is similar to the version you will have seen in Analysis, it just requires some more thought due to the multidimensional setting.

3.4 Functions of second-order tensors

In this section, we generalise the notions of derivatives which you know for functions which take scalars and vectors as input to the case where functions take second-order tensors a input. In general we consider functions $\psi : \mathcal{V}^2 \to \mathbb{R}$; examples of such functions include the trace, determinant and norm defined above.

3.4.1 Scalar-valued functions of second-order tensors

A function $\psi : \mathcal{V}^2 \to \mathbb{R}$ is said to be differentiable at $\boldsymbol{A} \in \mathcal{V}^2$ if there exists a second-order tensor $D\psi(\boldsymbol{A}) \in \mathcal{V}^2$ such that

$$\psi(\boldsymbol{A} + \boldsymbol{H}) = \psi(\boldsymbol{A}) + D\psi(\boldsymbol{A}) : \boldsymbol{H} + o(|\boldsymbol{H}|),$$

or equivalently,

$$D\psi(\boldsymbol{A}) : \boldsymbol{B} = \frac{d}{d\alpha}\psi(\boldsymbol{A} + \alpha\boldsymbol{B})\bigg|_{\alpha=0} \quad \text{for all } \boldsymbol{B} \in \mathcal{V}^2,$$

where $\alpha \in \mathbb{R}$. The tensor $D\psi(\boldsymbol{A})$ is called the *derivative* of ψ at $\boldsymbol{A}$.

If ψ is differentiable at $\boldsymbol{A}$ the derivative is unique, a fact which can be shown in much the same way as for the derivative of a real-valued function. The second characterisation follows from the first by choosing $\boldsymbol{H} = \alpha\boldsymbol{B}$, dividing by α and letting $\alpha \to 0$.

The following result provides a characterisation of the derivative in any coordinate frame.

Proposition 3.10 (Differentiating scalar functions of tensors)**.** *Let $\{\boldsymbol{e}_i\}$ be an arbitrary coordinate frame, and $\psi : \mathcal{V}^2 \to \mathbb{R}$. Then*

$$D\psi(\boldsymbol{A}) = \frac{\partial\psi}{\partial A_{ij}}(\boldsymbol{A})\boldsymbol{e}_i \otimes \boldsymbol{e}_j,$$

where A_{ij} are the components of $\boldsymbol{A} \in \mathcal{V}^2$ in the frame $\{\boldsymbol{e}_i\}$.

Proof. Writing ψ as a function of the components A_{ij} (and abusing notation), we have $\psi(\boldsymbol{A}) = \psi(A_{11}, A_{12} \ldots, A_{33})$. For any scalar α and tensor $\boldsymbol{B} = B_{kl}\boldsymbol{e}_k \otimes \boldsymbol{e}_l$, this gives

$$\psi(\boldsymbol{A} + \alpha\boldsymbol{B}) = \psi(A_{11} + \alpha B_{11}, \ldots, A_{33} + \alpha B_{33}).$$

Applying the chain rule, we have

$$\begin{aligned} D\psi(\boldsymbol{A}) : \boldsymbol{B} &= \frac{d}{d\alpha}\psi(\boldsymbol{A} + \alpha\boldsymbol{B})\bigg|_{\alpha=0} \\ &= \frac{\partial\psi}{\partial A_{ij}}(\boldsymbol{A})B_{ij} \\ &= \left(\frac{\partial\psi}{\partial A_{ij}}(\boldsymbol{A})\boldsymbol{e}_i \otimes \boldsymbol{e}_j\right) : B_{kl}\boldsymbol{e}_k \otimes \boldsymbol{e}_l, \end{aligned}$$

where we have used that $(\boldsymbol{e}_i \otimes \boldsymbol{e}_j) : (\boldsymbol{e}_k \otimes \boldsymbol{e}_l) = \delta_{ik}\delta_{jl}$. This implies the result. □

Note that the derivative of a scalar-valued function of a second-order tensor is a second-order tensor valued function, $D\psi : \mathcal{V}^2 \to \mathcal{V}^2$. Also, if third-order partial derivatives of ψ exist, then we can use Taylor's Theorem to deduce that for any small (in the sense of the norm) tensor $\boldsymbol{H} \in \mathcal{V}^2$, we have

$$\psi(\boldsymbol{A} + \boldsymbol{H}) = \psi(\boldsymbol{A}) + \frac{\partial\psi}{\partial A_{ij}}(\boldsymbol{A}) : H_{ij} + O(|\boldsymbol{H}|^2),$$

or equivalently

$$\psi(\boldsymbol{A}+\boldsymbol{H}) = \psi(\boldsymbol{A}) + D\psi(\boldsymbol{A}) : \boldsymbol{H} + O(|\boldsymbol{H}|^2),$$

which is strong than the assumption that the difference is $o(|\boldsymbol{H}|)$ as indicated in the definition.

Example: Consider the function $\psi : \mathcal{V}^2 \to \mathbb{R}$ defined by

$$\psi(\boldsymbol{A}) = \tfrac{1}{2}\boldsymbol{A} : \boldsymbol{A} = \tfrac{1}{2}A_{kl}A_{kl}.$$

To compute $D\psi(\boldsymbol{A})$, we compute the partial derivative with respect to a general component A_{ij}, obtaining via the product rule that

$$\frac{\partial \psi}{\partial A_{ij}} = \frac{1}{2}\frac{\partial A_{kl}}{\partial A_{ij}}A_{kl} + \frac{1}{2}A_{kl}\frac{\partial A_{kl}}{\partial A_{ij}} = \frac{\partial A_{kl}}{\partial A_{ij}}A_{kl}.$$

Now, since $\frac{\partial A_{kl}}{\partial A_{ij}} = \delta_{ik}\delta_{jl}$, we have

$$\frac{\partial \psi}{\partial A_{ij}} = \delta_{ik}\delta_{jl}A_{kl} = A_{ij},$$

and hence we deduce $D\psi(\boldsymbol{A}) = A_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j = \boldsymbol{A}$.

Proposition 3.11 (Derivative of determinant)**.** *Let $\psi(\boldsymbol{S}) = \det \boldsymbol{S}$. If $\boldsymbol{A}$ is invertible, then the derivative of ψ at $\boldsymbol{A}$ is*

$$D\psi(\boldsymbol{A}) = \det(\boldsymbol{A})\boldsymbol{A}^{-T}.$$

The second-order tensor $\det(\boldsymbol{A})\boldsymbol{A}^{-T}$ is also called the cofactor *tensor, denoted $\operatorname{cof}(\boldsymbol{A}) = \det(\boldsymbol{A})\boldsymbol{A}^{-T}$.*

Proof. Let $\boldsymbol{B} \in \mathcal{V}^2$ be arbitrary, and set $\lambda = -\frac{1}{\alpha}$. Then

$$\begin{aligned}\psi(\boldsymbol{A}+\alpha\boldsymbol{B}) &= \det(\boldsymbol{A}+\alpha\boldsymbol{B}) \\ &= \det\big(\alpha\boldsymbol{A}(\boldsymbol{A}^{-1}\boldsymbol{B} - \lambda\boldsymbol{I})\big) \\ &= \det\big(\alpha\boldsymbol{A}\big)\det(\boldsymbol{A}^{-1}\boldsymbol{B} - \lambda\boldsymbol{I}).\end{aligned}$$

For the first factor above, we note that $\det(\alpha\boldsymbol{A}) = \alpha^3 \det \boldsymbol{A}$. The second factor is the characteristic polynomial of $\boldsymbol{A}^{-1}\boldsymbol{B}$, and so recalling the facts about the principal invariants stated in Section 2.4.9, we have

$$\begin{aligned}\det(\boldsymbol{A}^{-1}\boldsymbol{B} - \lambda\boldsymbol{I}) &= -\lambda^3 + \lambda^2 I_1(\boldsymbol{A}^{-1}\boldsymbol{B}) - \lambda I_2(\boldsymbol{A}^{-1}\boldsymbol{B}) + I_3(\boldsymbol{A}^{-1}\boldsymbol{B}) \\ &= \frac{1}{\alpha^3} + \frac{1}{\alpha^2}I_1(\boldsymbol{A}^{-1}\boldsymbol{B}) + \frac{1}{\alpha}I_2(\boldsymbol{A}^{-1}\boldsymbol{B}) + I_3(\boldsymbol{A}^{-1}\boldsymbol{B}),\end{aligned}$$

and hence

$$\psi(\boldsymbol{A}+\alpha\boldsymbol{B}) = \det\boldsymbol{A}\big(1 + \alpha I_1(\boldsymbol{A}^{-1}\boldsymbol{B}) + \alpha^2 I_2(\boldsymbol{A}^{-1}\boldsymbol{B}) + \alpha^3 I_3(\boldsymbol{A}^{-1}\boldsymbol{B})\big).$$

Differentiating with respect to α and evaluating at $\alpha = 0$, we have

$$D\psi(\boldsymbol{A}) : \boldsymbol{B} = \det(\boldsymbol{A}) I_1(\boldsymbol{A}^{-1}\boldsymbol{B}).$$

Since $I_1(\boldsymbol{A}^{-1}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A}^{-1}\boldsymbol{B}) = \boldsymbol{A}^{-T} : \boldsymbol{B}$, we deduce that

$$D\psi(\boldsymbol{A}) : \boldsymbol{B} = \det(\boldsymbol{A}) \boldsymbol{A}^{-T} : \boldsymbol{B},$$

and so the result follows from the arbitrariness of $\boldsymbol{B}$. □

The following result is a handy consequence of the result above.

Proposition 3.12 (Time derivative of determinant)**.** *Let $\boldsymbol{S} : \mathbb{R} \to \mathcal{V}^2$ be a time-dependent second-order tensor. Then*

$$\frac{d}{dt} \det \boldsymbol{S} = \det(\boldsymbol{S}) \mathrm{tr} \left(\boldsymbol{S}^{-1} \frac{d\boldsymbol{S}}{dt} \right) = \det(\boldsymbol{S}) \boldsymbol{S}^{-T} : \frac{d\boldsymbol{S}}{dt},$$

where

$$\frac{d\boldsymbol{S}}{dt} = \frac{dS_{ij}}{dt} \boldsymbol{e}_i \otimes \boldsymbol{e}_j.$$

Exercise: Prove this result using the chain rule.

4 Mass, Forces and Stress

With a large amount of mathematical background now out of the way, we are able to begin thinking more carefully about the important physical concepts which will be central to developing models of solids and fluids. The central assumption we will make here is that it is reasonable to treat the objects we consider as *continuum* objects, i.e. that the objects are infinitely divisible, just like the real numbers that we use to describe them.

This assumption is of course false, since we know that solids and fluids are made up of atoms, which are not themselves easily divisible in everyday contexts. Nevertheless, the equations of continuum mechanics are one of the great triumphs of mathematical modelling, since the predictions which can made with them are highly accurate, and hold even down to scales which are only just above the atomistic length scale. Exploring the extent of their validity, and the connections between atomistic and continuum models remains an active research topic.

Aims. By the end of this chapter, you should be able to:

- Define what is meant by the terms *body, mass density, centre of mass, body force, resultant force, resultant torque, surface force,* and *traction.*
- State Cauchy's Postulate, and use it to show the existence of the Cauchy stress field.
- State the definition of static mechanical equilibrium, and derive the local equations satisfied by the Cauchy stress field in this case.
- Explain the concept of stress, and use various terms to describe particular stress states.

4.1 Bodies

Our assumption that a material body is a continuum means that we can identify the body with B, an open subset of Euclidean space $\mathbb{E}^3$. Each point $\boldsymbol{x} \in B$ is identified with a material *particle*. Note here that our use of the word particle does not refer to an atom, but rather an infinitesimal part of the body. We call B a configuration of the body in $\mathbb{E}^3$. Unless explicitly mentioned otherwise, we will always assume that B is a regular region of $\mathbb{E}^3$.

Our own experience tells us that material bodies move and change shape under the effects of physical processes, and we will indeed consider such effects later, which will necessitate thinking about how the body varies. For the moment however, we will ignore time variation, and focus only on a single configuration.

4.1.1 Volume and area

We recall that the volume $\mathrm{vol}(\Omega)$ of a set $\Omega \subseteq \mathbb{E}^3$ can be computed by integration, so that

$$\mathrm{vol}(\Omega) = \int_\Omega \mathrm{d}V_{\boldsymbol{x}}.$$

Similarly, the area $\mathrm{area}(\Gamma)$ of a surface $\Gamma \subset \mathbb{E}^3$ can also be computed by integration, so that

$$\mathrm{area}(\Gamma) = \int_\Gamma \mathrm{d}A_{\boldsymbol{x}}.$$

4.2 Mass, momentum, forces and torque

Mass is the property of matter which quantifies its resistance to acceleration, and forces are the influences acting on an object which may result in an acceleration. In classical mechanics, this is exemplified by Newton's second law of motion for objects with constant mass, which should be familiar:

$$\text{Net force acting} = \text{Mass} \times \text{Acceleration},$$
$$\boldsymbol{F} = m\boldsymbol{a}.$$

In its more general form, Newton's second law states that

$$\text{Net force acting} = \text{Rate of change of momentum},$$
$$\boldsymbol{F} = \frac{d\boldsymbol{p}}{dt} = \frac{d}{dt}(m\boldsymbol{v}),$$

where the (linear) momentum $\boldsymbol{p}$ is the product of the mass and the velocity of an object. Using the first of these relations, we see that the greater the mass of an object, the greater the force that is required to accelerate that object at the same rate.

When they are first introduced, the concepts of mass and force are typically applied to point particles. Here, we will consider bodies which are extended, and so as concepts, the mass of a single point or a force acting on a single point makes little sense. Instead, we introduce fields representing densities of mass and forces. To represent the mass, we define a *mass density*, and we focus on two particular types of forces: *body force* fields which act on points within the object, and *surface force* fields, which are act on surfaces, either within the object or on its exterior.

4.2.1 Mass density

We assume will assume that the mass of a body B is distributed across its volume, that any subset of B with non-zero volume must also have non-zero mass, and that the mass of any region of the body tends to zero as the volume of the region tends to zero.

If $\Omega \subseteq B$ is an open subset of B, we denote the mass of this subset $\mathrm{mass}(\Omega)$. We assume that there exists a *mass density field* $\rho : B \to \mathbb{R}$, with units of mass per unit volume, such that

$$\mathrm{mass}(\Omega) = \int_\Omega \rho(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}}.$$

Theoretically-speaking, if we know the mass and volume functions, we can formally define the mass density field ρ as follows. Let $\boldsymbol{x} \in B$ and let $\Omega_\delta(\boldsymbol{x}) \subseteq B$ denote a family of volumes with $\boldsymbol{x} \in \Omega_\delta(\boldsymbol{x})$ for all $\delta > 0$, and $\mathrm{vol}(\Omega_\delta(\boldsymbol{x})) \to 0$ as $\delta \to 0$. Then

$$\rho(\boldsymbol{x}) = \lim_{\delta \to 0} \frac{\mathrm{mass}(\Omega_\delta(\boldsymbol{x}))}{\mathrm{vol}(\Omega_\delta(\boldsymbol{x}))}.$$

Our basic assumptions on the mass distribution is that this limit exists, is strictly postive, i.e. $\rho(\boldsymbol{x}) > 0$ for each $\boldsymbol{x} \in B$, and this limit is independent of the precise choice of family $\Omega_\delta(\boldsymbol{x})$ which have the properties described.

4.2.2 Centre of Mass

The *centre of mass* and the *centre of volume* of an open subset $\Omega \subseteq B$ respectively mean the points

$$\mathrm{com}(\Omega) = \frac{1}{\mathrm{mass}(\Omega)} \int_\Omega \boldsymbol{x} \rho(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}}$$
$$\text{and} \quad \mathrm{cov}(\Omega) = \frac{1}{\mathrm{vol}(\Omega)} \int_\Omega \boldsymbol{x} \, \mathrm{d}V_{\boldsymbol{x}}.$$

In general, these points need not be points in B (although they will be if Ω is a convex set).

4.2.3 Body forces

Body forces are those forces which do not arise due to direct physical contact between bodies. Prototypical examples include gravity, or electromagnetic forces. The body force per unit volume exerted on a body B is assumed to be given by a function $\widehat{\boldsymbol{b}} : B \to \mathcal{V}$, called a *body force field per unit volume* acting on B.

Let Ω be an open subset of B. Then the *resultant force* on Ω due to the body force field is defined to be

$$\boldsymbol{r}_b(\Omega) = \int_\Omega \widehat{\boldsymbol{b}}(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}}.$$

The result force measures the total effect of the body forces acting on Ω. This interpretation is clear if we think of the integral as 'summing up' the contributions from the body force acting on every particle $\boldsymbol{x} \in \Omega$.

The *resultant torque about* $\boldsymbol{z}$ on Ω is defined to be

$$\boldsymbol{\tau}_b(\Omega; \boldsymbol{z}) = \int_\Omega (\boldsymbol{x} - \boldsymbol{z}) \times \widehat{\boldsymbol{b}}(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}}.$$

The torque measures the total turning effect of the body forces acting on Ω about the centre of rotation $\boldsymbol{z}$. Here, the integral sums up the contributions of each force moment $(\boldsymbol{x} - \boldsymbol{z}) \times \widehat{\boldsymbol{b}}(\boldsymbol{x})$ over Ω.

Instead of considering body force per unit volume, we can also consider body forces per unit mass. The *body force per unit mass* is

$$\boldsymbol{b}(x) = \frac{\widehat{\boldsymbol{b}}(\boldsymbol{x})}{\rho(\boldsymbol{x})},$$

which is well-defined since we assume $\rho(\boldsymbol{x}) > 0$. In terms of this new field, we have

$$\boldsymbol{r}_b(\Omega) = \int_\Omega \rho(\boldsymbol{x}) \boldsymbol{b}(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}} \quad \text{and} \quad \boldsymbol{\tau}_b(\Omega; \boldsymbol{z}) = \int_\Omega (\boldsymbol{x} - \boldsymbol{z}) \times \rho(\boldsymbol{x}) \boldsymbol{b}(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}}.$$

4.2.4 Angular momentum and torque

Before discussing forces further, we briefly discuss the definition of angular momentum and torque, which are the rotational versions of (linear) momentum and forces.

The *angular momentum* of an object is the rotational equivalent of linear momentum, and is defined to be the vector product of the objects displacement from some chosen point, $\boldsymbol{z}$, and its linear momentum, i.e.:

$$\boldsymbol{l} = (\boldsymbol{x} - \boldsymbol{z}) \times \boldsymbol{p} = m\big((\boldsymbol{x} - \boldsymbol{z}) \times \boldsymbol{v}\big),$$

Here, the point $\boldsymbol{z}$ used to define the displacement is the point about which we wish to consider the object's rotation. This could be the centre of mass (in which case, we consider an object's *spin*), or some other point about which we wish to consider rotation, such as the origin of our coordinate system.

The concept which accompanies the angular momentum and mirrors the notion of a force is the *torque* (or *moment*), which like angular momentum, must be taken about some point $\boldsymbol{z}$. This is denoted $\boldsymbol{\tau}$ and is defined to be

$$\boldsymbol{\tau} = (\boldsymbol{x} - \boldsymbol{z}) \times \boldsymbol{F}.$$

Torques are analogous to forces in that they are additive, and if there is no net torque acting on an object, then there will be no change in the angular momentum. Furthermore,

the equivalent of Newton's second law for the torque and the angular momentum is that the net torque is equal to the rate of change of the angular momentum, i.e.

$$\boldsymbol{\tau} = \frac{d\boldsymbol{l}}{dt},$$

(where both torque and angular momentum are taken about the same point $\boldsymbol{z}$).

4.2.5 Surface forces

A *surface force* is any force that arises due to direct physical contact between bodies. In the case of imaginary internal surfaces inside a body, we say the surface force is *internal*. Internal surface forces are what holds a body together, and they resist the tendency for one part of the body to be pulled away from, or pushed through, another part of the body. When the force acts along the exterior surface of the body, we say the surface force is *external*; these are contact forces applied to a body by its environment.

Understanding surface forces is a key step in unlocking your understanding of continuum mechanics, so you should take care to cover this material carefully.

Suppose that Γ is an oriented surface in B with a unit normal field $\boldsymbol{n} : \Gamma \to \mathcal{V}$. The fact that Γ is oriented allows us to define a positive and a negative side to Γ. The force per unit area exerted by the material on the positive side on the material on the negative side will be assumed to be given by a function $\boldsymbol{t}_\Gamma : \Gamma \to \mathcal{V}$, which we call the *traction* or *surface force field* for Γ. When Γ is part of the boundary ∂B, we always choose $\boldsymbol{n}$ to be the outward unit normal field. In this case, the traction field represents the external force per unit area applied to this portion of the boundary.

The *resultant traction* on Γ is defined to be

$$\boldsymbol{r}_s(\Gamma) = \int_\Gamma \boldsymbol{t}_\Gamma(\boldsymbol{x}) \, \mathrm{d}A_{\boldsymbol{x}},$$

where $\mathrm{d}A_{\boldsymbol{x}}$ represents the surface area element at $\boldsymbol{x} \in \Gamma$. The *resultant torque about* $\boldsymbol{z}$ due to the traction field is defined to be

$$\boldsymbol{\tau}_s(\Gamma; \boldsymbol{z}) = \int_\Gamma (\boldsymbol{x} - \boldsymbol{z}) \times \boldsymbol{t}_\Gamma(\boldsymbol{x}) \, \mathrm{d}A_{\boldsymbol{x}}.$$

4.2.6 Cauchy's postulate

The physical theory of surface forces in classical continuum mechanics is based on the following assumption, known as *Cauchy's postulate*.

Cauchy's Postulate

The traction field $\boldsymbol{t}$ on an oriented surface Γ which passes through a point $\boldsymbol{x} \in B$ depends only on the point $\boldsymbol{x}$, and the unit normal $\boldsymbol{n}(\boldsymbol{x})$ to the surface Γ at that point. In particular, there is a function $\boldsymbol{t} : \mathcal{N} \times B \to \mathcal{V}$ where $\mathcal{N} \subset \mathcal{V}$ is the set of all unit vectors, such that

$$\boldsymbol{t}_\Gamma(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{n}(\boldsymbol{x}), \boldsymbol{x}).$$

The function $\boldsymbol{t} : \mathcal{N} \times B \to \mathcal{V}$ is called the *traction function* for B.

The point of this postulate is that in general, it might hold that the traction $\boldsymbol{t}_\Gamma$ depends on the entirety of the surface Γ, or at the very least upon more details of the geometry of Γ near to $\boldsymbol{x}$, such as the curvature $\nabla \boldsymbol{n}(\boldsymbol{x})$. We could develop more sophisticated theories which take more geometry into account, but Cauchy's postulate is highly successful as an assumption to develop models in a wide range of contexts.

The next result can be interpreted as showing that the traction function satisfies a certain version of Newton's Third Law.

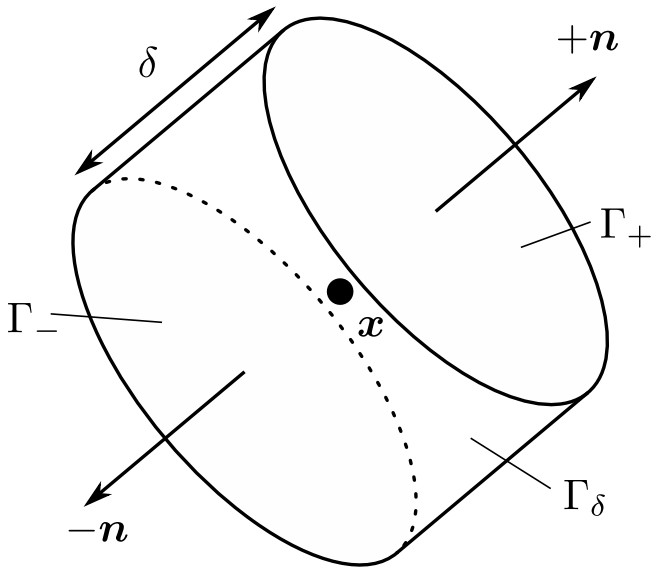


Figure 4.1: Illustration of Ω_δ used in the proof of Proposition 4.1.

Proposition 4.1. *Let $\boldsymbol{t} : \mathcal{N} \times B \to \mathcal{V}$ be the traction function for a body B. Suppose that $\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x})$ is continuous, and for any sequence of subsets Ω whose volumes tend to zero, we have*

$$\frac{1}{\text{area}(\partial\Omega)} \int_{\partial\Omega} \boldsymbol{t}(\boldsymbol{n}(\boldsymbol{x}), \boldsymbol{x}) \, \mathrm{d}A_{\boldsymbol{x}} \to \boldsymbol{0} \quad \textit{as} \quad \text{vol}(\Omega) \to 0. \tag{4.1}$$

Then it follows that

$$\boldsymbol{t}(-\boldsymbol{n}, \boldsymbol{x}) = -\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) \quad \textit{for any } \boldsymbol{n} \in \mathcal{N}, \boldsymbol{x} \in B.$$

Proof. Let $\boldsymbol{x} \in B$ and $\boldsymbol{n} \in \mathcal{N}$ be arbitrary, and let D be a disc of small radius r, centred at $\boldsymbol{x}$ with normal $\boldsymbol{n}$. For $\delta > 0$, let Ω_δ be the cylinder with centre of volume at $\boldsymbol{x}$, height δ, and axis parallel to $\boldsymbol{n}$. Let $\Gamma_\pm$ be the circular faces of the cylinder parallel to D, and the remaining surface be Γ_δ.

We note that $\mathrm{area}(\Gamma_\delta) = 2\pi r\delta$, which vanishes as $\delta \to 0$, and so $\mathrm{area}(\partial\Omega_\delta) \to 2\mathrm{area}(D) = 2\pi r^2$ as $\delta \to 0$. Note also that $\mathrm{vol}(\Omega_\delta) = \delta\pi r^2$, so $\mathrm{vol}(\Omega_\delta) \to 0$ as $\delta \to 0$, so this sequence of volumes fulfils the assumptions of the proposition, and since $\mathrm{area}(\partial\Omega_\delta) > 0$ for all $\delta > 0$, it follows that

$$\lim_{\delta\to 0} \int_{\partial\Omega_\delta} \boldsymbol{t}(\widehat{\boldsymbol{n}}(\boldsymbol{x}), \boldsymbol{x}')\, dA_{\boldsymbol{x}'} = 0$$

by assumption.

Next, note that the points in $\Gamma_\pm$ converge to points in D as $\delta \to 0$. Let $\widehat{\boldsymbol{n}} : \partial\Omega_\delta \to \mathcal{V}$ be the outward unit normal field on $\partial\Omega_\delta$, so $\widehat{\boldsymbol{n}} = \pm\boldsymbol{n}$ is constant on $\Gamma_\pm$. We can therefore express $\partial\Omega_\delta = \Gamma_\delta \cup \Gamma_+ \cup \Gamma_-$, and since $\mathrm{area}(\Gamma_\delta) \to 0$ as $\delta \to 0$ and $\boldsymbol{t}$ is continuous and so bounded, we have

$$\lim_{\delta\to 0} \int_{\Gamma_\delta} \boldsymbol{t}(\widehat{\boldsymbol{n}}(\boldsymbol{x}), \boldsymbol{x}')\, dA_{\boldsymbol{x}'} = \boldsymbol{0}.$$

It follows that

$$\int_D \boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}') + \boldsymbol{t}(-\boldsymbol{n}, \boldsymbol{x}')\, dA_{\boldsymbol{x}'} = \boldsymbol{0}.$$

Now, we apply the result of Proposition 3.9, which allows us to conclude that there is a point $\boldsymbol{x}_r \in D_r$ such that

$$\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}_r) + \boldsymbol{t}(-\boldsymbol{n}, \boldsymbol{x}_r) = \boldsymbol{0}.$$

Since this result holds for all $r > 0$, we can now let $r \to 0$. In this limit, $\boldsymbol{x}_r \to \boldsymbol{x}$, and since $\boldsymbol{t}$ was assumed to be continuous, it follows that

$$\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) + \boldsymbol{t}(-\boldsymbol{n}, \boldsymbol{x}) = \boldsymbol{0}.$$

This establishes the result. □

We have just shown that the traction exerted by material on the positive side of a surface on the negative side is equal and opposite to the traction exerted by the negative side on the positive. This is Newton's Third Law in action.

We note that the proof given above does not make complete use of the full strength of the condition Equation 4.1, since the sequence of volumes we consider does not have a vanishing surface area, but we will use the full strength of this condition in the results which follow.

4.3 The Cauchy stress tensor

Using Proposition 4.1 allows us to say more about the dependence of the function $\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x})$ upon $\boldsymbol{n}$. This result is often called *Cauchy's Theorem*, and is fundamental to Continuum Mechanics.

Proposition 4.2 (Stress Tensor)**.**

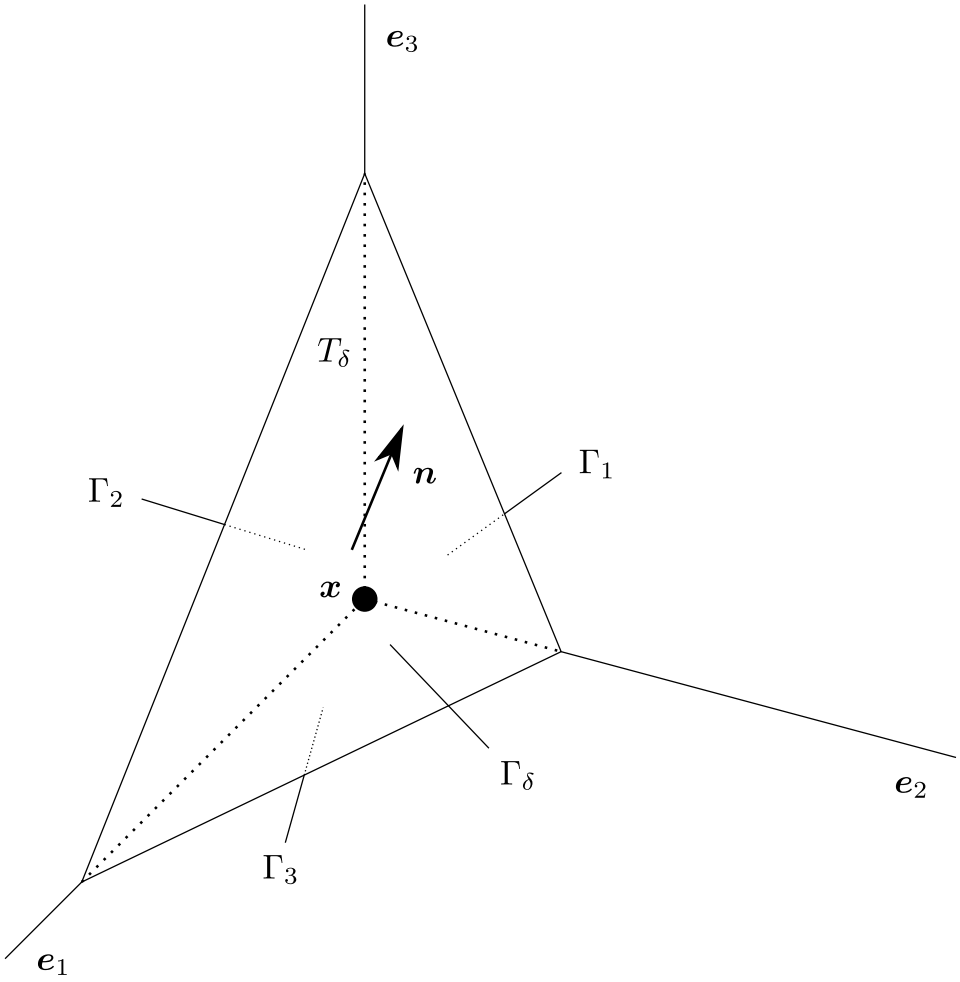


Figure 4.2: The tetrahedral region used in the proof of Proposition 4.2.

Let $\boldsymbol{t} : \mathcal{N} \times B \to \mathcal{V}$ be the traction function for a body B which satisfies the conditions of Proposition 4.1. Then $\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x})$ is a linear function of $\boldsymbol{n}$, i.e. for each $\boldsymbol{x} \in B$, there exists a second-order tensor $\boldsymbol{S}(\boldsymbol{x}) \in \mathcal{V}^2$ such that

$$\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) = \boldsymbol{S}(\boldsymbol{x})\boldsymbol{n}.$$

The field $\boldsymbol{S} : B \to \mathcal{V}^2$ *is called the* Cauchy stress field *for* B.

Proof. We will briefly suspend the summation convention for the first section of this proof.

Consider an arbitrary Cartesian coordinate frame $\{\boldsymbol{e}_i\}$, and suppose that $\boldsymbol{n} \cdot \boldsymbol{e}_i > 0$ for all i. Define the tetrahedral region

$$T_\delta = \big\{\boldsymbol{x}' \in B : 0 \leq (\boldsymbol{x}' - \boldsymbol{x}) \cdot \boldsymbol{e}_i \text{ and } (\boldsymbol{x}' - \boldsymbol{x}) \cdot \boldsymbol{n} \leq \delta\big\},$$

which is illustrated in Figure 1.1. T_δ has four faces; let the three faces with outward-pointing unit normal $-\boldsymbol{e}_i$ be denoted $\Gamma_{\delta,i}$, and let the final face be denoted $\Gamma_{\delta,n}$, which has outward-pointing unit normal $\boldsymbol{n}$.

We note that $\mathrm{vol}(T_\delta) \to 0$ as $\delta \to 0$, and hence condition Equation 4.1 implies that

$$\lim_{\delta \to 0} \frac{1}{\mathrm{area}(\partial T_\delta)} \int_{\partial T_\delta} \boldsymbol{t}(\widehat{\boldsymbol{n}}(\boldsymbol{x}'), \boldsymbol{x}')\, dA_{\boldsymbol{x}'} = \boldsymbol{0},$$

where $\widehat{\boldsymbol{n}}(\boldsymbol{x}')$ denotes the outward-pointing unit normal field on ∂T_δ. Using the explicit form for this field on each of the faces, we have

$$\lim_{\delta \to 0} \frac{1}{\mathrm{area}(\partial T_\delta)} \left(\int_{\Gamma_{\delta,n}} \boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}')\, dA_{\boldsymbol{x}'} + \sum_{i=1}^{3} \int_{\Gamma_{\delta,i}} \boldsymbol{t}(-\boldsymbol{e}_i, \boldsymbol{x}')\, dA_{\boldsymbol{x}'} \right) = \boldsymbol{0}. \tag{4.2}$$

It can be computed that the area of each face is

$$\mathrm{area}(\Gamma_{\delta,i}) = \frac{\delta^2}{2} \frac{n_i}{n_1 n_2 n_3} \quad \text{and} \quad \mathrm{area}(\Gamma_{\delta,n}) = \frac{\delta^2}{2} \frac{1}{n_1 n_2 n_3},$$

so that

$$\mathrm{area}(\Gamma_\delta) = n_1 \mathrm{area}(\Gamma_1) = n_2 \mathrm{area}(\Gamma_2) = n_3 \mathrm{area}(\Gamma_3), \tag{4.3}$$

and thus

$$\mathrm{area}(\partial T_\delta) = (1 + n_1 + n_2 + n_3)\mathrm{area}(\Gamma_{\delta,n}). \tag{4.4}$$

We now consider the integral over each face, and apply the Mean Value Theorem, which entails that there exist points $\boldsymbol{x}'_{\delta,i} \in \Gamma_{\delta,i}$ and a point $\boldsymbol{x}'_{\delta,n} \in \Gamma_{\delta,n}$ such that

$$\boldsymbol{t}(-\boldsymbol{e}_i, \boldsymbol{x}'_{\delta,i}) = \frac{1}{\mathrm{area}(\Gamma_{\delta,i})} \int_{\Gamma_{\delta,i}} \boldsymbol{t}(-\boldsymbol{e}_i, \boldsymbol{x}')\, dA_{\boldsymbol{x}'}$$
$$\text{and} \quad \boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}'_{\delta,n}) = \frac{1}{\mathrm{area}(\Gamma_{\delta,n})} \int_{\Gamma_{\delta,i}} \boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}')\, dA_{\boldsymbol{x}'}.$$

We note that as $\delta \to 0$, we have that $\boldsymbol{x}'_{\delta,n} \to \boldsymbol{x}$ and $\boldsymbol{x}'_{\delta,i} \to \boldsymbol{x}$, since the faces shrink to the point $\boldsymbol{x}$.

Using these facts and Equation 4.3, we have

$$\int_{\Gamma_{\delta,n}} \boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}')\, dA_{\boldsymbol{x}'} + \sum_{i=1}^{3} \int_{\Gamma_{\delta,i}} \boldsymbol{t}(-\boldsymbol{e}_i, \boldsymbol{x}')\, dA_{\boldsymbol{x}'}$$
$$= \operatorname{area}(\Gamma_{\delta,n}) \left(\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}'_{\delta,n}) + \sum_{i=1}^{3} n_i \boldsymbol{t}(-\boldsymbol{e}_i, \boldsymbol{x}_{\delta,i}) \right).$$

Dividing by $\operatorname{area}(\partial T_\delta)$, and applying Equation 4.4 with the limit, we deduce that

$$\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) + \sum_{i=1}^{3} n_i \boldsymbol{t}(-\boldsymbol{e}_i, \boldsymbol{x}) = \boldsymbol{0}.$$

Applying Proposition 4.1, and reinstating the summation convention, this can be rewritten as

$$\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) = n_i \boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x}) = \Big(\boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x}) \otimes \boldsymbol{e}_i \Big) \boldsymbol{n} = \boldsymbol{S}(\boldsymbol{x}) \boldsymbol{n}$$

where we define

$$\boldsymbol{S}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x}) \otimes \boldsymbol{e}_i. \tag{4.5}$$

This demonstrates the result for all $\boldsymbol{n}$ such that $\boldsymbol{n} \cdot \boldsymbol{e}_i > 0$.

To show the result for all remaining vectors, we can first repeat the argument above for any $\boldsymbol{n}$ such that $\boldsymbol{n} \cdot \boldsymbol{e}_i \neq 0$ by changing frame $\{\boldsymbol{e}_i\}$ to $\{\boldsymbol{e}'_i\}$ by a sequence of $90°$ rotations of the axes. To include vectors for which $\boldsymbol{n} \cdot \boldsymbol{e}_i = 0$ for at least one i, we can use the fact that $\boldsymbol{t}$ is a assumed to be continuous. □

From now on, we will abbreviate notation for $\boldsymbol{t}(\boldsymbol{n}(\boldsymbol{x}), \boldsymbol{x})$, writing $\boldsymbol{t}(\boldsymbol{x})$, where the dependence on the normal field $\boldsymbol{n}$ is kept implicit, and so we have $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{S}(\boldsymbol{x})\boldsymbol{n}$, or in some cases suppress all arguments to write $\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n}$.

The nine components of the stress tensor $\boldsymbol{S}(\boldsymbol{x})$ can be understood as the components of the three traction vectors $\boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x})$ acting across the coordinate planes at the point $\boldsymbol{x}$. In particular, taking components in Equation 4.5, we have

$$\boldsymbol{S}(\boldsymbol{x}) = S_{ij}(\boldsymbol{x}) \boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad \text{with} \quad S_{ij}(\boldsymbol{x}) = t_i(\boldsymbol{e}_j, \boldsymbol{x}),$$

and so

$$\boldsymbol{t}(\boldsymbol{e}_j, \boldsymbol{x}) = t_i(\boldsymbol{e}_j, \boldsymbol{x}) \boldsymbol{e}_i = S_{ij}(\boldsymbol{x}) \boldsymbol{e}_i.$$

4.4 Equilibrium

In this section, we define what is meant by a state of mechanical equilibrium, and use the definition to derive differential equations satisfied.

4.4.1 Preliminaries

Consider a body B_0 in Euclidean space $\mathbb{E}^3$, which is at a state of rest. Suppose the body is then subjected to an external traction and body force fields which cause the body to change shape and come to rest in a possibly different configuration B. The mass density field in the latter configuration is denoted $\rho : B \to \mathbb{R}$, the external traction field per unit area is $\boldsymbol{h} : \partial B \to \mathcal{V}$, and the body force field per unit mass is $\boldsymbol{b} : B \to \mathcal{V}$. All of these fields are assumed not to depend on time.

4.4.2 Necessary conditions

Let $\Omega \subseteq B$ be any open subset of B and let $\boldsymbol{t} : \partial\Omega \to \mathcal{V}$ be the traction field acting on its outer surface, with orientation determined by the outward-point normal field. The resultant force on Ω due to body and surface forces is

$$\boldsymbol{r}(\Omega) = \boldsymbol{r}_b(\Omega) + \boldsymbol{r}_s(\partial\Omega) = \int_\Omega \rho(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} + \int_{\partial\Omega} \boldsymbol{t}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}}, \tag{4.6}$$

and the resultant torque on Ω about the point $\boldsymbol{z} \in \mathbb{E}^3$ is

$$\begin{aligned}\boldsymbol{\tau}(\Omega;\boldsymbol{z}) &= \boldsymbol{\tau}_b(\Omega;\boldsymbol{z}) + \boldsymbol{\tau}_s(\partial\Omega;\boldsymbol{z}) \\ &= \int_\Omega (\boldsymbol{x}-\boldsymbol{z}) \times \rho(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} + \int_{\partial\Omega} (\boldsymbol{x}-\boldsymbol{x}) \times \boldsymbol{t}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}}.\end{aligned} \tag{4.7}$$

At a static equilibrium, we assume that both the resultant force and torque for any Ω vanish, which is encoded in the following axiom.

Static equilibrium

If a body B is in a state of *static mechanical equilibrium*, then the resultant force and resultant torque (taken about any fixed point) which act on any sub-body must vanish. That is, it holds that

$$\begin{aligned}\boldsymbol{r}(\Omega) = \int_\Omega \rho(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} + \int_{\partial\Omega} \boldsymbol{t}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}} = \boldsymbol{0} \\ \boldsymbol{\tau}(\Omega;\boldsymbol{z}) = \int_\Omega (\boldsymbol{x}-\boldsymbol{z}) \times \rho(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} + \int_{\partial\Omega} (\boldsymbol{x}-\boldsymbol{z}) \times \boldsymbol{t}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}} = \boldsymbol{0}\end{aligned} \tag{4.8}$$

for any $\Omega \subseteq B$.

The fact that we have freedom of choice in choosing the point $\boldsymbol{z}$ in the second equation above is a consequence of the first equation: see Exercises for details.

4.4.3 Local equations

We now use the condition of **Static equilibrium** to derive differential equations which are the 'local' form of the equilibrium conditions. Since Proposition 4.2 asserts that $\boldsymbol{t}$ is expressed in terms of the Cauchy stress tensor $\boldsymbol{S}$, it is unsurprising that these equations naturally involve this field.

Proposition 4.3. *If the Cauchy stress field $\boldsymbol{S} : B \to \mathcal{V}^2$ is continuously differentiable, and the density field $\rho : B \to \mathbb{R}$ and body force field $\boldsymbol{b}$ are continuous, then the equilibrium conditions Equation 4.8 are equivalent to*

$$\begin{aligned}(\nabla \cdot \boldsymbol{S})(\boldsymbol{x}) + \rho(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x}) &= \boldsymbol{0} \\ \boldsymbol{S}^T(\boldsymbol{x}) &= \boldsymbol{S}(\boldsymbol{x})\end{aligned} \tag{4.9}$$

for any $\boldsymbol{x} \in B$. In components, these equations are:

$$\begin{aligned}S_{ij,j}(\boldsymbol{x}) + \rho(\boldsymbol{x})b_i(\boldsymbol{x}) &= 0 \\ S_{ij}(\boldsymbol{x}) &= S_{ji}(\boldsymbol{x})\end{aligned}$$

Proof. To establish the first equation in Equation 4.9 assuming that the condition of **Static Equilibrium** holds, we first use the definition of $\boldsymbol{S}$ to write the first equation in Equation 4.8 as

$$\int_{\partial\Omega} \boldsymbol{S}\boldsymbol{n}\,\mathrm{d}A_{\boldsymbol{x}} + \int_{\Omega} \rho\boldsymbol{b}\,\mathrm{d}V_{\boldsymbol{x}} = \boldsymbol{0}.$$

Applying the Tensor Divergence Theorem Proposition 3.7 to the first integral, we obtain

$$\int_{\Omega} \left(\nabla \cdot \boldsymbol{S} + \rho\boldsymbol{b}\right) \mathrm{d}V_{\boldsymbol{x}} = \boldsymbol{0}.$$

Since this equation hold for an arbitrary open set $\Omega \subseteq B$, the Localisation Theorem (Proposition 3.8) allows us to conclude that the integrand vanishes, which is exactly the first result.

To establish the second result, we write the second equation in Equation 4.8 as

$$\int_{\partial\Omega} (\boldsymbol{x} - \boldsymbol{z}) \times (\boldsymbol{S}\boldsymbol{n})\,\mathrm{d}A_{\boldsymbol{x}} + \int_{\Omega} (\boldsymbol{x} - \boldsymbol{z}) \times \rho\boldsymbol{b}\,\mathrm{d}V_{\boldsymbol{x}} = \boldsymbol{0}.$$

Since we have just shown $\rho\boldsymbol{b} = -\nabla \cdot \boldsymbol{S}$, we can substitute, and rewrite

$$\int_{\partial\Omega} (\boldsymbol{x} - \boldsymbol{z}) \times (\boldsymbol{S}\boldsymbol{n})\,\mathrm{d}A_{\boldsymbol{x}} - \int_{\Omega} (\boldsymbol{x} - \boldsymbol{z}) \times \left(\nabla \cdot \boldsymbol{S}\right) \mathrm{d}V_{\boldsymbol{x}} = \boldsymbol{0}. \tag{4.10}$$

Next, note that we may define the tensor $\boldsymbol{R} = R_{il}\boldsymbol{e}_i \otimes \boldsymbol{e}_l$ to be

$$R_{il} = \epsilon_{ijk}(x_j - z_j)S_{kl},$$

which has the property that $\boldsymbol{Rn} = (\boldsymbol{x} - \boldsymbol{z}) \times (\boldsymbol{Sn})$, and hence Equation 4.10 can be written as

$$\int_{\partial\Omega} \boldsymbol{Rn}\, \mathrm{d}A_{\boldsymbol{x}} - \int_{\Omega} (\boldsymbol{x} - \boldsymbol{z}) \times (\nabla \cdot \boldsymbol{S})\, \mathrm{d}V_{\boldsymbol{x}} = \boldsymbol{0}.$$

Applying the Tensor Divergence Theorem to the first of these integrals, we find that

$$\int_{\Omega} \nabla \cdot \boldsymbol{R} - (\boldsymbol{x} - \boldsymbol{z}) \times (\nabla \cdot \boldsymbol{S})\, \mathrm{d}V_{\boldsymbol{x}} = \boldsymbol{0}.$$

Applying the Localisation Theorem in the same way as we did above, it follows that

$$\big(\nabla \cdot \boldsymbol{R}\big)(\boldsymbol{x}) - (\boldsymbol{x} - \boldsymbol{z}) \times (\nabla \cdot \boldsymbol{S})(\boldsymbol{x}) = \boldsymbol{0}$$

for all $\boldsymbol{x} \in B$, which in components becomes

$$(\epsilon_{ijk}(x_j - z_j)S_{kl})_{,l} - \epsilon_{ijk}(x_j - z_j)S_{kl,l} = 0.$$

Using the product rule, we have

$$\begin{aligned}
&(\epsilon_{ijk}(x_j - z_j)S_{kl})_{,l} - \epsilon_{ijk}(x_j - z_j)S_{kl,l} \\
&\qquad = \epsilon_{ijk}x_{j,l}S_{kl} + \epsilon_{ijk}(x_j - z_j)S_{kl,l} - \epsilon_{ijk}(x_j - z_j)S_{kl,l} \\
&\qquad = \epsilon_{ijk}\delta_{jl}S_{kl} \\
&\qquad = \epsilon_{ilk}S_{kl}
\end{aligned}$$

Considering each of these equations for $i = 1, 2, 3$, we see that

$$S_{32} - S_{23} = 0, \quad S_{13} - S_{31} = 0 \quad \text{and} \quad S_{21} - S_{12} = 0.$$

In summary, we have shown that $S_{ij} = S_{ji}$, and hence $\boldsymbol{S}(\boldsymbol{x}) = \boldsymbol{S}^T(\boldsymbol{x})$ for all $\boldsymbol{x} \in B$, which is precisely the second equation in Equation 4.9.

To show that Equation 4.8 hold starting from Equation 4.9, we need simply reverse the arguments above. □

1. The Cauchy stress field is always symmetric, even when a body is not in equilibrium (see later).
2. The local equilibrium equations Equation 4.9 do not completely determine the stress field for a body in equilibrium, since there are 3 PDEs and 3 algebraic equations for 9 unknown components of $\boldsymbol{S}$. This demonstrates that we need additional information to determine the Cauchy stress, and we may address this by prescribing *constitutive equations* characterising the specific material properties of a body.
3. The traction field $\boldsymbol{h}$ acting on ∂B represents the surface force per unit area exterted on B by its environment. Applying Proposition 4.2, we have $\boldsymbol{Sn} = \boldsymbol{h}$ for any $\boldsymbol{x} \in \partial B$, where $\boldsymbol{n}$ is the outward-point normal vector on ∂B. This equation provides a boundary condition for the local equilibrium equations Equation 4.9.

4. In deriving the equilibrium equations, we assumed that the mass density ρ and body force field $\boldsymbol{b}$ are continuous, and that $\boldsymbol{S}$ is continuously differentiable. In practice, establishing such regularity properties is an important part of the subject, and a topic of active research. In this module, we will assume that all fields are sufficiently regular to allow us to exchange integral laws for differential equations and vice versa.
5. The SI unit for stress is the Pascal, denoted Pa, and is equivalent to 1 Newton per square metre. In reasonable real-world cases, stresses can often range up to megapascals (MPa) or gigapascals (GPa).

4.5 Stress concepts

We now study stress in more detail, describing various states of stress which may exist at a point in a body. We also discuss the decomposition of the stress into various components which have important physical meaning.

4.5.1 Simple stress states

If the stress tensor $\boldsymbol{S}$ at a point $\boldsymbol{x} \in B$ takes the form

$$\boldsymbol{S} = -p\boldsymbol{I},$$

where p is a scalar pressure, we say that a *spherical* state of stress exists at $\boldsymbol{x}$. In this stress state, the traction on any surface is parallel to the normal vector $\boldsymbol{n}$:

$$\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} = -p\boldsymbol{n}.$$

The sign choice for p here ensures that positive pressures push inwards on an object.

The stress at a point $\boldsymbol{x} \in B$ is said to be *uniaxial* is there exists a unit vector $\boldsymbol{e}$ and a scalar σ such that

$$\boldsymbol{S} = \sigma \boldsymbol{e} \otimes \boldsymbol{e}.$$

If $\sigma > 0$, we call this state a *pure tension*, and if $\sigma < 0$, a *pure compression*. In this case, the traction on a surface with normal $\boldsymbol{n}$ at $\boldsymbol{x}$ is

$$\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} = (\boldsymbol{e} \cdot \boldsymbol{n})\sigma \boldsymbol{e}.$$

The traction is always parallel to $\boldsymbol{e}$, and vanishes if $\boldsymbol{n}$ is orthogonal to $\boldsymbol{e}$.

If there are a pair of orthogonal unit vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ and a scalar τ such that the stress at a point $\boldsymbol{x} \in B$ takes the form

$$\boldsymbol{S} = \tau\big(\boldsymbol{a} \otimes \boldsymbol{b} + \boldsymbol{b} \otimes \boldsymbol{a}\big),$$

then we say that a state of *pure shear* exists at $\boldsymbol{x}$. For this stress state, the traction on a surface with normal $\boldsymbol{n}$ is

$$\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} = \tau(\boldsymbol{b}\cdot\boldsymbol{n})\boldsymbol{a} + \tau(\boldsymbol{a}\cdot\boldsymbol{n})\boldsymbol{b}.$$

When $\boldsymbol{n} = \boldsymbol{a}$, $\boldsymbol{t} = \tau\boldsymbol{b}$ and when $\boldsymbol{n} = \boldsymbol{b}$, $\boldsymbol{t} = \tau\boldsymbol{a}$.

If at a point $\boldsymbol{x} \in B$ there are a pair of orthogonal unit vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ such that the matrix representation of $\boldsymbol{S}$ with respect to the basis $\boldsymbol{e}_1 = \boldsymbol{a}$, $\boldsymbol{e}_2 = \boldsymbol{b}$ and $\boldsymbol{e}_3 = \boldsymbol{a}\times\boldsymbol{b}$ is

$$= \begin{pmatrix} S_{11} & S_{12} & 0 \\ S_{21} & S_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

we say that a state of *plane stress* exists at $\boldsymbol{x}$.

4.5.2 Principal, normal and shear stresses

The eigenvalues of the Cauchy stress $\boldsymbol{S}$ evaluated at a point $\boldsymbol{x} \in B$ are called the *principal stresses* at $\boldsymbol{x}$. The corresponding eigenvectors are called the *principal stress directions* at $\boldsymbol{x}$. We note that since the stress tensor is symmetric, there exist three real principal stresses, and three orthogonal principal stress directions for each point.

Consider a surface with normal in direction $\boldsymbol{n}$ at $\boldsymbol{x}$. Then the corresponding traction vector can be decomposed into two parts:

$$\begin{aligned} \text{a } \textit{normal traction}\text{:} \qquad & \boldsymbol{t}_n = (\boldsymbol{t}\cdot\boldsymbol{n})\boldsymbol{n}, \\ \text{and a } \textit{shear traction}\text{:} \qquad & \boldsymbol{t}_s = \boldsymbol{t} - (\boldsymbol{t}\cdot\boldsymbol{n})\boldsymbol{n}, \end{aligned}$$

In particular, we have $\boldsymbol{t} = \boldsymbol{t}_n + \boldsymbol{t}_s$, and we call $\sigma_n = |\boldsymbol{t}_n|$ the *normal stress* and $\sigma_s = |\boldsymbol{t}_s|$ the *shear stress* on the surface with normal $\boldsymbol{n}$ at $\boldsymbol{x}$.

4.5.3 Maximum normal and shear stresses

Given a point $\boldsymbol{x}$ in a body of interest, it is often of interest to understand what surfaces passing through $\boldsymbol{x}$ experience the largest normal and shear stresses. In practice, this may be relevant due to knowledge about the level at which a particular material will undergo failure due to stresses of these types. For example, both high levels of tension (a normal stress state) and high levels of shear stress can induce cracking, and the threshold for these two different modes of failure is often different.

Proposition 4.4. *Suppose that the principal stresses σ_i at a point $\boldsymbol{x} \in B$ are distinct and ordered, with*

$$\sigma_1 > \sigma_2 > \sigma_3.$$

Then

- *The maximum normal stress σ_n is $\max_i |\sigma_i|$.*
- *The maximum shear stress σ_s is $\frac{1}{2}|\sigma_1 - \sigma_3|$, and this is achieved for for the two pairs of normals*
$$\boldsymbol{n} = \pm\frac{1}{\sqrt{2}}(\boldsymbol{e}_1 + \boldsymbol{e}_3) \quad \textit{and} \quad \boldsymbol{n} = \pm\frac{1}{\sqrt{2}}(\boldsymbol{e}_1 - \boldsymbol{e}_3).$$

Proof. We prove only the first statement. First note that in terms of the Cauchy stress tensor, we have
$$\sigma_n = |\boldsymbol{n} \cdot (\boldsymbol{S}\boldsymbol{n})|.$$
Since $\boldsymbol{S}$ is symmetric, we can decompose it in terms of the principal stresses σ_i (the eigenvalues of $\boldsymbol{S}$) and the corresponding eigenvectors, so
$$\sigma_n = \left|\boldsymbol{n} \cdot \left(\left(\sum_{i=1}^{3} \sigma_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i\right)\boldsymbol{n}\right)\right| = \left|\sum_{i=1}^{3} \sigma_i (\boldsymbol{e}_i \cdot \boldsymbol{n})^2\right|.$$
Now, applying the triangle inequality and using the fact that $\boldsymbol{n}$ is a unit vector, we have
$$\sigma_n \leq \sum_{i=1}^{3} |\sigma_i|(\boldsymbol{e}_i \cdot \boldsymbol{n})^2 \leq \max_i |\sigma_i|.$$
This shows that the normal stress is at most the value required. To show the maximum is attained, set $\boldsymbol{n} = \boldsymbol{e}_j$, where $|\sigma_j|$ is maximal. □

4.5.4 Spherical and deviatoric stress tensors

At any point $\boldsymbol{x}$, the Cauchy stress tensor can be decomposed into two parts, a *spherical stress tensor*
$$\boldsymbol{S}_S = -p\boldsymbol{I},$$
and a *deviatoric stress tensor*
$$\boldsymbol{S}_D = \boldsymbol{S} + p\boldsymbol{I} = \boldsymbol{S} - \boldsymbol{S}_S,$$
where $p = -\frac{1}{3}\mathrm{tr}\boldsymbol{S}$ is called the pressure. Note that $\boldsymbol{S} = \boldsymbol{S}_S + \boldsymbol{S}_D$.

4.6 An example: atmospheric pressure

To round out this chapter, let's show how we can use the force balance equations in practice by deriving a crude model of atmospheric pressure in the Earth's atmosphere.

A good assumption describing the surface forces in a gas is to assume that the Cauchy stress tensor is spherical everywhere, which means that
$$\boldsymbol{S}(\boldsymbol{x}) = -p(\boldsymbol{x})\boldsymbol{I},$$

where p is the pressure, which is an unknown. This choice is automatically symmetric, so we don't need to concern ourselves further with this condition from the local equations of static equilibrium.

Now, since the atmosphere is a thin layer relative to the curvature of the Earth, we can take the body force per unit mass to be constant. It is sensible to choose a coordinate system so that

$$\boldsymbol{b} = -g\boldsymbol{e}_3.$$

Putting these together, the equilibrium equations become

$$-\nabla_{\boldsymbol{x}} p(\boldsymbol{x}) - g\boldsymbol{e}_3\rho(\boldsymbol{x}) = \mathbf{0}. \tag{4.11}$$

This equation couples the pressure p and the density ρ, but we still need a further constitutive relation between these quantities to get an equation we can solve. One possible option is to treat the air in the atmosphere as an ideal gas. This means that we have the relationship

$$pV = nRT, \tag{4.12}$$

where:

- p is the pressure of the gas,
- V is the volume of the gas,
- n is the number of moles of gas,
- R is the gas constant, and
- T is the absolute temperature.

This is not yet in a form we can use, but rearranging and multiply and dividing by m_0, the mass of the air per mole (the molar mass), we have

$$p = \frac{nm_0}{V}\frac{RT}{m_0} = \rho\frac{RT}{m_0}.$$

Substituting in to Equation 4.11, we can now find a class of solutions to our equation by taking p and ρ to depend only on x_3, the vertical height. In this case, we find that

$$-\frac{\partial\rho}{\partial x_3} - \frac{gm_0}{RT}\rho = 0.$$

This ODE has solutions of the form

$$\rho(x_3) = \rho_0 \exp\left(-\frac{gm_0}{RT}x_3\right),$$

where ρ_0 is the density of the air measured at ground level. We can equally convert this prediction to one concerning the pressure:

$$p(x_3) = \underbrace{\frac{\rho_0 RT}{m_0}}_{=p_0} \exp\left(-\frac{gm_0}{RT}x_3\right),$$

where the p_0 factor in front of the exponential is the pressure observed at ground level. A nice observation we can make from this formula is that the total mass M of our model atmosphere per unit area on the ground is given by the integral

$$M = \int_0^\infty \rho(x_3)\, dx_3 = \rho_0 \frac{RT}{gm_0}.$$

Looking at the expression for the pressure at ground level p_0, we find that

$$p_0 = gM,$$

i.e. the air pressure at ground level is caused by the weight of the air in the column above it.

As we might expect from our experience, the density and pressure decrease with increasing height x_3, and the exponential rate is a reasonably good approximation to the observed behaviour at least in the lowest layer of the atmosphere. On the other hand, there are numerous possible ways we could improve the model. Some potential issues include:

- The fact that air is not an ideal gas;
- The fact that the atmosphere is not at static equilibrium (there is definitely motion occuring!)
- For the same reason, the pressure and density are not independent of the horizontal position; and
- Perhaps worst of all, the temperature of the air reduces with height, whereas our model assumed it to be constant!

There are various standard models used for practical engineering that make use of the basic ideas we discuss here. To find out more, search for the term 'Standard Atmosphere'.

5 Kinematics

Kinematics is the study of describing motion independently of considering mass, forces and stress. We describe the notion of strain in a body which changes shape over time. In the remainder of the module, we will discuss particular relationships between stress and strain which characterise different types of materials.

Aims. By the end of this chapter, you should be able to:

- Define and geometrically interpret the *deformation gradient*, *Cauchy-Green strain tensor*, and the *infinitesimal strain tensor*.
- Use the concept of a *motion* to describe the change in a body.
- Describe the relationship between *material* and *spatial fields*.
- Define the *total time derivative* of material and spatial fields.
- Define the *velocity*, *acceleration*, *vorticity*, *rate of strain* and *spin* fields.
- Explain how volume and surface integrals of fields transform and change due to motion thanks to *change of variables formulae* and the *Reynolds Transport Theorem*.

5.1 Configurations and deformations

At any particular time, a material body occupies an open subset $B \subseteq \mathbb{E}^3$, as discussed in Chapter 4. The identification of material particles with point of B defines what is called a **configuration** of the body.

A **deformation** is a mapping between two configurations, usually a fixed region B called the **reference configuration** and another varying configuration B' called the **deformed configuration**. We will denote points in B by $\boldsymbol{x}$, and points in B' by $\boldsymbol{y}$. If we view B and B' as two configurations of the same material body, we expect that there is a one-to-one correspondence between points in the two configurations.

This leads us naturally to the idea of the deformation map. We will describe the mapping from B to B' by a function $\boldsymbol{\varphi} : B \to B'$ which maps each point $\boldsymbol{x}$ to a point $\boldsymbol{\varphi}(\boldsymbol{x}) \in B'$. The displacement of a material particle from its initial location $\boldsymbol{x}$ to final location $\boldsymbol{\varphi}(\boldsymbol{x})$ is

$$\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{\varphi}(\boldsymbol{x}) - \boldsymbol{x}.$$

We call $\boldsymbol{u} : B \to \mathcal{V}$ the **displacement field** associated to $\boldsymbol{\varphi}$.

We will assume that deformation maps satisfy the following two important conditions:

Admissible deformation

A deformation $\boldsymbol{\varphi} : B \to B'$ is said to be **admissible** if:

- $\boldsymbol{\varphi} : B \to B'$ is one-to-one, and
- $\det \nabla \boldsymbol{\varphi}(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in B$.

The first of these assumptions means that two material particles cannot occupy the same location, and the second ensures that a body cannot be continuously mapped onto its own mirror image. Note that the first assumption is sufficient to ensure that $\boldsymbol{\varphi}$ is a bijection between B and B'.

Throughout, we will assume that all deformations are admissible, and that all are smooth enough to allow us to justify the operations of differentiation we need. Finding (and guaranteeing) that these assumptions hold in particular cases are the topic of active research.

5.2 Measures of strain

Consider the open ball Ω of radius δ centred at $\boldsymbol{x}_0$ within the body we consider. Under the deformation $\boldsymbol{\varphi}$, $\boldsymbol{x}_0$ is mapped to $\boldsymbol{y}_0 = \boldsymbol{\varphi}(\boldsymbol{x}_0)$, and Ω_δ is mapped to a region $\Omega' = \boldsymbol{\varphi}(\Omega)$. Any difference in shaped between Ω' and Ω as $\delta \to 0$ is called *strain* at $\boldsymbol{x}_0$. Strain refers to the local stretching of a body induced by the deformation $\boldsymbol{\varphi}$. The concept of strain plays a central roles in modelling solid materials.

5.2.1 The deformation gradient

One way to quantify strain is by taking the derivative of the deformation:

Deformation gradient

The **deformation gradient** is a second-order tensor field $\boldsymbol{F} : B \to \mathcal{V}^2$ defined to be

$$\boldsymbol{F}(\boldsymbol{x}) = \nabla \boldsymbol{\varphi}(\boldsymbol{x}).$$

The field $\boldsymbol{F}$ naturally provides information on the local behaviour of $\boldsymbol{\varphi}$, since Taylor expanding, we see

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{\varphi}(\boldsymbol{x}_0) + \boldsymbol{F}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + O\big(|\boldsymbol{x} - \boldsymbol{x}_0|^2\big),$$

or equivalently

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{c} + \boldsymbol{F}(\boldsymbol{x}_0)\boldsymbol{x} + O\big(|\boldsymbol{x} - \boldsymbol{x}_0|^2\big) \quad \text{where} \quad \boldsymbol{c} = \boldsymbol{\varphi}(\boldsymbol{x}_0) - \boldsymbol{F}(\boldsymbol{x}_0)\boldsymbol{x}_0.$$

To understand better how $\boldsymbol{F}$ measures strain, we will look at the following simple class of deformations.

Homogeneous deformation

A deformation $\boldsymbol{\varphi}$ is **homogeneous** if its deformation gradient field is constant, i.e. it can be written as

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{c} + \boldsymbol{F}\boldsymbol{x}$$

for a constant vector $\boldsymbol{c} \in \mathcal{V}$ and a constant tensor $\boldsymbol{F} \in \mathcal{V}^2$.

In other words, homogeneous deformations are **affine maps**. We note that by virtue of this affine nature, any line segment in the reference configuration B is mapped onto a corresponding line segment in the deformed configuration B'. We also remark that the admissibility conditions we require of deformations imply that $\det \nabla\boldsymbol{\varphi}(\boldsymbol{x}) > 0$, which ensures that $\det \boldsymbol{F} > 0$, so $\boldsymbol{F}$ is invertible.

5.2.2 Translations, Rotations and Rigid body deformations

We now define some special classes of homogeneous deformation.

Translations

A homogeneous deformation $\boldsymbol{\varphi}$ is called a **translation** if $\boldsymbol{F} = \boldsymbol{I}$, i.e.

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{c} \quad \text{for some fixed } \boldsymbol{c} \in \mathcal{V}.$$

Under the action of a translation, all points in B' are simply moved through a displacement $\boldsymbol{c}$, so there is no change in shape or orientation of the body.

Rotation

A homogeneous deformation $\boldsymbol{\varphi}$ is called a **rotation about** $\boldsymbol{x}'$ if it takes the form

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{x}')$$

for some rotation tensor $\boldsymbol{Q} \in \mathcal{V}^2$.

As for translations, rotations change the orientation of the body, but not its shape. Together, we can combine these two classes to get:

Rigid deformation

A homogeneous deformation φ is called a **rigid deformation** if it can be written as

$$\varphi(\boldsymbol{x}) = \boldsymbol{c} + \boldsymbol{Q}\boldsymbol{x}$$

where $\boldsymbol{c} \in \mathcal{V}$ and $\boldsymbol{Q} \in \mathcal{V}^2$ is a rotation tensor.

Rigid deformations are important since they correspond to moving a body in three dimensions with changing its shape. Similarly, they encode important symmetries of Euclidean space, allowing us to change the origin and orthonormal basis of our coordinate system.

Looking back at our definition of what it means to be a rotation about a point $\boldsymbol{x}'$, we see that it makes sense to define $\boldsymbol{x}'$ as a fixed point of the deformation, since $\varphi(\boldsymbol{x}') = \boldsymbol{x}'$. This gives us the following definition:

Fixed point

A homogeneous deformation has a **fixed point** at $\boldsymbol{x}'$ if

$$\varphi(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{F}(\boldsymbol{x} - \boldsymbol{x}'). \tag{5.1}$$

For a deformation with a fixed point, any point which is not $\boldsymbol{x}'$ in the body is displaced by an amount determined by $\boldsymbol{F}$ and the position relative to the fixed point $\boldsymbol{x}'$. For such a deformation, the body can change both shape and orientation, and we now define some classes of homogeneous deformation with fixed points.

Stretch

A homogeneous deformation φ is called a **stretch about** $\boldsymbol{x}'$ if it takes the form

$$\varphi(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{S}(\boldsymbol{x} - \boldsymbol{x}')$$

for some symmetric positive definite tensor $\boldsymbol{S} \in \mathcal{V}^2$.

When undergoing a stretch, the orientation of the body is not changed, but it is extended by different amounts in different directions with the point $\boldsymbol{x}'$ remaining fixed.

The following result shows that an arbitrary homogeneous deformation can be expressed as a composition of a translation and deformation with a given fixed point.

Proposition 5.1. *Let φ be a homogeneous deformation. Then, given any point $\boldsymbol{x}' \in \mathbb{E}^3$, we can decompose φ as*

$$\varphi = \boldsymbol{t}_1 \circ \boldsymbol{g} = \boldsymbol{g} \circ \boldsymbol{t}_2,$$

*where $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are **translations**, and $\boldsymbol{g}$ is a homogeneous deformation with a **fixed point** at $\boldsymbol{x}'$. In particular, $\boldsymbol{g}$ takes the form*

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{F}(\boldsymbol{x} - \boldsymbol{x}').$$

The following result shows that an arbitrary homogeneous deformation with a fixed point can always be decomposed as a rotation and a stretch about the same fixed point, using the polar decomposition theorem.

Proposition 5.2. *If $\boldsymbol{\varphi}$ is a homogeneous deformation with a fixed point $\boldsymbol{x}'$ and deformation gradient $\boldsymbol{F}$, let $\boldsymbol{F} = \boldsymbol{RU} = \boldsymbol{VR}$ be the right and left polar decompositions of $\boldsymbol{F}$. Then $\boldsymbol{\varphi}$ can be decomposed as*

$$\boldsymbol{\varphi} = \boldsymbol{r} \circ \boldsymbol{s}_1 = \boldsymbol{s}_2 \circ \boldsymbol{r},$$

where $\boldsymbol{r}$ is a rotation about $\boldsymbol{x}'$, and $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are stretches about $\boldsymbol{x}'$. In particular:

$$\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{R}(\boldsymbol{x} - \boldsymbol{x}'), \quad \boldsymbol{s}_1(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{U}(\boldsymbol{x} - \boldsymbol{x}')$$
$$\text{and} \qquad \boldsymbol{s}_2(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{V}(\boldsymbol{x} - \boldsymbol{x}').$$

We can go further in decomposing stretches, by making the following definition.

Extension

A homogeneous deformation $\boldsymbol{\varphi}$ is called an **extension** about $\boldsymbol{x}'$ in the direction of the unit vector $\boldsymbol{e}$ if

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{x}' + \boldsymbol{F}(\boldsymbol{x} - \boldsymbol{x}') \quad \text{with} \quad \boldsymbol{F} = \boldsymbol{I} + (\lambda - 1)\boldsymbol{e} \otimes \boldsymbol{e},$$

for some $\lambda > 0$.

This terminology is based on the observation that $\boldsymbol{F}$ changes the length of any vector parallel to $\boldsymbol{e}$ by a factor of λ, that is

$$\boldsymbol{F}(\alpha \boldsymbol{e}) = \alpha \boldsymbol{e} + (\lambda - 1)\alpha(\boldsymbol{e} \cdot \boldsymbol{e})\boldsymbol{e} = \lambda \alpha \boldsymbol{e}$$

This is a particular case of the stretch deformation introduced earlier. The following result shows that the stretches appearing in Proposition 5.2 can be expressed as the composition of three extensions defined by the eigenvalues and eigenvectors of $\boldsymbol{U}$ and $\boldsymbol{V}$.

Proposition 5.3. *Let $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ be the stretches defined in Proposition 5.2, and let $\{\lambda_i, \boldsymbol{u}_i\}$ and $\{\lambda_i, \boldsymbol{v}_i\}$ be the eigenpairs associated with the tensors $\boldsymbol{U}$ and $\boldsymbol{V}$ respectively. Then*

$$\boldsymbol{s}_1 = \boldsymbol{f}_1 \circ \boldsymbol{f}_2 \circ \boldsymbol{f}_3 \quad \text{and} \quad \boldsymbol{s}_2 = \boldsymbol{h}_1 \circ \boldsymbol{h}_2 \circ \boldsymbol{h}_3,$$

where $\boldsymbol{f}_i$ is the extension about $\boldsymbol{x}'$ by λ_i in the direction $\boldsymbol{u}_i$, and $\boldsymbol{h}_i$ is the extension about $\boldsymbol{x}'$ by λ_i in the direction $\boldsymbol{v}_i$.

We make a few notes about this result:

- The tensors $\boldsymbol{U}$ and $\boldsymbol{V}$ which appear in the right and left polar decompositions of $\boldsymbol{F}$ have the same eigenvalues, but (in general) have different eigenvectors. It follows that $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ give rise to extensions by the same amounts, but in different directions.
- We call the eigenvalues λ_i the **principal stretches** associated with the deformation gradient $\boldsymbol{F}$, and the eigenvectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are respectively referred to as the right and left **principal directions**. Similarly, we call $\boldsymbol{U}$ and $\boldsymbol{V}$ the right and left **stretch tensors**.

In summary, we have shown that a homogeneous deformation can be decomposed in various ways, for example as a translation, then a rotation, then a series of extensions, or as a translation, then a series of extensions, and then a rotation. These different decompositions allow us to think about how a body might be affected by such a sequence of operations, and so devise physically-appropriate models, especially since a Taylor expansion of the deformation $\boldsymbol{\varphi}$ allows us to treat deformations close to a given point as approximately homogeneous.

5.2.3 The Cauchy–Green strain tensor

A second measure of strain is:

Cauchy-Green strain

The **(right) Cauchy-Green strain tensor** field is $\boldsymbol{C} : B \to \mathcal{V}^2$, defined by

$$\boldsymbol{C} = \boldsymbol{F}^T \boldsymbol{F}.$$

Notice that $\boldsymbol{C}$ is symmetric and positive definite at every point in B. While $\boldsymbol{F}$ contains information on both rotation and stretching, $\boldsymbol{C}$ includes information about stretches only. Since $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$,

$$\boldsymbol{C} = \boldsymbol{F}^T \boldsymbol{F} = \boldsymbol{U}^2,$$

we see the definition of $\boldsymbol{C}$ does not contain $\boldsymbol{R}$.

- Note that $\boldsymbol{U}$ is also independent of $\boldsymbol{R}$, and contains information on stretches only, but $\boldsymbol{C}$ is easier to compute than the square root $\boldsymbol{U} = \sqrt{\boldsymbol{F}^T \boldsymbol{F}} = \sqrt{\boldsymbol{C}}$.
- Recall from the Spectral Decomposition Theorem that

$$\boldsymbol{U} = \sum_{i=1}^{3} \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i,$$

where $\lambda_i > 0$ and $\boldsymbol{e}_i$ are orthonormal eigenvectors of $\boldsymbol{U}$. As $\boldsymbol{C} = \boldsymbol{U}^2$, we see that

$$\boldsymbol{C} = \sum_{i=1}^{3} \lambda_i^2 \boldsymbol{e}_i \otimes \boldsymbol{e}_i,$$

so the eigenvalues of $\boldsymbol{C}$ are the squares of the principal stretches.

- Another measure of strain is the **left Cauchy-Green strain tensor**, $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T = \boldsymbol{V}^2$.

We now consider the interpretation of $\boldsymbol{C}$. Consider an arbitrary point $\boldsymbol{x}_0$ in the reference configuration, and let Ω be the open ball of radius $\alpha > 0$ centred at $\boldsymbol{x}_0$. Consider the points

$$\boldsymbol{x}' = \boldsymbol{x}_0 + \alpha \boldsymbol{e}' \quad \text{and} \quad \boldsymbol{x}'' = \boldsymbol{x}_0 + \alpha \boldsymbol{e}'',$$

where $\boldsymbol{e}'$ and $\boldsymbol{e}''$ are unit vectors. Let

$$\boldsymbol{y}_0 = \boldsymbol{\varphi}(\boldsymbol{x}_0), \quad \boldsymbol{y}' = \boldsymbol{\varphi}(\boldsymbol{x}') \quad \text{and} \quad \boldsymbol{y}'' = \boldsymbol{\varphi}(\boldsymbol{x}'')$$

be the corresponding points in the deformed configuration. Let $\phi \in [0, \pi]$ be the angle between the displacement vectors $\boldsymbol{y}' - \boldsymbol{y}_0$ and $\boldsymbol{y}'' - \boldsymbol{y}_0$ in the deformed configuration. Then we have the following result:

Proposition 5.4. *For any point $\boldsymbol{x}_0 \in B$ and unit vectors $\boldsymbol{e}'$ and $\boldsymbol{e}''$, define*

$$\lambda(\boldsymbol{e}') = \sqrt{\boldsymbol{e}' \cdot \boldsymbol{C}\boldsymbol{e}'} > 0 \quad \textit{and} \quad \theta(\boldsymbol{e}', \boldsymbol{e}'') = \arccos\left(\frac{\boldsymbol{e}' \cdot \boldsymbol{C}\boldsymbol{e}''}{\lambda(\boldsymbol{e}')\lambda(\boldsymbol{e}'')}\right) \in [0, \pi].$$

Then, as $\alpha \to 0$, we have that

$$\frac{|\boldsymbol{y}' - \boldsymbol{y}_0|}{|\boldsymbol{x}' - \boldsymbol{x}_0|} \to \lambda(\boldsymbol{e}'), \quad \frac{|\boldsymbol{y}'' - \boldsymbol{y}_0|}{|\boldsymbol{x}'' - \boldsymbol{x}_0|} \to \lambda(\boldsymbol{e}'')$$

and

$$\phi \to \theta(\boldsymbol{e}', \boldsymbol{e}'').$$

Proof. Note that $\boldsymbol{y}' - \boldsymbol{y}_0 = \boldsymbol{\varphi}(\boldsymbol{x}_0 + \alpha \boldsymbol{e}') - \boldsymbol{\varphi}(\boldsymbol{x}_0)$, so Taylor expanding, we have

$$\boldsymbol{y}' - \boldsymbol{y}_0 = \alpha \boldsymbol{F}(\boldsymbol{x}_0)\boldsymbol{e}' + O(\alpha^2).$$

It follows that

$$|\boldsymbol{y}' - \boldsymbol{y}_0|^2 = \alpha^2 \Big(\boldsymbol{F}(\boldsymbol{x}_0)\boldsymbol{e}'\Big) \cdot \Big(\boldsymbol{F}(\boldsymbol{x}_0)\boldsymbol{e}'\Big) + O(\alpha^3) = \alpha^2 \boldsymbol{e}' \cdot \boldsymbol{C}(\boldsymbol{x}_0)\boldsymbol{e}' + O(\alpha^3).$$

Dividing through by $\alpha^2 = |\boldsymbol{x}' - \boldsymbol{x}_0|^2$, we have

$$\frac{|\boldsymbol{y}' - \boldsymbol{y}_0|^2}{|\boldsymbol{x}' - \boldsymbol{x}_0|^2} = \boldsymbol{e}' \cdot \boldsymbol{C}(\boldsymbol{x}_0)\boldsymbol{e}' + O(\alpha),$$

which, after taking a square root, proves the first result. The second result follows by a similar argument.

To establish the final result, we note that

$$\cos\phi = \frac{(\boldsymbol{y}' - \boldsymbol{y}_0)\cdot(\boldsymbol{y}'' - \boldsymbol{y}_0)}{|\boldsymbol{y}' - \boldsymbol{y}_0||\boldsymbol{y}'' - \boldsymbol{y}_0|}.$$

Using the facts that $\boldsymbol{y}' - \boldsymbol{y}_0 = \alpha \boldsymbol{F}(\boldsymbol{x})\boldsymbol{e}' + O(\alpha^2)$ and $\boldsymbol{y}'' - \boldsymbol{y}_0 = \alpha \boldsymbol{F}(\boldsymbol{x})\boldsymbol{e}'' + O(\alpha^2)$, and using the first part of the result, we have

$$\cos\phi = \boldsymbol{e}' \cdot \boldsymbol{C}(\boldsymbol{x})\boldsymbol{e}'' + O(\alpha).$$

Letting $\alpha \to 0$, we conclude. □

Remarks:

- This result states that $\lambda(\boldsymbol{e})$ approximates the ratio of lengths between deformed points and reference points, where the reference points are separated by a short displacement in the direction $\boldsymbol{e}$. This is the reason for naming $\lambda(\boldsymbol{e})$ the *stretch* in the direction $\boldsymbol{e}$ at the reference point $\boldsymbol{x}$.
- At any fixed reference point $\boldsymbol{x}$, the extreme values of $\lambda(\boldsymbol{e})$ occur when $\boldsymbol{e}$ is an eigenvector of $\boldsymbol{C}$ (i.e. a right principal direction). As a result the extreme values of $\lambda(\boldsymbol{e})$ are given by the maximum and minimum principal stretches.
- $\theta(\boldsymbol{e}', \boldsymbol{e}'')$ approximates the angle between points in the deformed configuration which lie at short distances from $\boldsymbol{x}$ along the directions $\boldsymbol{e}'$ and $\boldsymbol{e}''$. If $\Theta(\boldsymbol{e}', \boldsymbol{e}'') = \arccos(\boldsymbol{e}' \cdot \boldsymbol{e}'')$ is the angle between these directions in the reference configuration, we define
$$\gamma(\boldsymbol{e}', \boldsymbol{e}'') = \Theta(\boldsymbol{e}', \boldsymbol{e}'') - \theta(\boldsymbol{e}', \boldsymbol{e}''),$$
which we call the **shear** between the directions $\boldsymbol{e}'$ and $\boldsymbol{e}''$ at the point $\boldsymbol{x}$. The shear measures the change in angle between directions due to deformation.

The following result shows that $\boldsymbol{C}$ explicitly characterises the stretch and shear caused by a deformation when considered in a particular coordinate frame.

Proposition 5.5. *Let C_{ij} be the components of $\boldsymbol{C}$ with respect to an arbitrary coordinate frame $\{\boldsymbol{e}_i\}$. Then, for any point $\boldsymbol{x} \in B$, we have*

$$C_{ii} = \lambda(\boldsymbol{e}_i)^2 \quad \textit{and} \quad C_{ij} = \lambda(\boldsymbol{e}_i)\lambda(\boldsymbol{e}_j)\sin\big(\gamma(\boldsymbol{e}_i, \boldsymbol{e}_j)\big),$$

where no summation is implied. The diagonal components of $\boldsymbol{C}$ are therefore the squares of the stretches along coordinate directions, and the off-diagonal components capture the shear between corresponding pairs of coordinate directions.

To finish our discussion of the Cauchy-Green strain, we note that if $\boldsymbol{\varphi}$ is a rigid deformation, we have $\boldsymbol{F} = \boldsymbol{Q}$ where $\boldsymbol{Q}$ is a rotation tensor, and hence

$$\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}.$$

It follows that rigid deformations produce no strain measured by $\boldsymbol{C}$. Looking carefully at the proof of Proposition 5.4, we see that the relative position and orientation of any three points in B are unchanged by a rigid deformation (hence the name). It is possible to show that a deformation is rigid if and only if $\boldsymbol{C}(\boldsymbol{x}) = \boldsymbol{I}$ for all $\boldsymbol{x} \in B$.

5.2.4 The infinitesimal strain tensor

Consider a deformation $\boldsymbol{\varphi} : B \to B'$ with displacement field $\boldsymbol{u}$ and displacement gradient $\nabla\boldsymbol{u}$. We define:

Infinitesimal strain

The **infinitesimal strain tensor** field associated with a deformation is $\boldsymbol{\varepsilon} : B \to \mathcal{V}^2$ defined by

$$\boldsymbol{\varepsilon} = \operatorname{sym}(\nabla\boldsymbol{u}) = \tfrac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T).$$

Note that, by definition, $\boldsymbol{\varepsilon}$ is symmetric. $\boldsymbol{\varepsilon}$ can related to the deformation gradient $\boldsymbol{F}$ and the Cauchy-Green tensor $\boldsymbol{C}$. From the definition of $\boldsymbol{u}$, we see that $\nabla\boldsymbol{u} = \boldsymbol{F} - \boldsymbol{I}$, and hence $\boldsymbol{\varepsilon} = \operatorname{sym}(\boldsymbol{F} - \boldsymbol{I})$. Since $\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F}$, we see that

$$\boldsymbol{\varepsilon} = \tfrac{1}{2}(\boldsymbol{C} - \boldsymbol{I}) - \tfrac{1}{2}\nabla\boldsymbol{u}^T\nabla\boldsymbol{u}.$$

The tensor $\boldsymbol{\varepsilon}$ is particularly useful in the case of small deformations. A deformation $\boldsymbol{\varphi}$ induces **small strains** if the displacement gradient is small, i.e. $|\nabla\boldsymbol{u}(\boldsymbol{x})| \ll 1$ for all points $\boldsymbol{x} \in B$. In this case, we see that that

$$\boldsymbol{\varepsilon} = \tfrac{1}{2}(\boldsymbol{C} - \boldsymbol{I}) + O\left(|\nabla\boldsymbol{u}|^2\right).$$

If we drop the latter terms, then $\boldsymbol{\varepsilon}$ is equivalent to $\boldsymbol{C}$ up to a multiplicative factor and offset, and $\boldsymbol{\varepsilon}$ will arise when studying linearised models of stress in elastic solids.

- Note that if $\nabla\boldsymbol{u} = \boldsymbol{O}$ everywhere in B, we have $\boldsymbol{F} = \boldsymbol{I}$, and hence $\boldsymbol{\varphi}$ is a translation. It follows that $\boldsymbol{\varphi}$ is a small deformation when it deviates only slightly from a pure *translation* (not a rigid body motion!).
- For small deformations, the tensor $\boldsymbol{\varepsilon}$ contains similar information to $\boldsymbol{C}$. However, we note that $\boldsymbol{\varepsilon}$ depends linearly on $\boldsymbol{u}$ (and hence on $\boldsymbol{y}$), whereas $\boldsymbol{C}$ depends non-linearly on $\boldsymbol{u}$.

To interpret $\boldsymbol{\varepsilon}$, we have the following result.

Proposition 5.6. *Let ε_{ij} be the components of $\boldsymbol{\varepsilon}$ taken with respect to a Cartesian frame $\{\boldsymbol{e}_i\}$. Then, with no sum over repeated indices implies, we have*

$$\varepsilon_{ii} = \lambda(\boldsymbol{e}_i) - 1 + O(|\nabla \boldsymbol{u}|^2) \quad \textit{and} \quad \varepsilon_{ij} = \tfrac{1}{2}\sin\gamma(\boldsymbol{e}_i, \boldsymbol{e}_j) + O(|\nabla \boldsymbol{u}|^2),$$

where $\lambda(\boldsymbol{e}_i)$ is the stretch in the direction $\boldsymbol{e}_i$, and $\gamma(\boldsymbol{e}_i, \boldsymbol{e}_j)$ is the shear between the directions $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$.

The result above can be proven using Proposition 5.5.

- As shown in Proposition 5.4, the stretch $\lambda(\boldsymbol{e}_i)$ is the ratio of the distance between deformed points which were close to one another in the $\boldsymbol{e}_i$ direction in the reference configuration. This means that $\lambda(\boldsymbol{e}_i) - 1$ is approximately the relative change in the length of an infinitesimal line segment pointing in the $\boldsymbol{e}_i$ direction in the reference configuration.

- When the shear angle $\gamma(\boldsymbol{e}_i, \boldsymbol{e}_j)$ is small, we have that

  $$\varepsilon_{ij} \approx \tfrac{1}{2}\sin\gamma(\boldsymbol{e}_i, \boldsymbol{e}_j) \approx \tfrac{1}{2}\gamma(\boldsymbol{e}_i, \boldsymbol{e}_j),$$

  so for small deformations, the off-diagonal components of $\boldsymbol{\varepsilon}$ are approximately half the shear angle between to line segments which pointed in the directions $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ in the reference configuration.

5.2.5 Infinitesimal rigid deformations

A homogeneous deformation $\boldsymbol{\varphi}$ is called *infinitesimally rigid* if the associated displacement field $\boldsymbol{u}$ takes the form

$$\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{c} + \boldsymbol{W}\boldsymbol{x}$$

for some vector $\boldsymbol{c} \in \mathcal{V}$ and a skew-symmetric tensor $\boldsymbol{W} \in \mathcal{V}^2$. Recalling the result of Proposition 2.7, we may write

$$\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{c} + \boldsymbol{w} \times \boldsymbol{x}$$

where $\boldsymbol{w}$ is the axial vector of $\boldsymbol{W}$. For an infinitesimally rigid deformation, the displacement gradient is $\nabla \boldsymbol{u} = \boldsymbol{W}$, and

$$\boldsymbol{\varepsilon} = \operatorname{sym}(\nabla \boldsymbol{u}) = \tfrac{1}{2}(\boldsymbol{W} + \boldsymbol{W}^T) = \boldsymbol{O}.$$

As a result, we see that infinitesimal rigid deformations produce no strain measured by $\boldsymbol{\varepsilon}$.

5.3 Motion

The continuous deformation of a body over time is called a *motion*. The motion of a body with reference configuration B can be described by a continuous map $\boldsymbol{\varphi} : B \times [0, T] \to \mathbb{E}^3$, where for any fixed time t, the function $\boldsymbol{\varphi}(\cdot, t) : B \to \mathbb{E}^3$ is a deformation of B. At time t, the body undergoing this motion is in the configuration $B_t = \boldsymbol{\varphi}(B, t)$, and we call B_t the *current* or *deformed* configuration at time t.

If $\boldsymbol{\varphi}(\cdot, 0)$ is the identity map, i.e. $\boldsymbol{\varphi}(\boldsymbol{x}, 0) = \boldsymbol{x}$ for all $\boldsymbol{x} \in B$, we have that $B_0 = B$. Assuming continuity ensures that the body cannot break apart into pieces, and cannot move by an instantaneous jump. We will assume that $\boldsymbol{\varphi}(\cdot, t)$ is admissible in the sense described above, and hence there exists an inverse deformation $\boldsymbol{\psi}(\cdot, t) : B_t \to B$ with $\boldsymbol{\psi}(\cdot, t) = \boldsymbol{\varphi}^{-1}(\cdot, t)$, i.e. $\boldsymbol{\psi}(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) = \boldsymbol{x}$.

Throughout the remainder of this module, we will assume that all motions and their inverses are smooth in the sense that partial derivatives of all orders used exist and are continuous.

5.3.1 Material and spatial fields

In our study of the motion of continuum bodies, we wish to study fields defined on the current configuration B_t, whose points we label $\boldsymbol{y}$. Since we assume such points can be expressed in terms of the deformation mapping $\boldsymbol{y} = \boldsymbol{\varphi}(\boldsymbol{x}, t)$, any function defined on B_t can also be expressed as a function of B through 'pre-composition' with the motion $\boldsymbol{\varphi} : B \times [0, T] \to \mathbb{E}^3$, and any function defined on the reference configuration B can be expressed as a function defined on B_t through precomposition with the inverse mapping $\boldsymbol{\psi}(\cdot, t) : B_t \to \mathbb{E}^3$.

By a *material field*, we mean a field which is expressed in terms of the points $\boldsymbol{x} \in B$; for example $\Omega = \Omega(\boldsymbol{x}, t)$ for $\boldsymbol{x} \in B$. By a *spatial field*, we mean a field expressed in terms of the point $\boldsymbol{y} \in B_t$, for example $\Gamma = \Gamma(\boldsymbol{y}, t)$ for $\boldsymbol{y} \in B_t$.

We can freely map back and forth between these descriptions. To any material field $\Omega(\boldsymbol{x}, t)$, we can associate a spatial field

$$\Omega_s(\boldsymbol{y}, t) = \Omega(\boldsymbol{\psi}(\boldsymbol{y}, t), t).$$

We call Ω_s the *spatial description* of the material field Ω. Likewise, given a spatial field $\Gamma(\boldsymbol{y}, t)$, we can define an associated material field

$$\Gamma_m(\boldsymbol{x}, t) = \Gamma(\boldsymbol{\varphi}(\boldsymbol{x}, t), t),$$

and we refer to Γ_m as the material description of the spatial field Γ.

5.3.2 Derivatives

Given that we have two different descriptions of the fields of interest, we need to distinguish differentiation with respect to the material coordinates $\boldsymbol{x} = (x_1, x_2, x_3)$ labelling points in B, and the spatial coordinates $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3)$, labelling points in B_t.

To keep these derivatives separate, we use $\nabla_{\boldsymbol{x}}$ to refer to the gradient, divergence and curl taken in the material coordinates, and $\nabla_{\boldsymbol{y}}$ to refer to the gradient, divergence and curl taken in the spatial coordinates.

We will also frequently compute the rate of change of a field with respect to time. The *total time derivative* or *convective derivative* of a field is the rate of change of a field as measured by a stationary observer who is tracking the motion of each particle in the body. Since we keep the reference configuration B fixed, it follows that the material coordinates $\boldsymbol{x} \in B$ of each particle are fixed. On the other hand, as the current configuration varies in time, the spatial coordinates $\boldsymbol{y} \in B_t$ will change with time.

We note that any material field $\Omega(\boldsymbol{x}, t)$ satisfies

$$\frac{\mathrm{D}}{\mathrm{D}t}\Omega(\boldsymbol{x}, t) = \frac{\partial}{\partial t}\Omega(\boldsymbol{x}, t),$$

where as for a spatial field $\Gamma(\boldsymbol{y}, t)$, we have

$$\frac{\mathrm{D}}{\mathrm{D}t}\Gamma(\boldsymbol{y}, t) = \frac{\partial}{\partial t}\Gamma\big(\boldsymbol{\varphi}(\boldsymbol{x}, t), t\big)\Big|_{\boldsymbol{x}=\boldsymbol{\psi}(\boldsymbol{y},t)} = \frac{\partial}{\partial t}\Gamma_m(\boldsymbol{x}, t)\Big|_{\boldsymbol{x}=\boldsymbol{\psi}(\boldsymbol{y},t)}.$$

We can write this as

$$\frac{\mathrm{D}}{\mathrm{D}t}\Gamma = \left[\frac{\mathrm{D}}{\mathrm{D}t}\Gamma_m\right]_s.$$

Remarks:

- The total time derivative of a field is sometimes called the *material*, *substantial*, or *convective* derivative.
- In general, we note that $\frac{\mathrm{D}}{\mathrm{D}t}\Gamma(\boldsymbol{y}, t) \neq \frac{\partial}{\partial t}\Gamma(\boldsymbol{y}, t)$, i.e. the total time derivative and partial time derivative of a spatial field are not the same. This distinction arises since the partial derivative treats spatial coordinates as fixed, but the time derivative takes into account the motion of the coordinates themselves as they follow a given particle.

5.3.3 Velocity and acceleration fields

Consider a motion $\boldsymbol{\varphi}$, and consider a particle $\boldsymbol{x} \in B$. At later times, the particle has position $\boldsymbol{y} = \boldsymbol{\varphi}(\boldsymbol{x}, t) \in B_t$. The velocity of this particle and its acceleration both labelled by its material coordinates, are

$$\frac{\partial}{\partial t}\boldsymbol{\varphi}(\boldsymbol{x}, t) \quad \text{and} \quad \frac{\partial^2}{\partial t^2}\boldsymbol{\varphi}(\boldsymbol{x}, t).$$

By definition, these are both material fields, but we will also wish to consider the spatial description of these fields. Setting $\boldsymbol{v}$ and $\boldsymbol{a}$ to be the spatial forms of these fields, we have

$$\boldsymbol{v}(\boldsymbol{y}, t) = \frac{\partial}{\partial t}\boldsymbol{\varphi}(\boldsymbol{x}, t)\bigg|_{\boldsymbol{x}=\boldsymbol{\psi}(\boldsymbol{y},t)} \quad \text{and} \quad \boldsymbol{a}(\boldsymbol{y}, t) = \frac{\partial^2}{\partial t^2}\boldsymbol{\varphi}(\boldsymbol{x}, t)\bigg|_{\boldsymbol{x}=\boldsymbol{\psi}(\boldsymbol{y},t)}.$$

These two fields correspond to the velocity and acceleration of a particle whose spatial position is $\boldsymbol{y}$ at time t. The following result uses these functions to provide a convenient formula for the total time derivative of a spatial field.

Proposition 5.7. *If $\boldsymbol{\varphi} : B \times [0, T] \to \mathbb{E}^3$ is a motion of a continuum body with associated spatial velocity field $\boldsymbol{v}$, and $\phi = \phi(\boldsymbol{y}, t)$ and $\boldsymbol{w} = \boldsymbol{w}(\boldsymbol{x}, t)$ are spatial fields, then*

$$\frac{\mathrm{D}}{\mathrm{D}t}\phi = \frac{\partial}{\partial t}\phi + \nabla_{\boldsymbol{y}}\phi \cdot \boldsymbol{v} \quad \textit{and} \quad \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{w} = \frac{\partial}{\partial t}\boldsymbol{w} + \big(\nabla_{\boldsymbol{y}}\boldsymbol{w}\big)\boldsymbol{v}.$$

Proof. Taking components in a Cartesian frame $\{\boldsymbol{e}_i\}$, we have

$$\nabla_{\boldsymbol{y}}\phi = \frac{\partial \phi}{\partial y_i}\boldsymbol{e}_i \quad \text{and} \quad \nabla_{\boldsymbol{y}}\boldsymbol{w} = \frac{\partial w_i}{\partial y_j}\boldsymbol{e}_i \otimes \boldsymbol{e}_j.$$

By the definition of the total time derivative, we have

$$\begin{aligned}
\frac{\mathrm{D}}{\mathrm{D}t}\phi(\boldsymbol{y}, t)\bigg|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} &= \frac{\partial}{\partial t}\phi(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) \\
&= \frac{\partial}{\partial t}\phi(\boldsymbol{y}, t)\bigg|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} + \frac{\partial}{\partial y_i}\phi(\boldsymbol{y}, t)\bigg|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} \frac{\partial}{\partial t}\varphi_i(\boldsymbol{y}, t)\bigg|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} \\
&= \frac{\partial}{\partial t}\phi(\boldsymbol{y}, t)\bigg|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} + \frac{\partial}{\partial y_i}\phi(\boldsymbol{y}, t)\bigg|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} v_i(\boldsymbol{y}, t)\bigg|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)}.
\end{aligned}$$

Expressing this result in terms of the spatial coordinates $\boldsymbol{y}$, we have the result. Writing $\boldsymbol{w} = w_i\boldsymbol{e}_i$ and applying the same argument to each coordinate, we obtain the result for $\boldsymbol{w}$. □

Remarks:

- This result shows that is we know the spatial velocity field $\boldsymbol{v}$, then we can compute the total time derivative of a spatial field ϕ without explicit knowledge of $\boldsymbol{\varphi}$ or $\boldsymbol{\psi}$.
- The spatial acceleration field satisfies $\boldsymbol{a} = \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v}$, and so applying Proposition 5.7, we have
$$\boldsymbol{a} = \frac{\partial \boldsymbol{v}}{\partial t} + (\nabla_{\boldsymbol{y}}\boldsymbol{v})\boldsymbol{v}.$$
The spatial acceleration is therefore a nonlinear function of the spatial velocity and its derivatives.
- Many texts write $\boldsymbol{v} \cdot \nabla_{\boldsymbol{y}}\boldsymbol{w}$ in place of $(\nabla_{\boldsymbol{y}}\boldsymbol{w})\boldsymbol{v}$, but we prefer our notation here since $\nabla_{\boldsymbol{y}}\boldsymbol{w}$ is a second-order tensor, and so our notation remains consistent with the usual application of a second-order tensor to a vector.

5.4 Rate of strain and spin

Sometimes, we wish to quantify the rate at which points in a body of interest are distorting. Any measure of the rate of change of shape is called a *rate of strain*, while any change of orientation is called a *rate of rotation* or *spin*. It is important to realise that rates of strain and rotation are independent of the reference configuration B (in contrast to measures of strain), since these measures only compare the change of the body over short times. As such, these measures play an important role in the study of fluids.

If $\boldsymbol{\varphi} : B \times [0,T] \to \mathbb{E}^3$ is a motion with spatial velocity field $\boldsymbol{v}$, then the *rate of strain* tensor is
$$\boldsymbol{L} = \operatorname{sym}(\nabla_{\boldsymbol{y}}\boldsymbol{v}) = \tfrac{1}{2}\big(\nabla_{\boldsymbol{y}}\boldsymbol{v} + (\nabla_{\boldsymbol{y}}\boldsymbol{v})^T\big).$$
This is a spatial field $\boldsymbol{L}(\cdot,t) : B_t \to \mathcal{V}^2$, and is symmetric for every point $\boldsymbol{y} \in B_t$ and each time t.

We also define the *spin tensor*
$$\boldsymbol{W} = \operatorname{skew}(\nabla_{\boldsymbol{y}}\boldsymbol{v}) = \tfrac{1}{2}\big(\nabla_{\boldsymbol{y}}\boldsymbol{v} - (\nabla_{\boldsymbol{y}}\boldsymbol{v})^T\big).$$
This too is a spatial field $\boldsymbol{W}(\cdot,t) : B_t \to \mathcal{V}^2$, which is skew for all points $\boldsymbol{y} \in B_t$ and times t.

We can use arguments analogous to those used to analyse the infinitesimal strain tensor $\boldsymbol{\varepsilon}$ to study $\boldsymbol{L}$, which leads us to conclude that the diagonal components of $\boldsymbol{L}$ quantify the instantaneous rate of stretch at a spatial point in the coordinate axis directions. Similarly, the off-diagonal components of $\boldsymbol{L}$ quantify the instantaneous rate of shearing. In particular, if we consider a motion $\widehat{\boldsymbol{\varphi}} : B_t \to B_{t+s}$ where $s \ll 1$, then $\boldsymbol{L}$ is the rate of change of both the corresponding infinitesimal strain tensor $\widehat{\boldsymbol{\varepsilon}}$ and of the right stretch tensor $\widehat{\boldsymbol{U}}$.

The skew-symmetric tensor $\boldsymbol{W}$ quantifies the instantaneous rate of rigid rotation. In particular, considering the same motion $\widehat{\boldsymbol{\varphi}} : B_t \to B_{t+s}$, then $\boldsymbol{W}$ is the rate of change of

$\widehat{\boldsymbol{R}}$ which is the rotation tensor in the polar decomposition of the deformation gradient $\nabla_{\boldsymbol{x}}\widehat{\boldsymbol{\varphi}}$.

5.4.1 Vorticity

The vorticity of a motion is the spatial vector field

$$\boldsymbol{w} = \nabla_{\boldsymbol{y}} \times \boldsymbol{v}.$$

We note that in view of result Proposition 2.7, $\boldsymbol{w}$ is the axial vector associated with the (multiple of the) spin tensor $2\boldsymbol{W}$. The vorticity therefore measures the rotation or spin at a given spatial point.

5.4.2 Rigid motion

A motion $\boldsymbol{\varphi}$ is **rigid** if it is a rigid deformation at each time t, i.e.

$$\boldsymbol{\varphi}(\boldsymbol{x}, t) = \boldsymbol{c}(t) + \boldsymbol{Q}(t)\boldsymbol{x}$$

for some time-dependent vector $\boldsymbol{c}(t)$ and rotation tensor $\boldsymbol{Q}(t)$. For such motions it can be shown that the spatial velocity field takes the form

$$\boldsymbol{v}(\boldsymbol{y}, t) = \boldsymbol{\omega} \times (\boldsymbol{y} - \boldsymbol{c}) + \frac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t},$$

where $\boldsymbol{\omega}(t)$ is a time-dependent vector called the **spatial angular velocity** of the motion. We can use Proposition 2.7 to write this as

$$\boldsymbol{v}(\boldsymbol{y}, t) = \boldsymbol{\Omega}(t)\big(\boldsymbol{y} - \boldsymbol{c}(t)\big) + \frac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t}(t),$$

where $\boldsymbol{\Omega}(t)$ is the second-order tensor with axial vector $\boldsymbol{\omega}(t)$. It follows that

$$\nabla_{\boldsymbol{y}}\boldsymbol{v}(\boldsymbol{y}, t) = \boldsymbol{\Omega}(t),$$

from which we deduce $\boldsymbol{L} = \boldsymbol{O}$, $\boldsymbol{W} = \boldsymbol{\Omega}(t)$, and $\boldsymbol{w} = 2\boldsymbol{\omega}(t)$. It follows that rigid motions produce no rate of strain measured by $\boldsymbol{L}$, but do produce spin as measured by the spin $\boldsymbol{W}$ or vorticity $\boldsymbol{w}$.

5.5 Change of variables

As we saw in the previous chapter, it is useful to express various physical quantities as integrals. However, we now have two sets of coordinates to integrate with respect to: the material coordinates $\boldsymbol{x} \in B$, and the spatial coordinates $\boldsymbol{y} \in B_t$. For fixed t, the mappings $\boldsymbol{\varphi}(\cdot, t)$ and $\boldsymbol{\psi}(\cdot, t)$ act as changes of variable between these two sets of coordinates.

5.5.1 Transformation of volume integrals

As you will have seen in multivariable calculus modules, we can express the volume element in the spatial coordinates, $dV_{\boldsymbol{y}}$, in terms of material coordinates $\boldsymbol{x}$ by considering the volume change transformation of an infinitesimal cube at a fixed material point. This volume change factor is given by the *Jacobian (determinant)* of the transformation, which in the case of $\boldsymbol{y} = \boldsymbol{\varphi}(\boldsymbol{x}, t)$ is

$$\det(\nabla_{\boldsymbol{x}} \boldsymbol{\varphi}(\boldsymbol{x}, t)) = \det \boldsymbol{F}(\boldsymbol{x}, t),$$

and hence $\mathrm{d}V_{\boldsymbol{y}} = \det \boldsymbol{F}(\boldsymbol{x}, t)\, \mathrm{d}V_{\boldsymbol{x}}$. This leads to the following change of variable formula for volume integrals.

Proposition 5.8. *If $\phi(\boldsymbol{y}, t)$ is a spatial scalar field defined on B_t, and let $\Omega_t \subset B_t$, with $\Omega = \boldsymbol{\psi}(\Omega_t, t) \subset B$. Then*

$$\int_{\Omega_t} \phi(\boldsymbol{y}, t)\, \mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega} \phi_m(\boldsymbol{x}, t) \det \boldsymbol{F}(\boldsymbol{x}, t)\, \mathrm{d}V_{\boldsymbol{x}}.$$

Remarks:

1. By arguments similar to those applied to the expressions of the Divergence Theorem and Stokes' Theorem, we can use the change of variable formula to show that

   $$\det \boldsymbol{F}(\boldsymbol{x}_0, t) = \lim_{\delta \to 0} \frac{\mathrm{vol}(\Omega_{\delta,t})}{\mathrm{vol}(\Omega_{\delta,0})},$$

   where $\Omega_{\delta,0}$ is a ball of radius δ centred at $\boldsymbol{x}_0 \in B$, and $\Omega_{\delta,t} = \boldsymbol{\varphi}(\Omega_{\delta,0}, t)$ is the image of the set after deformation. This means that the determinant of the deformation gradient measures the volume change due to deformation at a fixed material point.

2. The field $J(\boldsymbol{x}, t) = \det \boldsymbol{F}(\boldsymbol{x}, t)$ is called the Jacobian field of a deformation, and is a measure of the *volumetric strain* at a material point $\boldsymbol{x} \in B$ and time t. If $J(\boldsymbol{x}, t) > 1$, the material volume has expanded, while if $J(\boldsymbol{x}, t) < 1$, the material volume has contracted.

The latter remark leads us to define a *volume-preserving* or *isochoric* motion to be any motion such that $\mathrm{vol}(\boldsymbol{\varphi}(\Omega, t)) = \mathrm{vol}(\Omega)$ for all open subsets $\Omega \subset B$, and in this case, we have the following result:

Proposition 5.9. *Suppose $\boldsymbol{\varphi} : B \times [0, T] \to \mathbb{E}^3$ is a motion with spatial velocity field $\boldsymbol{v}$ and deformation gradient $\boldsymbol{F}$. Then $\boldsymbol{\varphi}$ is volume-preserving in the sense above if and only if $\det \boldsymbol{F}(\boldsymbol{x}, t) = 1$ for all $\boldsymbol{x} \in B$ and $t \in [0, T]$, or equivalently,*

$$\nabla_{\boldsymbol{y}} \cdot \boldsymbol{v}(\boldsymbol{y}, t) = 0 \quad \textit{for all } \boldsymbol{y} \in B_t \textit{ and } t \in [0, T].$$

We note that simple translations and rotations are volume-preserving since they produce no distortion, but motions which do distort a body can also be volume-preserving.

5.5.2 Derivatives of time-dependent integrals

We sometimes need to compute the time derivative of a spatial integral, and for this, the following result is useful.

Proposition 5.10. *If $\boldsymbol{\varphi} : B \times [0,T] \to \mathbb{E}^3$ is a motion with spatial velocity field $\boldsymbol{v}(\boldsymbol{y},t)$ and deformation gradient $\boldsymbol{F}(\boldsymbol{x},t)$, then*

$$\frac{\partial}{\partial t} \det \boldsymbol{F}(\boldsymbol{x},t) = \det \boldsymbol{F}(\boldsymbol{x},t)(\nabla_{\boldsymbol{y}} \cdot \boldsymbol{v})(\boldsymbol{y},t)\Big|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)}.$$

Proof. Using Proposition 3.11 and the chain rule, we have

$$\frac{\partial}{\partial t} \det \boldsymbol{F}(\boldsymbol{x},t) = \det \boldsymbol{F}(\boldsymbol{x},t) \operatorname{tr}\left(\boldsymbol{F}(\boldsymbol{x},t)^{-1} \frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{x},t)\right), \tag{5.2}$$

and the components of the spatial velocity field satisfy

$$v_i(\boldsymbol{y},t)\Big|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} = \frac{\partial}{\partial t}\varphi_i(\boldsymbol{x},t).$$

Our aim is now to express the trace term in Equation 5.2 in terms of $\boldsymbol{v}$. Notice that

$$F_{ij}(\boldsymbol{x},t) = \frac{\partial \varphi_i}{\partial x_j}(\boldsymbol{x},t),$$

so differentiating in time,

$$\begin{aligned}
\frac{\partial F_{ij}}{\partial t}(\boldsymbol{x},t) = \frac{\partial}{\partial t}\frac{\partial \varphi_i}{\partial x_j}(\boldsymbol{x},t) &= \frac{\partial}{\partial x_j}\frac{\partial \varphi_i}{\partial t}(\boldsymbol{x},t) \\
&= \frac{\partial}{\partial x_j}\boldsymbol{v}(\boldsymbol{\varphi}(\boldsymbol{x},t),t) \\
&= \frac{\partial}{\partial y_k}\boldsymbol{v}(\boldsymbol{y},t)\Big|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} \frac{\partial \varphi_k}{\partial x_j}(\boldsymbol{x},t) \\
&= \frac{\partial}{\partial y_k}\boldsymbol{v}(\boldsymbol{y},t)\Big|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} F_{kj}(\boldsymbol{x},t)
\end{aligned}$$

In tensor notation, we have shown that

$$\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{x},t) = \nabla_{\boldsymbol{y}}\boldsymbol{v}(\boldsymbol{y},t)\Big|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} \boldsymbol{F}(\boldsymbol{x},t),$$

so multiplying by $\boldsymbol{F}^{-1}$ on the right and taking the trace, we have

$$\operatorname{tr}\left(\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{x},t)\boldsymbol{F}^{-1}(\boldsymbol{x},t)\right) = \operatorname{tr}\big(\nabla_{\boldsymbol{y}}\boldsymbol{v}(\boldsymbol{y},t)\big)\Big|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)} = \nabla_{\boldsymbol{y}} \cdot \boldsymbol{v}(\boldsymbol{y},t)\Big|_{\boldsymbol{y}=\boldsymbol{\varphi}(\boldsymbol{x},t)}.$$

We can now conclude by using this expression in Equation 5.2, and noting that $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ for any two tensors $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{V}^2$. □

This result shows that the time derivative of the Jacobian field depends only upon the field itself and the divergence of the spatial velocity. This leads to the following famous result.

Proposition 5.11 (Reynolds' Transport Theorem). *Suppose $\boldsymbol{\varphi} : B \times [0,T] \to \mathbb{E}^3$ is a motion with an associated spatial velocity field $\boldsymbol{v}$. Let $\Omega_t \subset B_t$ be an arbitrary volume with boundary $\partial\Omega_t$ which has outward normal field $\boldsymbol{n}$. Then for any spatial scalar field ϕ we have*

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t} \phi \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega_t} \frac{\mathrm{D}}{\mathrm{D}t}\phi + \phi(\nabla_{\boldsymbol{y}} \cdot \boldsymbol{v}) \,\mathrm{d}V_{\boldsymbol{y}},$$

or equivalently

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t} \phi \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega_t} \frac{\partial \phi}{\partial t} \,\mathrm{d}V_{\boldsymbol{y}} + \int_{\partial\Omega_t} \phi \boldsymbol{v} \cdot \boldsymbol{n} \,\mathrm{d}A_{\boldsymbol{y}}.$$

Remarks:

1. Once more, this theorem shows that if the spatial velocity $\boldsymbol{v}$ is known, then the time derivative of a volume integral over any subset $\Omega_t \subset B_t$ can be determined without explicit knowledge of $\boldsymbol{\varphi}$ or its inverse $\boldsymbol{\psi}$.

2. The fact this is called a 'transport theorem' is motivated by the second equality. The first term on the RHS is the change of the field ϕ in the volume Ω_t, while the second term reflect the amount of ϕ which is transported across the boundary $\partial\Omega_t$, as determined by the normal component of the velocity $\boldsymbol{v} \cdot \boldsymbol{n}$.

5.5.3 Transformation of surface integrals

Along with volume integrals, we also consider how surface integrals transform under the mapping from the reference configuration to the deformed configuration. Suppose that $\Gamma \subset B$ is a referential surface with a unit normal field $\boldsymbol{\nu} : \Gamma \to \mathcal{V}$, and let $\gamma_t = \boldsymbol{\varphi}(\Gamma, t)$ be the corresponding surface in the deformed configuration, with accompanying normal field $\boldsymbol{n} : \gamma_t \to \mathcal{V}$.

It can be shown that in this case that the oriented surface area element satisfies

$$\boldsymbol{n}(\boldsymbol{y}) \,\mathrm{d}A_{\boldsymbol{y}} = \operatorname{cof}\big(\boldsymbol{F}(\boldsymbol{x},t)\big)\boldsymbol{\nu}(\boldsymbol{x}) \,\mathrm{d}A_{\boldsymbol{x}}.$$

The argument is similar to that made for a transformation of surface integrals, but since it is relatively technical, we omit it. This leads us to the following result.

Suppose ϕ, $\boldsymbol{w}$, and $\boldsymbol{T}$ are spatial scalar, vector and tensor fields, and that $\Omega \subset B$ is a subset which is mapped to $\Omega_t = \boldsymbol{\varphi}(B,t) \subset B_t$ under the motion $\boldsymbol{\varphi}$, where $\partial\Omega \subset B$ has

outward-pointing normal field $\boldsymbol{\nu}$, and $\partial\Omega_t \subset B_t$ has outward-pointing normal field $\boldsymbol{n}$. Then

$$\int_{\partial\Omega_t} \phi(\boldsymbol{y},t)\boldsymbol{n}(\boldsymbol{y})\,\mathrm{d}A_{\boldsymbol{y}} = \int_{\partial\Omega} \phi_m(\boldsymbol{x},t)\mathrm{cof}\boldsymbol{F}(\boldsymbol{x},t)\boldsymbol{\nu}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}}$$

$$\int_{\partial\Omega_t} \boldsymbol{w}(\boldsymbol{y},t)\cdot\boldsymbol{n}(\boldsymbol{y})\,\mathrm{d}A_{\boldsymbol{y}} = \int_{\partial\Omega} \boldsymbol{w}_m(\boldsymbol{x},t)\cdot\big(\mathrm{cof}\boldsymbol{F}(\boldsymbol{x},t)\boldsymbol{\nu}(\boldsymbol{x})\big)\,\mathrm{d}A_{\boldsymbol{x}}$$

$$\int_{\partial\Omega_t} \boldsymbol{T}(\boldsymbol{y},t)\boldsymbol{n}(\boldsymbol{y})\,\mathrm{d}A_{\boldsymbol{y}} = \int_{\partial\Omega} \boldsymbol{T}_m(\boldsymbol{x},t)\mathrm{cof}\boldsymbol{F}(\boldsymbol{x},t)\boldsymbol{\nu}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}},$$

where we recall that the cofactor tensor (as previously defined in the statement of Proposition 3.11) is defined to be

$$\mathrm{cof}(\boldsymbol{F}) = (\det\boldsymbol{F})\boldsymbol{F}^{-T}.$$

6 Balance Laws

In this chapter, we state a variety of axioms for a mechanical theory of continuum bodies. As in Chapter 4, we first state laws in integral form, and then derive local versions, which yield (systems of) PDEs. It is important to state at the outset that these laws apply to all continuum bodies, regardless of whether they are solid or fluid.

Aims. By the end of this chapter, you should be able to:

- Provide definitions of the *mass*, *linear* and *angular momentum* of a subset of body as integrals.
- Explain why we must include a description of temperature in a continuum theory of matter.
- State the physical principles of the *conservation of mass*, *laws of inertia*, and the *first law of thermodynamics* as used in continuum mechanics.
- Using physical principles and theorems to derive local forms of the balance laws, in both the Eulerian and Lagrangian formalisms.
- Explain how the principle of *frame indifference* limits the constitutive laws which are possible to close the equations derived.

6.1 Motivation

Consider N particles with masses $m_i \in \mathbb{R}$ and positions $\boldsymbol{x}_i \in \mathbb{E}^3$. We may think of these as the atoms (or perhaps molecules) making up a continuum body. Suppose that the potential energy of interaction between particles i and j is described by an energy U_{ij}, which is symmetric in exchange of particle indices, i.e. $U_{ij} = U_{ji}$. The force on particle i due to interaction with particle j is denoted

$$\boldsymbol{f}_{ij}^{\text{int}} = -\nabla_{\boldsymbol{x}_i} U_{ij}.$$

Assuming that the environment of the particle also acts on each particle through a force $\boldsymbol{f}_i^{\text{env}}$, and these are the only forces acting, then a complete description of the motion of the particles is:

$$\dot{m}_i = 0, \qquad m_i \ddot{\boldsymbol{x}}_i = \boldsymbol{f}_i^{\text{env}} + \sum_{j \neq i} \boldsymbol{f}_{ij}^{\text{int}},$$

where we do not apply the summation convention, and i ranges from $1, \dots, N$. These two sets of equations state that the mass of each particle is constant, and each particle obeys Newton's Second Law.

Consider an arbitrary subset of indices $I \subset \{1, \dots, N\}$, and let Ω_t be the current configuration of particles with indices I at time $t \geq 0$. We can associate to Ω_t a total mass $M[\Omega_t]$, a linear momentum $\boldsymbol{l}[\Omega_t]$, an angular momentum $\boldsymbol{j}[\Omega_t]$, an internal energy $U[\Omega_t]$ and a kinetic energy $K[\Omega_t]$ as follows:

$$M[\Omega_t] = \sum_{i \in I} m_i, \quad \boldsymbol{l}[\Omega_t] = \sum_{i \in I} m_i \dot{\boldsymbol{x}}_i, \quad \boldsymbol{j}[\Omega_t] = \sum_{i \in I} \boldsymbol{x}_i \times m_i \dot{\boldsymbol{x}}_i,$$
$$U[\Omega_t] = \sum_{\substack{i,j \in I \\ i<j}} U_{ij}, \quad K[\Omega_t] = \sum_{i \in I} \tfrac{1}{2} m_i |\dot{\boldsymbol{x}}_i|^2.$$

If we make the assumption that U_{ij} depends only on $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ through the distance $|\boldsymbol{x}_i - \boldsymbol{x}_j|$, then we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} M[\Omega_t] &= 0, \\
\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{l}[\Omega_t] &= \sum_{i \in I} \left(\boldsymbol{f}_i^{\text{env}} + \sum_{j \notin I} \boldsymbol{f}_{ij}^{\text{int}} \right) \\
\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{j}[\Omega_t] &= \sum_{i \in I} \boldsymbol{x}_i \times \left(\boldsymbol{f}_i^{\text{env}} + \sum_{j \notin I} \boldsymbol{f}_{ij}^{\text{int}} \right) \\
\frac{\mathrm{d}}{\mathrm{d}t} \big(U[\Omega_t] + K[\Omega_t] \big) &= \sum_{i \in I} \dot{\boldsymbol{x}}_i \cdot \left(\boldsymbol{f}_i^{\text{env}} + \sum_{j \notin I} \boldsymbol{f}_{ij}^{\text{int}} \right).
\end{aligned}$$

These equations state that the mass of Ω_t does not change with time, the rate of change of linear momentum is equal to the resultant external force acting on Ω_t, the rate of change of the angular momentum is equal to the resultant external torque acting on Ω_t, and the rate of change of the energy is equal to the power of the external forces acting on Ω_t.

The equations above are balance laws, and describe how various physically important quantities are conserved or change due to external influences. These four laws provide a fundamental starting point for the modelling of continuum bodies, by (at least informally) letting the number of particles tend to infinity, and replacing sums by integrals. In the case of mass and linear and angular momentum, this limit is straightforward. However, for the energy balance, things are more complicated: a continuum field cannot directly represent fluctuations on the atomic scale. As such, continuum fields should be treated as 'averages over' sufficiently large collections of atoms, where fluctuations are smoothed out.

6.2 Balance laws in integral form

Consider a motion $\boldsymbol{\varphi}$, which maps a reference configuration B to the current configuration $B_t = \boldsymbol{\varphi}(B, t)$. Let $\rho(\boldsymbol{y}, t)$ be the mass density and $\boldsymbol{v}(\boldsymbol{y}, t)$ be the velocity at a spatial point $\boldsymbol{y} \in B_t$, and let $\Omega_t \subseteq B_t$ be an arbitrary open subset of B_t. Then:

$$\begin{aligned}
\text{The } \textit{mass} \text{ of } \Omega_t \text{ is:} \quad & \operatorname{mass}[\Omega_t] = \int_{\Omega_t} \rho(\boldsymbol{y}, t)\, \mathrm{d}V_{\boldsymbol{y}}, \\
\text{The } \textit{linear momentum} \text{ of } \Omega_t \text{ is:} \quad & \boldsymbol{l}[\Omega_t] = \int_{\Omega_t} \rho(\boldsymbol{y}, t)\boldsymbol{v}(\boldsymbol{y}, t)\, \mathrm{d}V_{\boldsymbol{y}}, \\
\text{The } \textit{angular momentum} \text{ of } \Omega_t \text{ about } \boldsymbol{z} \text{ is:} \quad & \boldsymbol{j}[\Omega_t; \boldsymbol{z}] = \int_{\Omega_t} (\boldsymbol{y} - \boldsymbol{z}) \times \rho(\boldsymbol{y}, t)\boldsymbol{v}(\boldsymbol{y}, t)\, \mathrm{d}V_{\boldsymbol{y}}.
\end{aligned}$$

6.2.1 Conservation of mass and the laws of inertia

Assuming that we are neglecting relativistic effects and possible chemical reactions, the balance laws can be generalised to continuum bodies as follows.

Conservation of Mass

The mass of any open subset of a continuum body is invariant as the body changes position and shape, i.e.

$$\frac{\mathrm{D}}{\mathrm{D}t}\operatorname{mass}[\Omega_t] = 0 \quad \text{for all } \Omega_t \subseteq B_t.$$

Now, recalling the definitions of the resultant force $\boldsymbol{r}[\Omega]$ and resultant torque $\boldsymbol{\tau}[\Omega; \boldsymbol{z}]$, we can state the laws of inertia:

The Laws of Inertia

The rate of change of linear momentum of any open subset of a continuum body equals the resultant force applied to it, and the rate of change of angular momentum equals the resultant torque, i.e.

$$\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{l}[\Omega_t] = \boldsymbol{r}[\Omega_t] \quad \text{and} \quad \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{j}[\Omega_t; \boldsymbol{z}] = \boldsymbol{\tau}[\Omega_t; \boldsymbol{z}] \qquad \text{for all } \Omega_t \subseteq B_t.$$

Both **Conservation of Mass** and **The Laws of Inertia** are assumed to hold for every motion and every subset of a continuum body.

6.2.2 Conservation of energy

Temperature is a physical property of matter that is based on our perceptions of hot and cold, and quantifies the microscopic fluctuations in the velocity of the atoms making up the material about their mean value.

To see this more clearly, we note that the continuum velocity field $\boldsymbol{v}(\boldsymbol{y}, t)$ measures the average velocity of the atoms in the body at a spatial point $\boldsymbol{y} \in B_t$, but it is perfectly possible that, at some instant in time, many of the atoms close to this point are moving in different directions. This variation in the microscopic velocities of the atoms is what temperature is. To describe temperature at the continuum level, we will assume the existence of a (scalar) absolute *temperature field* $\theta(\boldsymbol{y}, t) > 0$ for all points $\boldsymbol{y} \in B_t$ and times $t \geq 0$. This field is positive, since we always assume that there is some (possibly very small) variation about the mean.

By the *thermal energy* or *heat content* of a body we mean an energy associated with the velocity fluctuations of atoms within the body, and so with its temperature. We note however that heat and temperature are not the same! We know from our own experience that heat can be converted into mechanical work: deforming a body can affect its temperature and therefore its heat content. Similarly, heating a body can affect its temperature and set it in motion, for example through thermal expansion. This tells us that we may use the same units to describe heat energy as we use to describe mechanical work.

Heat energy can be transformed in two basic ways:

- It can be produced or consumed by mechanical, chemical, or electromagnetic processes. For example, a flowing electrical current produces heat as it encounters resistance in the body.
- It can be transferred from one body to another through physical contact or thermal radiation. Heat transfer through contact can be further classified as heat conduction or convection depending on the circumstances and nature of the bodies involved.

Given $\Omega_t \subseteq B_t$, we define the *net rate of heating* $Q[\Omega_t]$ to be the rate at which heat is being added to Ω_t. As suggested above (and in analogy with the treatment of forces in @cha-mass-forces, we assume that the net heating can be decomposed into a bulk and surface heating:

$$Q[\Omega_t] = Q_b[\Omega_t] + Q_s[\partial\Omega_t].$$

As with our treatment of mass density and body forces, we assume there exists a *heat supply field* per unit volume $\hat{r}(\boldsymbol{y}, t) \in \mathbb{R}$ such that

$$Q_b[\Omega_t] = \int_{\Omega_t} \hat{r}(\boldsymbol{y}, t)\, \mathrm{d}V_{\boldsymbol{y}},$$

and equivalently, a heat supply field per unit mass $r(\boldsymbol{y},t) \in \mathbb{R}$ with $r(\boldsymbol{y},t) = \hat{r}(\boldsymbol{y},t)/\rho(\boldsymbol{y},t)$ so that, equivalently,

$$Q_b[\Omega_t] = \int_{\Omega_t} \rho(\boldsymbol{y},t) r(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}}.$$

We similarly assume that there is a *heat transfer field* per unit area $h(\boldsymbol{y},t) \in \mathbb{R}$ so that

$$Q_s[\partial\Omega_t] = \int_{\partial\Omega_t} h(\boldsymbol{y},t)\,\mathrm{d}A_{\boldsymbol{y}}.$$

More specifically, we assume that the heat transfer field takes the form $h(\boldsymbol{y},t) = -\boldsymbol{q}(\boldsymbol{y},t)\cdot \boldsymbol{n}(\boldsymbol{y})$, where $\boldsymbol{n}$ is the outward-pointing unit normal field for the surface $\partial\Omega_t$, and $\boldsymbol{q}$ is the *heat flux vector field* in B_t.

The direction and intensity of the heat flow across any surface is determined by the field $\boldsymbol{q}$, and the negative sign in the relation $h = -\boldsymbol{q}\cdot\boldsymbol{n}$ is due to choosing the outward rather than the inward pointing normal (we compute the flow of heat *into* the region Ω_t). The relationship between the surface heat transfer h and the flux $\boldsymbol{q}$ is analogous to the relationship between the traction $\boldsymbol{t}$ and the Cauchy stress tensor $\boldsymbol{S}$. Indeed, similar conditions and arguments to those deployed in Proposition 4.1 and Proposition 4.2 show that heat transfer must take this form.

Together, we now have that

$$Q[\Omega_t] = \int_{\Omega_t} \rho(\boldsymbol{y},t) r(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}} - \int_{\partial\Omega_t} \boldsymbol{q}(\boldsymbol{y},t)\cdot\boldsymbol{n}(\boldsymbol{y})\,\mathrm{d}A_{\boldsymbol{y}}.$$

The *kinetic energy* of Ω_t is

$$K[\Omega_t] = \int_{\Omega_t} \tfrac{1}{2}\rho(\boldsymbol{y},t)|\boldsymbol{v}(\boldsymbol{y},t)|^2\,\mathrm{d}V_{\boldsymbol{y}},$$

and the *power of external forces* acting on Ω_t is

$$P[\Omega_t] = \int_{\Omega_t} \rho(\boldsymbol{y},t)\boldsymbol{b}(\boldsymbol{y},t)\cdot\boldsymbol{v}(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}} + \int_{\partial\Omega_t} \boldsymbol{t}(\boldsymbol{y},t)\cdot\boldsymbol{v}(\boldsymbol{y},t)\,\mathrm{d}A_{\boldsymbol{y}}.$$

The *net rate of internal mechanical work* of external forces acting on Ω_t is the power not used up in producing motion, i.e.

$$W[\Omega_t] = P[\Omega_t] - \frac{\mathrm{D}}{\mathrm{D}t}K[\Omega_t].$$

It $W[\Omega_t] = 0$ over some time interval, then all of the mechanical energy delivered to the body by external forces is converted to kinetic energy. If $W[\Omega_t] > 0$ over some time interval, then some of the energy delivered to the body is stored, and if $W[\Omega_t] < 0$, energy is released from the body.

From the above discussion, we see that there may be some part of the energy of a continuum body which is not kinetic energy. This component of the energy is called the *internal energy*, and represents a store of energy that can be increased in various ways (e.g. by heating or applying forces to the body), or released (e.g. by heat transfer or through performing mechanical work). We will assume that the internal energy of a body consists only of thermal and mechanical energy, neglecting chemical and electromagnetic effects (which can be treated within a similar framework).

Given a subset $\Omega_t \subseteq B_t$, we denote its internal energy $U[\Omega_t]$, and assume that there exists an *internal energy density field* per unit volume $\hat{\phi}(\boldsymbol{y}, t) \in \mathbb{R}$ such that

$$U[\Omega_t] = \int_{\Omega_t} \hat{\phi}(\boldsymbol{y}, t) \, \mathrm{d}V_{\boldsymbol{y}}.$$

As usual, we can define a corresponding internal energy density per unit mass $\phi(\boldsymbol{y}, t) = \hat{\phi}(\boldsymbol{y}, t)/\rho(\boldsymbol{y}, t)$ so that

$$U[\Omega_t] = \int_{\Omega_t} \rho(\boldsymbol{y}, t)\phi(\boldsymbol{y}, t) \, \mathrm{d}V_{\boldsymbol{y}}.$$

The First Law of Thermodynamics (Conservation of Energy)

The rate of change of the internal energy of any open subset of a continuum body is the sum of the net rate of heating and net rate of work applied to it, i.e.

$$\frac{\mathrm{D}}{\mathrm{D}t} U[\Omega_t] = Q[\Omega_t] + W[\Omega_t] \quad \text{for all } \Omega_t \subseteq B_t,$$

or using the relationship between W, P and K:

$$\frac{\mathrm{D}}{\mathrm{D}t} \Big(U[\Omega_t] + K[\Omega_t] \Big) = Q[\Omega_t] + P[\Omega_t] \quad \text{for all } \Omega_t \subseteq B_t.$$

There are various fundamental differences between the continuum versions of the energy balance and the corresponding law for particles. These arise from the fact that our continuum description cannot capture the atomic fluctuations, which are separated from the kinetic energy.

The energy balance relates the rate of change of internal energy with the net rate of heating and net rate of work due to external influences on a body. While these must be balanced, there is no restriction on the rates of these processes themselves. For example, if the net rate of work is zero over some time interval, then there is no limit on the rate at which a body may absorb heat and store it as internal energy. While we will not focus on it here, The Second Law of Thermodynamics expresses the fact that bodies *do* have limits on the rate at which heat can be absorbed, but no limit on the rate at which heat can be released. This also entails constitutive restrictions, but we will focus mainly on *isothermal* processes, so will not discuss this further (unless I find time later).

6.3 Localised Eulerian form of balance laws

So far, we have summarised balance laws for the mass, momentum, and energy in terms of arbitrary open subsets of the current configuration B_t. We will now exploit the Localisation Theorem, Proposition 3.8, to derive partial differential equations for the various fields we have considered. When expressed in terms of the spatial coordinates, we refer to such balance laws as *Eulerian*, and when expressed in material coordinates, we refer to the balance laws as being in *Lagrangian* form. Our focus is first on the Eulerian form, and we will assume in this section that $\boldsymbol{\varphi} : B \times [0,T] \to \mathbb{E}^3$ is a motion which satisfies $\boldsymbol{\varphi}(\boldsymbol{x}, 0) = \boldsymbol{x}$.

6.3.1 Conservation of mass

Consider $\Omega_t \subseteq B_t$, and let $\Omega \subseteq B$ such that $\boldsymbol{\varphi}(\Omega, t) = \Omega_t$. **Conservation of Mass** entails that

$$\text{mass}[\Omega_t] = \text{mass}[\Omega_0],$$

where $\Omega_0 = \Omega$. Furthermore, we have

$$\text{mass}[\Omega_t] = \int_{\Omega_t} \rho(\boldsymbol{y}, t)\, \mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega} \rho_m(\boldsymbol{x}, t) \det \boldsymbol{F}(\boldsymbol{x}, t)\, \mathrm{d}V_{\boldsymbol{x}},$$

and as we have assumed that $\boldsymbol{\varphi}(\boldsymbol{x}, 0) = \boldsymbol{x}$, we have

$$\text{mass}[\Omega_0] = \int_{\Omega} \rho(\boldsymbol{x}, 0)\, \mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega} \rho_0(\boldsymbol{x})\, \mathrm{d}V_{\boldsymbol{x}},$$

where $\rho_0(\boldsymbol{x}) = \rho(\boldsymbol{x}, 0)$. We must therefore have

$$\int_{\Omega} \rho_m(\boldsymbol{x}, t) \det \boldsymbol{F}(\boldsymbol{x}, t) - \rho_0(\boldsymbol{x})\, \mathrm{d}V_{\boldsymbol{x}} = 0$$

for any $t \geq 0$ and any $\Omega \subseteq B$, so the Localisation Theorem entails that

$$\rho_m(\boldsymbol{x}, t) \det \boldsymbol{F}(\boldsymbol{x}, t) = \rho_0(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x} \in B, t \geq 0. \tag{6.1}$$

To convert this form to Eulerian form, we take the total time derivative of both sides of this equation, finding that

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\rho(\boldsymbol{\varphi}(\boldsymbol{x}, t), t)\right) \det \boldsymbol{F}(\boldsymbol{x}, t) + \rho(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) \frac{\mathrm{D}}{\mathrm{D}t} \det \boldsymbol{F}(\boldsymbol{x}, t) = 0.$$

For the latter term, we apply Proposition 5.10, so that

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\rho(\boldsymbol{\varphi}(\boldsymbol{x}, t), t)\right) \det \boldsymbol{F}(\boldsymbol{x}, t) + \rho(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) \det \boldsymbol{F}(\boldsymbol{x}, t) \nabla_{\boldsymbol{y}} \cdot \boldsymbol{v}(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) = 0.$$

Dividing through by the non-zero determinant term, and taking the total derivative in the first term, then re-expressing everything in spatial coordinates by substituting $\boldsymbol{y} = \boldsymbol{\varphi}(\boldsymbol{x}, t)$, we have

$$\frac{\partial \rho}{\partial t}(\boldsymbol{y}, t) + \nabla_{\boldsymbol{y}} \rho(\boldsymbol{y}, t) \cdot \boldsymbol{v}(\boldsymbol{y}, t) + \rho(\boldsymbol{y}, t) \nabla_{\boldsymbol{y}} \cdot \boldsymbol{v}(\boldsymbol{y}, t) = 0.$$

Finally, using the property of the divergence that $\nabla \cdot (\phi \boldsymbol{w}) = \nabla \phi \cdot \boldsymbol{w} + \phi \nabla \cdot \boldsymbol{w}$, we have:

$$\frac{\partial \rho}{\partial t}(\boldsymbol{y}, t) + \nabla_{\boldsymbol{y}} \cdot \big(\rho(\boldsymbol{y}, t) \boldsymbol{v}(\boldsymbol{y}, t)\big) = 0.$$

We have just proved the following result:

Proposition 6.1. *If $\boldsymbol{\varphi} : B \times [0, T] \to \mathbb{E}^3$ is a motion with associated spatial velocity field $\boldsymbol{v}$ and spatial mass density field ρ, then the conservation of mass entails that*

$$\frac{\partial \rho}{\partial t} + \nabla_{\boldsymbol{y}} \cdot (\rho \boldsymbol{v}) = 0$$

for all $\boldsymbol{y} \in B_t$ and $t \in [0, T]$. Equivalently, we have

$$\frac{\mathrm{D}}{\mathrm{D}t} \rho + \rho \nabla_{\boldsymbol{y}} \cdot \boldsymbol{v} = 0$$

for all $\boldsymbol{y} \in B_t$ and $t \in [0, T]$.

The following result allows us to express the time derivative of integrals involving the mass density ρ.

Proposition 6.2. *If $\boldsymbol{\varphi} : B \times [0, T] \to \mathbb{E}^3$ is a motion with associated spatial velocity field $\boldsymbol{v}$ and spatial mass density field ρ, Φ is any spatial scalar, vector or tensor field, and $\Omega_t \subseteq B_t$, then*

$$\frac{\mathrm{D}}{\mathrm{D}t} \int_{\Omega_t} \Phi(\boldsymbol{y}, t) \rho(\boldsymbol{y}, t) \, \mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega_t} \frac{\mathrm{D}}{\mathrm{D}t} \Phi(\boldsymbol{y}, t) \rho(\boldsymbol{y}, t) \, \mathrm{d}V_{\boldsymbol{y}}.$$

Proof. Applying Proposition 5.8, we have

$$\int_{\Omega_t} \Phi(\boldsymbol{y}, t) \rho(\boldsymbol{y}, t) \, \mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega} \Phi(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) \rho(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) \det \boldsymbol{F}(\boldsymbol{x}, t) \, \mathrm{d}V_{\boldsymbol{x}},$$

where $\Omega_t = \boldsymbol{\varphi}(\Omega, t)$. Applying the equality Equation 6.1, i.e. $\rho(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) \det \boldsymbol{F}(\boldsymbol{x}, t) = \rho_0(\boldsymbol{x})$, we can rewrite the right-hand side as

$$\int_{\Omega_t} \Phi(\boldsymbol{y}, t) \rho(\boldsymbol{y}, t) \, \mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega} \Phi(\boldsymbol{\varphi}(\boldsymbol{x}, t), t) \rho_0(\boldsymbol{x}) \, \mathrm{d}V_{\boldsymbol{x}}.$$

Differentiating with respect to time, and exchanging the order of integration and differentiation on the RHS, we have

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t}\Phi(\boldsymbol{y},t)\rho(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}}=\int_{\Omega}\frac{\mathrm{D}}{\mathrm{D}t}\Big(\Phi(\boldsymbol{\varphi}(\boldsymbol{x},t),t)\Big)\rho_0(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}}.$$

But now, using equality Equation 6.1 again, and then applying the definition of the total derivative and Proposition 5.8, we have

$$\begin{aligned}\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t}\Phi(\boldsymbol{y},t)\rho(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}}&=\int_{\Omega}\frac{\mathrm{D}}{\mathrm{D}t}\Big(\Phi(\boldsymbol{\varphi}(\boldsymbol{x},t),t)\Big)\rho(\boldsymbol{\varphi}(\boldsymbol{x},t),t)\det\boldsymbol{F}(\boldsymbol{x},t)\,\mathrm{d}V_{\boldsymbol{x}}\\&=\int_{\Omega_t}\frac{\mathrm{D}}{\mathrm{D}t}\Phi(\boldsymbol{y},t)\rho(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}}.\end{aligned}$$

□

6.3.2 Balance of linear momentum

Recalling the **Laws of Inertia**, we have

$$\int_{\Omega_t}\rho\boldsymbol{v}\,\mathrm{d}V_{\boldsymbol{y}}=\int_{\partial\Omega_t}\boldsymbol{t}\,\mathrm{d}A_{\boldsymbol{y}}+\int_{\Omega_t}\rho\boldsymbol{b}\,\mathrm{d}V_{\boldsymbol{y}},$$

where all of the integrated quantities are spatial fields. The results of Chapter 4 state that the traction field $\boldsymbol{t}=\boldsymbol{S}\boldsymbol{n}$, so that

$$\int_{\Omega_t}\rho\boldsymbol{v}\,\mathrm{d}V_{\boldsymbol{y}}=\int_{\partial\Omega_t}\boldsymbol{S}\boldsymbol{n}\,\mathrm{d}A_{\boldsymbol{y}}+\int_{\Omega_t}\rho\boldsymbol{b}\,\mathrm{d}V_{\boldsymbol{y}}.$$

Using the tensorial version of the Divergence Theorem, we have

$$\int_{\Omega_t}\rho\boldsymbol{v}\,\mathrm{d}V_{\boldsymbol{y}}=\int_{\Omega_t}\nabla_{\boldsymbol{y}}\cdot\boldsymbol{S}+\rho\boldsymbol{b}\,\mathrm{d}V_{\boldsymbol{y}}.$$

Now, as Ω_t is an arbitrary subset of B_t, the Localisation Theorem entails that we have the following result:

Proposition 6.3. *If $\boldsymbol{\varphi}:B\times[0,T]\to\mathbb{E}^3$ is a motion with associated spatial velocity field $\boldsymbol{v}$ and spatial mass density field ρ, then the balance of linear momentum entails that*

$$\rho\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v}=\nabla_{\boldsymbol{y}}\cdot\boldsymbol{S}+\rho\boldsymbol{b}$$

for all $\boldsymbol{y}\in B_t$ and $t\in[0,T]$, where $\boldsymbol{S}(\boldsymbol{y},t)$ is the Cauchy stress field and $\boldsymbol{b}(\boldsymbol{y},t)$ is the spatial body force field per unit mass.

This is the dynamic version of the equilibrium condition established in Proposition 4.3: the static result corresponds to the case where $\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v} = 0$ for all $\boldsymbol{y} \in B_t$ and $t \in [0, T]$, i.e. there is no motion.

Recall that

$$\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v} = \frac{\partial \boldsymbol{v}}{\partial t} + (\nabla_{\boldsymbol{y}}\boldsymbol{v})\boldsymbol{v},$$

so we may equivalently write

$$\rho\left(\frac{\partial \boldsymbol{v}}{\partial t} + (\nabla_{\boldsymbol{y}}\boldsymbol{v})\boldsymbol{v}\right) = \nabla_{\boldsymbol{y}} \cdot \boldsymbol{S} + \rho\boldsymbol{b}.$$

6.3.3 Balance of angular momentum

Taking the point about which we compute angular momentum to be $\boldsymbol{z} = \boldsymbol{0}$, recall that the balance of angular momentum is

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t} \boldsymbol{y} \times (\rho\boldsymbol{v})\,\mathrm{d}V_{\boldsymbol{y}} = \int_{\partial\Omega_t} \boldsymbol{y} \times \boldsymbol{t}\,\mathrm{d}A_{\boldsymbol{y}} + \int_{\Omega_t} \boldsymbol{y} \times \rho\boldsymbol{b}\,\mathrm{d}V_{\boldsymbol{y}}, \tag{6.2}$$

where as before, all fields in the integrands are spatial. To simplify the LHS of this equation, note that $\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{y} = \boldsymbol{v}$, and hence

$$\frac{\mathrm{D}}{\mathrm{D}t}(\boldsymbol{y} \times \boldsymbol{v}) = \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{y} \times \boldsymbol{v} + \boldsymbol{y} \times \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v} = \underbrace{\boldsymbol{v} \times \boldsymbol{v}}_{=\boldsymbol{0}} + \boldsymbol{y} \times \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v}.$$

Moreover, Proposition 6.2 entails that

$$\begin{aligned}\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t} (\boldsymbol{y} \times \boldsymbol{v})\rho\,\mathrm{d}V_{\boldsymbol{y}} &= \int_{\Omega_t} \rho\frac{\mathrm{D}}{\mathrm{D}t}(\boldsymbol{y} \times \boldsymbol{v})\,\mathrm{d}V_{\boldsymbol{y}} \\ &= \int_{\Omega_t} \rho\,\boldsymbol{y} \times \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v}\,\mathrm{d}V_{\boldsymbol{y}}\end{aligned}$$

Substituting into Equation 6.2 and using the fact that $\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n}$, we obtain

$$\int_{\Omega_t} \boldsymbol{y} \times \left(\rho\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v}\right)\,\mathrm{d}V_{\boldsymbol{y}} = \int_{\partial\Omega_t} \boldsymbol{y} \times \boldsymbol{S}\boldsymbol{n}\,\mathrm{d}A_{\boldsymbol{y}} + \int_{\Omega_t} \boldsymbol{y} \times \rho\boldsymbol{b}\,\mathrm{d}V_{\boldsymbol{y}}.$$

Now, using Proposition 6.3 to substitute for $\rho\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v}$ on the LHS, we have

$$\int_{\Omega_t} \boldsymbol{y} \times (\nabla_{\boldsymbol{y}} \cdot \boldsymbol{S})\,\mathrm{d}V_{\boldsymbol{y}} = \int_{\partial\Omega_t} \boldsymbol{y} \times \boldsymbol{S}\boldsymbol{n}\,\mathrm{d}A_{\boldsymbol{y}}.$$

Now, proceeding as in the proof of Proposition 4.3, we have:

Proposition 6.4. *Let $\boldsymbol{\varphi} : B \times [0,T] \to \mathbb{E}^3$ be a motion, and let $\boldsymbol{S}$ be the Cauchy stress field in B_t. Then the balance of angular momentum entails that*

$$\boldsymbol{S}^T(\boldsymbol{y},t) = \boldsymbol{S}(\boldsymbol{y},t) \quad \textit{for all } \boldsymbol{y} \in B_t, t \in [0,T].$$

The balance of angular momentum therefore requires that the Cauchy stress tensor is always symmetric, no matter whether the body is in motion or not. Note also that the argument used works with any fixed point $\boldsymbol{z}$ about which we compute the angular momentum.

6.3.4 Net work

To study localised version of the First and Second Laws of Thermodynamics, we derive a relationship between the rate of change of kinetic energy and the power of internal and external forces.

We begin by taking the local balance of linear momentum, taking the dot product of each term with the velocity field and integrating to obtain

$$\int_{\Omega_t} \rho \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v} \cdot \boldsymbol{v} \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega_t} (\nabla_{\boldsymbol{y}} \cdot \boldsymbol{S}) \cdot \boldsymbol{v} + \rho \boldsymbol{b} \cdot \boldsymbol{v} \,\mathrm{d}V_{\boldsymbol{y}}.$$

We recall that $\nabla_{\boldsymbol{y}} \cdot (\boldsymbol{S}^T\boldsymbol{v}) = (\nabla_{\boldsymbol{y}} \cdot \boldsymbol{S}) \cdot \boldsymbol{v} + \boldsymbol{S} : \nabla_{\boldsymbol{y}}\boldsymbol{v}$, so applying this equality and the Divergence Theorem, we obtain

$$\int_{\Omega_t} \rho \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v} \cdot \boldsymbol{v} \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega_t} -\boldsymbol{S} : \nabla_{\boldsymbol{y}}\boldsymbol{v} \,\mathrm{d}V_{\boldsymbol{y}} + \int_{\partial\Omega_t} \boldsymbol{S}^T\boldsymbol{v} \cdot \boldsymbol{n} \,\mathrm{d}A_{\boldsymbol{y}} + \int_{\Omega_t} \rho \boldsymbol{b} \cdot \boldsymbol{v} \,\mathrm{d}V_{\boldsymbol{y}}.$$

Recalling that $\boldsymbol{S}$ is symmetric and $\boldsymbol{L} = \mathrm{sym}(\nabla_{\boldsymbol{y}}\boldsymbol{v})$, we have that $\boldsymbol{S} : \nabla_{\boldsymbol{y}}\boldsymbol{v} = \boldsymbol{S} : \boldsymbol{L}$, so:

$$\int_{\Omega_t} \rho \frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{v} \cdot \boldsymbol{v} + \boldsymbol{S} : \boldsymbol{L} \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\partial\Omega_t} \boldsymbol{v} \cdot \boldsymbol{S}\boldsymbol{n} \,\mathrm{d}A_{\boldsymbol{y}} + \int_{\Omega_t} \rho \boldsymbol{b} \cdot \boldsymbol{v} \,\mathrm{d}V_{\boldsymbol{y}}.$$

This leads us to the following result.

Proposition 6.5. *If $\boldsymbol{\varphi}$ is a motion, $\boldsymbol{S}$ is the Cauchy stress field, $\boldsymbol{L}$ is the rate of strain field, and $\Omega_t \subset B_t$, then*

$$\frac{\mathrm{D}}{\mathrm{D}t}K[\Omega_t] + \int_{\Omega_t} \boldsymbol{S} : \boldsymbol{L} \,\mathrm{d}V_{\boldsymbol{y}} = P[\Omega_t] \quad \textit{for all } t \in [0,T],$$

where $K[\Omega_t]$ is the kinetic energy of Ω_t, and $P[\Omega_t]$ is the power of external forces acting on Ω_t. The net rate of work done on $W[\Omega_t]$ is therefore

$$W[\Omega_t] = \int_{\Omega_t} \boldsymbol{S} : \boldsymbol{L} \,\mathrm{d}V_{\boldsymbol{y}}.$$

The quantity $\boldsymbol{S} : \boldsymbol{L}$ is called the *stress power* associated with a motion. It corresponds to the rate at which work is done by internal forces (i.e. stresses) acting at each point in a body.

6.3.5 First Law of Thermodynamics

Using the expressions derived in the previous section, we can invoke **Conservation of Energy** to deduce that the balance law for energy in $\Omega_t \subseteq B_t$ is

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t} \rho\phi \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega_t} \boldsymbol{S} : \boldsymbol{L} \,\mathrm{d}V_{\boldsymbol{y}} - \int_{\partial\Omega_t} \boldsymbol{q}\cdot\boldsymbol{n} \,\mathrm{d}A_{\boldsymbol{y}} + \int_{\Omega_t} \rho r \,\mathrm{d}V_{\boldsymbol{y}}.$$

Recall that:

- ρ is the mass density field,
- ϕ is the internal energy field per unit mass,
- $\boldsymbol{q}$ is the heat flux vector field,
- r is the heat supply field per unit mass,
- $\boldsymbol{L}$ is the rate of strain field, and
- $\boldsymbol{S}$ is the Cauchy stress field.

We can apply Proposition 6.2 along with the divergence theorem to deduce that

$$\int_{\Omega_t} \rho\frac{\mathrm{D}}{\mathrm{D}t}\phi \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega_t} \boldsymbol{S} : \boldsymbol{L} - \nabla_{\boldsymbol{y}}\cdot\boldsymbol{q} + \rho r \,\mathrm{d}V_{\boldsymbol{y}},$$

and then an application of the Localisation Theorem entails:

Proposition 6.6. *If $\boldsymbol{\varphi} : B\times[0,T] \to \mathbb{E}^3$ is a motion, then the **First Law of Thermodynamics** entails that*

$$\rho\frac{\mathrm{D}}{\mathrm{D}t}\phi = \boldsymbol{S} : \boldsymbol{L} - \nabla_{\boldsymbol{y}}\cdot\boldsymbol{q} + \rho r \quad \textit{for all } \boldsymbol{y}\in B_t, t\in[0,T].$$

6.3.6 Summary

We now summarise our Eulerian equations. Taking components with respect to a particular coordinate frame, there are 21 scalar unknowns:

$\varphi_i(\boldsymbol{x},t)$	3 components of motion,
$v_i(\boldsymbol{y},t)$	3 components of velocity,
$\rho(\boldsymbol{y},t)$	1 mass density field,
$S_{ij}(\boldsymbol{y},t)$	9 components of Cauchy stress,
$\theta(\boldsymbol{y},t)$	1 temperature field,
$q_i(\boldsymbol{y},t)$	3 components of heat flux, and
$\phi(\boldsymbol{y},t)$	1 internal energy field per unit mass.

Eulerian equations

From physical principles, we have deduced 11 equations:

$$
\begin{aligned}
[v_i]_m &= \frac{\partial \varphi_i}{\partial t} && \text{3 kinematic relations,} \\
\frac{\partial \rho}{\partial t} + (\rho v_i)_{,i} &= 0 && \text{1 conservation of mass equation,} \\
\rho \frac{\mathrm{D}}{\mathrm{D}t} v_i &= S_{ij,j} + \rho b_i && \text{3 balance of linear momentum equations,} \\
S_{ij} &= S_{ji} && \text{3 balance of angular momentum equations, and} \\
\rho \frac{\mathrm{D}}{\mathrm{D}t} \phi &= S_{ij} v_{i,j} - q_{i,i} + \rho r && \text{1 energy balance equation.}
\end{aligned}
$$

Notes:

- The motion map $\boldsymbol{\varphi}$ is often ignored within the Eulerian framework. In this case, we only care about the spatial velocity field $\boldsymbol{v}$, and 3 unknowns and equations disappear.
- Since the number of unknowns is greater than the number of equations, more equations are required to close the system. This leads to the idea of constitutive equations which relate the fields $(\boldsymbol{S}, \boldsymbol{q}, \phi)$ to the fields $(\rho, \boldsymbol{v}, \theta)$. Such relationships reflect the specific material properties of a body, and are often inherently much less certain than the physical laws we have used here.

6.4 Localised Lagrangian form of balance laws

Recall that the Lagrangian form of the local balance laws are expressed in terms of fields defined on the reference configuration, B, or in material coordinates. The PDEs we derive are often used in solid mechanics, where material points do not get 'mixed-up' in the way that they do in fluids.

6.4.1 Conservation of mass

We have already covered conservation of mass in Lagrangian form through the arguments applied in Section 6.3.1:

Proposition 6.7. *If $\boldsymbol{\varphi} : B \times [0,T] \to \mathbb{E}^3$ is a motion with associated deformation gradient $\boldsymbol{F}(\boldsymbol{x}, t)$, the spatial mass density field is $\rho(\boldsymbol{y}, t)$, and $\rho_0(\boldsymbol{x})$ is the mass density in the reference*

configuration B, then the conservation of mass entails that

$$\rho_m(\boldsymbol{x},t) = \frac{\rho_0(\boldsymbol{x})}{\det \boldsymbol{F}(\boldsymbol{x},t)}$$

for all $\boldsymbol{x} \in B$ and $t \in [0,T]$, where ρ_m is the material description of the mass density.

6.4.2 Balance of linear momentum

We recall from the **Laws of Inertia** that the balance of linear momentum for $\Omega_t \subseteq B_t$ entails that

$$\frac{\mathrm{D}}{\mathrm{D}t}\boldsymbol{l}[\Omega_t] = \boldsymbol{r}[\Omega_t].$$

In order to derive PDEs in material coordinates, we therefore need to express the quantities $\boldsymbol{l}[\Omega_t]$ and $\boldsymbol{r}[\Omega_t]$ in terms of those same material coordinates.

Changing variable, and recalling that $\Omega_t = \boldsymbol{\varphi}(\Omega,t)$, and applying Proposition 6.7, we find

$$\begin{aligned}\boldsymbol{l}[\Omega_t] &= \int_{\Omega_t} \rho(\boldsymbol{y},t)\boldsymbol{v}(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}} \\ &= \int_{\Omega} \rho(\boldsymbol{\varphi}(\boldsymbol{x},t),t)\boldsymbol{v}(\boldsymbol{\varphi}(\boldsymbol{x},t),t)\det \boldsymbol{F}(\boldsymbol{x},t)\,\mathrm{d}V_{\boldsymbol{x}} \\ &= \int_{\Omega} \rho_0(\boldsymbol{x})\frac{\partial}{\partial t}\boldsymbol{\varphi}(\boldsymbol{x},t)\,\mathrm{d}V_{\boldsymbol{x}}.\end{aligned}$$

A similar argument for $\boldsymbol{r}[\Omega_t]$ yields

$$\begin{aligned}\boldsymbol{r}[\Omega_t] &= \int_{\partial\Omega_t} \boldsymbol{S}(\boldsymbol{y},t)\boldsymbol{n}(\boldsymbol{y})\,\mathrm{d}A_{\boldsymbol{y}} + \int_{\Omega_t} \rho(\boldsymbol{y},t)\boldsymbol{b}(\boldsymbol{y},t)\,\mathrm{d}V_{\boldsymbol{y}} \\ &= \int_{\partial\Omega} \boldsymbol{S}(\boldsymbol{\varphi}(\boldsymbol{x},t),t)\mathrm{cof}\boldsymbol{F}(\boldsymbol{x},t)\boldsymbol{\nu}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}} + \int_{\Omega} \rho_m(\boldsymbol{x},t)\boldsymbol{b}(\boldsymbol{\varphi}(\boldsymbol{x},t),t)\det \boldsymbol{F}(\boldsymbol{x},t)\,\mathrm{d}V_{\boldsymbol{x}}.\end{aligned}$$

Inspecting the first integral, we see that if we want to apply the Divergence Theorem, we will end up with a divergence of $\boldsymbol{S}\mathrm{cof}\boldsymbol{F}$. This motivates the definition of the *(first) Piola-Kirchhoff stress tensor field*, $\boldsymbol{P} : B \times [0,T] \to \mathcal{V}^2$, to be

$$\boldsymbol{P}(\boldsymbol{x},t) := \boldsymbol{S}_m(\boldsymbol{x},t)\mathrm{cof}\boldsymbol{F}(\boldsymbol{x},t).$$

To interpret this stress tensor, note that $\boldsymbol{P}$ transforms the normal $\boldsymbol{\nu}$ to an infinitesimal surface in the reference configuration to the traction $\boldsymbol{t} = \boldsymbol{P}\boldsymbol{\nu}$ acting act the equivalent point in the deformed configuration.

For reference, we note that there is another measure of stress called the *second Piola-Kirchhoff stress tensor field*, $\boldsymbol{\Sigma} : B \times [0,T] \to \mathcal{V}^2$, defined to be

$$\boldsymbol{\Sigma}(\boldsymbol{x},t) := \boldsymbol{F}^{-1}(\boldsymbol{x},t)\boldsymbol{P}(\boldsymbol{x},t) = \boldsymbol{F}^{-1}(\boldsymbol{x},t)\boldsymbol{S}_m(\boldsymbol{x},t)\boldsymbol{F}^{-T}(\boldsymbol{x},t)\det(\boldsymbol{F}(\boldsymbol{x},t)).$$

This stress tensor field

We now proceed by using the above definition and Proposition 6.7 to write

$$\boldsymbol{r}[\Omega_t] = \int_{\partial\Omega} \boldsymbol{P}(\boldsymbol{x},t)\boldsymbol{\nu}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}} + \int_{\Omega_t} \rho_0(\boldsymbol{x})\boldsymbol{b}_m(\boldsymbol{x},t)\,\mathrm{d}V_{\boldsymbol{x}},$$

where $\boldsymbol{b}_m(\boldsymbol{x},t)$ is the material description of the body force field $\boldsymbol{b}$. The conservation of linear momentum therefore entails that

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega} \rho_0(\boldsymbol{x})\frac{\partial}{\partial t}\boldsymbol{\varphi}(\boldsymbol{x},t)\,\mathrm{d}V_{\boldsymbol{x}} = \int_{\partial\Omega} \boldsymbol{P}(\boldsymbol{x},t)\boldsymbol{\nu}(\boldsymbol{x})\,\mathrm{d}A_{\boldsymbol{x}} + \int_{\Omega_t} \rho_0(\boldsymbol{x})\boldsymbol{b}_m(\boldsymbol{x},t)\,\mathrm{d}V_{\boldsymbol{x}}.$$

Using the fact that ρ_0 is independent of time, and applying the Divergence Theorem, we obtain

$$\int_{\Omega} \rho_0 \frac{\partial^2\boldsymbol{\varphi}}{\partial t^2}\,\mathrm{d}V_{\boldsymbol{x}} = \int_{\Omega} \nabla_{\boldsymbol{x}}\cdot\boldsymbol{P} + \rho_0\boldsymbol{b}_m\,\mathrm{d}V_{\boldsymbol{x}}.$$

Now, since this holds for all $\Omega \subset B$ (as $\boldsymbol{\varphi}$ is a bijection), we have the following result.

Proposition 6.8. *If $\boldsymbol{\varphi} : B\times[0,T] \to \mathbb{E}^3$ is a motion and $\rho_0(\boldsymbol{x})$ is the mass density in the reference configuration B, then the balance of linear momentum entails that*

$$\rho_0\frac{\partial^2\boldsymbol{\varphi}}{\partial t^2} = \nabla_{\boldsymbol{x}}\cdot\boldsymbol{P} + \rho_0\boldsymbol{b}_m$$

for all $\boldsymbol{x}\in B$ and $t\in[0,T]$, where $\boldsymbol{P}$ is the Piola-Kirchhoff stress tensor and $\boldsymbol{b}_m$ is the material description of the spatial body force field $\boldsymbol{b}$.

6.4.3 Balance of angular momentum

We note that the balance of angular momentum in Eulerian form entails that $\boldsymbol{S} = \boldsymbol{S}^T$. Noting that $(\mathrm{cof}\boldsymbol{F})\boldsymbol{F}^T = (\det\boldsymbol{F})\boldsymbol{I}$, we deduce that

$$\boldsymbol{P}\boldsymbol{F}^T = \boldsymbol{S}\boldsymbol{F}^{-T}\boldsymbol{F}^T\det(\boldsymbol{F}) = \boldsymbol{S}^T\det(\boldsymbol{F}) = \boldsymbol{F}\boldsymbol{F}^{-1}\boldsymbol{S}^T\det(\boldsymbol{F}) = \boldsymbol{F}\boldsymbol{P}^T.$$

We therefore find:

If $\boldsymbol{\varphi} : B\times[0,T]\to\mathbb{E}^3$ is a motion with associated deformation gradient $\boldsymbol{F} : B\times[0,T]\to\mathcal{V}^2$, then the balance of linear momentum entails that

$$\boldsymbol{P}\boldsymbol{F}^T = \boldsymbol{F}\boldsymbol{P}^T.$$

for all $\boldsymbol{x}\in B$ and $t\in[0,T]$, where $\boldsymbol{P}$ is the Piola-Kirchhoff stress tensor.

6.4.4 Net rate of work

We now consider energies and power as considered when discussing the Laws of Thermodynamics. Recall that for $\Omega_t \subseteq B_t$, we have that the kinetic energy is

$$K[\Omega_t] = \int_{\Omega_t} \tfrac{1}{2}\rho|\boldsymbol{v}|^2 \,\mathrm{d}V_{\boldsymbol{y}},$$

which, by applying Proposition 6.7 and a change of variable, transforms to

$$K[\Omega_t] = \int_{\Omega} \tfrac{1}{2}\rho_0(\boldsymbol{x}) \left|\frac{\partial}{\partial t}\boldsymbol{\varphi}(\boldsymbol{x},t)\right|^2 \,\mathrm{d}V_{\boldsymbol{x}}.$$

Changing variable, the power of the external forces can be expressed as

$$\begin{aligned} P[\Omega_t] &= \int_{\Omega_t} \rho\boldsymbol{b}\cdot\boldsymbol{v}\,\mathrm{d}V_{\boldsymbol{y}} + \int_{\partial\Omega_t} \boldsymbol{v}\cdot(\boldsymbol{S}\boldsymbol{n})\,\mathrm{d}A_{\boldsymbol{y}} \\ &= \int_{\Omega} \rho_m\boldsymbol{b}_m\cdot\boldsymbol{v}_m \det\boldsymbol{F}\,\mathrm{d}V_{\boldsymbol{x}} + \int_{\partial\Omega} \boldsymbol{v}_m\cdot(\boldsymbol{S}_m\mathrm{cof}\boldsymbol{F}\boldsymbol{\nu})\,\mathrm{d}A_{\boldsymbol{x}} \\ &= \int_{\Omega} \rho_0\boldsymbol{b}_m\cdot\dot{\boldsymbol{\varphi}}\,\mathrm{d}V_{\boldsymbol{x}} + \int_{\partial\Omega} \dot{\boldsymbol{\varphi}}\cdot(\boldsymbol{P}\boldsymbol{\nu})\,\mathrm{d}A_{\boldsymbol{x}} \\ &= \int_{\Omega} \rho_0\boldsymbol{b}_m\cdot\dot{\boldsymbol{\varphi}}\,\mathrm{d}V_{\boldsymbol{x}} + \int_{\partial\Omega} (\boldsymbol{P}^T\dot{\boldsymbol{\varphi}})\cdot\boldsymbol{\nu}\,\mathrm{d}A_{\boldsymbol{x}}. \end{aligned}$$

The Divergence Theorem and product rule for the divergence of a tensor applied to a vector now gives

$$\begin{aligned} P[\Omega_t] &= \int_{\Omega} \rho_0\boldsymbol{b}_m\cdot\frac{\partial}{\partial t}\boldsymbol{\varphi} + \nabla_{\boldsymbol{x}}\cdot\left(\boldsymbol{P}^T\frac{\partial}{\partial t}\boldsymbol{\varphi}\right)\,\mathrm{d}V_{\boldsymbol{x}} \\ &= \int_{\Omega} (\nabla_{\boldsymbol{x}}\cdot\boldsymbol{P} + \rho_0\boldsymbol{b}_m)\cdot\frac{\partial}{\partial t}\boldsymbol{\varphi} + \boldsymbol{P} : \nabla_{\boldsymbol{x}}\frac{\partial}{\partial t}\boldsymbol{\varphi}\,\mathrm{d}V_{\boldsymbol{x}}. \end{aligned}$$

Differentiating and substituting for $\rho_0\frac{\partial^2\boldsymbol{\varphi}}{\partial t^2}$ using Proposition 6.3, we have

$$\frac{\mathrm{D}}{\mathrm{D}t}K[\Omega_t] = \int_{\Omega} \rho_0\frac{\partial}{\partial t}\boldsymbol{\varphi}\cdot\frac{\partial^2\boldsymbol{\varphi}}{\partial t^2}\,\mathrm{d}V_{\boldsymbol{x}} = \int_{\Omega} \frac{\partial}{\partial t}\boldsymbol{\varphi}\cdot(\nabla_{\boldsymbol{x}}\cdot\boldsymbol{S} + \rho_0\boldsymbol{b}_m)\,\mathrm{d}V_{\boldsymbol{x}}.$$

Equating this expression with the expression for the power $P[\Omega_t]$, we have shown the following result.

Proposition 6.9. *If $\boldsymbol{\varphi} : B\times[0,T] \to \mathbb{E}^3$ is a motion with associated deformation gradient $\boldsymbol{F} : B\times[0,T] \to \mathcal{V}^2$ and the Piola-Kirchhoff stress is $\boldsymbol{P}(\boldsymbol{x},t)$, then*

$$\frac{\mathrm{D}}{\mathrm{D}t}K[\Omega_t] + \int_{\Omega} \boldsymbol{P} : \frac{\partial}{\partial t}\boldsymbol{F}\,\mathrm{d}V_{\boldsymbol{x}} = P[\Omega_t]$$

for any $\Omega_t \subseteq B_t$ and $t\in[0,T]$. Put differently, the net rate of work done on Ω_t is

$$W[\Omega_t] = \int_{\Omega} \boldsymbol{P} : \frac{\partial}{\partial t}\boldsymbol{F}\,\mathrm{d}V_{\boldsymbol{x}} \quad \textit{for any } t\in[0,T].$$

We compare this result with the Eulerian version, and note that $\boldsymbol{S} : \boldsymbol{L}$ and $\boldsymbol{P} : \frac{\partial}{\partial t}\boldsymbol{F}$ are both stress power combinations. The former is per unit volume in the deformed configuration B_t, while the latter is per unit volume in the reference configuration B.

6.4.5 First Law of Thermodynamics

The **First Law of Thermodynamics** states that

$$\frac{\mathrm{D}}{\mathrm{D}t}U[\Omega_t] = Q[\Omega_t] + W[\Omega_t].$$

Recalling the definition of U, we can change variable to the reference configuration and use the conservation of mass in Lagrangian form to write

$$U[\Omega_t] = \int_{\Omega_t} \phi\rho \,\mathrm{d}V_{\boldsymbol{y}} = \int_{\Omega} \rho_0\phi_m \,\mathrm{d}V_{\boldsymbol{x}}.$$

For the net rate of heating, we have

$$Q[\Omega_t] = \int_{\Omega_t} \rho r \,\mathrm{d}V_{\boldsymbol{y}} - \int_{\partial\Omega_t} \boldsymbol{q}\cdot\boldsymbol{n}\,\mathrm{d}A_{\boldsymbol{y}} = \int_{\Omega} \rho_0 r_m \,\mathrm{d}V_{\boldsymbol{x}} - \int_{\partial\Omega} \boldsymbol{q}_m \cdot (\operatorname{cof}\boldsymbol{F}\boldsymbol{\nu})\,\mathrm{d}A_{\boldsymbol{x}}.$$

Defining the material heat flux vector field $\boldsymbol{\kappa} : B\times[0,T]\to\mathcal{V}$ via

$$\boldsymbol{\kappa}(\boldsymbol{x},t) = \big(\operatorname{cof}\boldsymbol{F}(\boldsymbol{x},t)\big)^T\boldsymbol{q}_m(\boldsymbol{x},t),$$

we have

$$Q[\Omega_t] = \int_{\Omega} \rho_0 r_m \,\mathrm{d}V_{\boldsymbol{x}} - \int_{\partial\Omega} \boldsymbol{\kappa}\cdot\boldsymbol{\nu}\,\mathrm{d}A_{\boldsymbol{x}}.$$

Applying the Divergence Theorem, we have:

Proposition 6.10. *If $\boldsymbol{\varphi} : B\times[0,T]\to\mathbb{E}^3$ is a motion with material internal energy field ϕ_m, material heat flux vector field $\boldsymbol{\kappa}$ and material heat supply field r_m, then*

$$\rho_0\frac{\partial}{\partial t}\phi_m = \boldsymbol{P} : \frac{\partial}{\partial t}\boldsymbol{F} - \nabla_{\boldsymbol{x}}\cdot\boldsymbol{\kappa} + \rho_0 r_m$$

for all $\boldsymbol{x}\in B$ and $t\in[0,T]$.

6.4.6 Summary

In the Lagrangian description of the motion of a general continuum body, then if we take component with respect to a given Cartesian frame, there are 17 unknown fields:

$\varphi_i(\boldsymbol{x},t)$	3 components of motion,
$P_{ij}(\boldsymbol{x},t)$	9 components of Piola-Kirchhoff stress,
$\theta_m(\boldsymbol{x},t)$	1 temperature field,
$\kappa_i(\boldsymbol{x},t)$	3 components of heat flux, and
$\phi_m(\boldsymbol{x},t)$	1 internal energy field per unit mass.

Lagrangian equations

From physical principles, we have deduced 7 equations:

$$\rho_0 \frac{\partial^2 \varphi_i}{\partial t^2} = P_{ij,j} + \rho_0 [b_i]_m \qquad \text{3 balance of linear momentum equations,}$$

$$P_{ij} F_{jk} = F_{ik} P_{jk} \qquad \text{3 balance of angular momentum equations, and}$$

$$\rho_m \frac{\partial}{\partial t} \phi_m = P_{ij} \frac{\partial}{\partial t} F_{ij} - \kappa_{i,i} + \rho_0 r_m \qquad \text{1 energy balance equation.}$$

Notes:

- In contrast with the Eulerian formulation, the mass density field is a known quantity, since we assume it is prescribed to be ρ_0.
- We note that the velocity field is absent here, and this is because it can be obtained directly by differentiating $\boldsymbol{\varphi}$ in time.
- Once more, the number of unknowns exceeds the number of equations by 11, and so we need additional constitutive relations to relate $(\boldsymbol{P}, \boldsymbol{\kappa}, \phi_m, \eta_m)$ to $(\boldsymbol{\varphi}, \theta_m)$.

6.5 Frame indifference

Frame indifference is a physical principle which is assumed in classical (i.e. non-relativistic) mechanics. In this case, we assume that certain fields are independent of the observer, and this principle results in limitations on the constitutive equations which are valid for a material body.

6.5.1 Superposed rigid motions

In order to describe two observers viewing the same motion, we introduce the notion of a superposed rigid motion. We will say that two motions $\boldsymbol{\varphi}, \boldsymbol{\varphi}^* : B \times [0,T] \to \mathbb{E}^3$ are related by a *superposed rigid motion* if it holds that

$$\boldsymbol{\varphi}^*(\boldsymbol{x}, t) = \boldsymbol{Q}(t)\boldsymbol{\varphi}(\boldsymbol{x}, t) + \boldsymbol{c}(t) \quad \text{for all } \boldsymbol{x} \in B, t \in [0,T],$$

where $\boldsymbol{Q} : [0,T] \to \mathcal{V}^2$ is a rotation-valued mapping, and $\boldsymbol{c} : [0,T] \to \mathcal{V}$.

In this definition, $\boldsymbol{Q}(t)$ and $\boldsymbol{c}(t)$ describe the movement of one observer with respect to the other: $\boldsymbol{Q}(t)$ describes the relative rotation, and $\boldsymbol{c}(t)$ describes the relative translation. Now, consider each of the various fields we have discussed, and consider their representation in the frames of the two observers. To denote those fields observed which in the frame where the motion is $\boldsymbol{\varphi}^*$, we use a star. Applying the definition above, we then find the following:

Proposition 6.11. *The relationship between the material fields* $(\boldsymbol{F}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{C})$ *and* $(\boldsymbol{F}^*, \boldsymbol{R}^*, \boldsymbol{U}^*, \boldsymbol{V}^*, \boldsymbol{C}^*)$ *is given by*

$$\begin{aligned} \boldsymbol{F}^* &= \boldsymbol{Q}\boldsymbol{F}, \quad \boldsymbol{R}^* = \boldsymbol{Q}\boldsymbol{R}, \\ \boldsymbol{U}^* &= \boldsymbol{U}, \quad \boldsymbol{V}^* = \boldsymbol{Q}\boldsymbol{V}\boldsymbol{Q}^T, \quad \boldsymbol{C} = \boldsymbol{C}^*, \end{aligned}$$

where these equalities hold for all $\boldsymbol{x} \in B$ *and* $t \in [0, T]$.

6.5.2 The Axiom of Frame-indifference

Let B_t be the current configuration of B under motion $\boldsymbol{\varphi}$ at time t, and let B_t^* be the current configuration of B under motion $\boldsymbol{\varphi}^*$ at time t.

The scalar field ϕ, vector field $\boldsymbol{w}$, and tensor field $\boldsymbol{T}$ are *frame-indifferent* if for any superposed rigid motion $\boldsymbol{g}(\cdot, t) : B_t \to B_t^*$, we have

$$\begin{aligned} \phi^*(\boldsymbol{y}^*, t) &= \phi(\boldsymbol{y}, t), \\ \boldsymbol{w}^*(\boldsymbol{y}^*, t) &= \boldsymbol{Q}(t)\boldsymbol{w}(\boldsymbol{y}, t), \\ \boldsymbol{T}^*(\boldsymbol{y}^*, t) &= \boldsymbol{Q}(t)\boldsymbol{T}(\boldsymbol{y}, t)\boldsymbol{Q}(t)^T, \end{aligned}$$

for all $\boldsymbol{y} \in B_t$ and $t \in [0, T]$, where $\boldsymbol{y}^* = \boldsymbol{g}(\boldsymbol{y}, t)$.

This notion of frame-indifference captures the idea that some physical quantities are inherent to the body, and are independent of the observer. Given that they are referring to the same point in the body, two observers should agree on the mass density and the temperature at a given material point. Up to a change of basis, they should also agree on the heat flux and stress tensors. Some fields will not satisfy this assumption however: the velocity of a body will depend upon the frame, as will the acceleration.

It is natural to make the following assumption of frame-indifference for the fields we have considered:

> **Frame Indifference**
>
> The spatial mass density ρ, temperature θ, internal energy per unit mass ϕ, Cauchy stress $\boldsymbol{S}$, and heat flux vector $\boldsymbol{q}$ are all frame-indifferent fields.

This assumption places strict limitations on the constitutive equations which it is possible to use to model the material properties of a body. For example, if we express the internal energy, heat flux and Cauchy stress as functions

$$\phi = \widehat{\phi}(\rho, \theta, \boldsymbol{F}), \quad \boldsymbol{q} = \widehat{\boldsymbol{q}}(\rho, \theta, \boldsymbol{F}), \quad \text{and} \quad \boldsymbol{S} = \widehat{\boldsymbol{S}}(\rho, \theta, \boldsymbol{F}),$$

then in order to comply with the principle of frame indifference, we must have

$$\begin{aligned}\widehat{\phi}(\rho^*, \theta^*, \boldsymbol{F}^*) &= \widehat{\phi}(\rho, \theta, \boldsymbol{F}),\\ \widehat{\boldsymbol{q}}(\rho^*, \theta^*, \boldsymbol{F}^*) &= \boldsymbol{Q}\widehat{\boldsymbol{q}}(\rho, \theta, \boldsymbol{F}),\\ \widehat{\boldsymbol{S}}(\rho^*, \theta^*, \boldsymbol{F}^*) &= \boldsymbol{Q}\widehat{\boldsymbol{S}}(\rho, \theta, \boldsymbol{F})\boldsymbol{Q}^T.\end{aligned}$$

Substituting for the starred fields on the LHS, we have

$$\begin{aligned}\widehat{\phi}(\rho, \theta, \boldsymbol{QF}) &= \widehat{\phi}(\rho, \theta, \boldsymbol{F}),\\ \widehat{\boldsymbol{q}}(\rho, \theta, \boldsymbol{QF}) &= \boldsymbol{Q}\widehat{\boldsymbol{q}}(\rho, \theta, \boldsymbol{F}),\\ \widehat{\boldsymbol{S}}(\rho, \theta, \boldsymbol{QF}) &= \boldsymbol{Q}\widehat{\boldsymbol{S}}(\rho, \theta, \boldsymbol{F})\boldsymbol{Q}^T,\end{aligned}$$

which must hold for all rotations $\boldsymbol{Q}$ and all admissible values of ρ, θ and $\boldsymbol{F}$.

It is important to note that linearised theories will typically violate frame-indifference, since the relationships we have derived above are nonlinear. It is therefore important to assess whether any linearisation could be breaking down when using a linearised model to make predictions.

7 Solid Mechanics

In this chapter, we apply the principles we have developed to the study of solid materials, focusing specifically on elasticity theory. As mentioned, in solid mechanics the canonical choice is to use the *Lagrangian* or *material* formulation of the balance laws. In particular, we focus on the momentum balance equations, which are:

$$\rho_0 \frac{\partial^2 \boldsymbol{\varphi}}{\partial t^2} = \nabla_{\boldsymbol{x}} \cdot \boldsymbol{P} + \rho_0 \boldsymbol{b}_m, \quad \text{and} \quad \boldsymbol{P}\boldsymbol{F}^T = \boldsymbol{F}\boldsymbol{P}^T. \tag{7.1}$$

We note that we can immediately deduce *force balance* equations for statice problems from these equations by neglecting the momentum change term, in which case they become

$$\nabla_{\boldsymbol{x}} \cdot \boldsymbol{P} + \rho_0 \boldsymbol{b}_m = \boldsymbol{0}, \quad \text{and} \quad \boldsymbol{P}\boldsymbol{F}^T = \boldsymbol{F}\boldsymbol{P}^T. \tag{7.2}$$

To make further progress with either of these sets of equations, we must introduce a *constitutive law* for the stress in terms of strain, often referred to as *stress–strain* relation.

7.1 Elastic solids

Elasticity refers to the property of solid objects to stretch and then return to their original shape, releasing energy in an (approximately) reversible process. To model such cases, we make the following definition.

Elastic solid

A continuum body B is said to be an **elastic solid** if the following 2 properties are satisfied:

1. The Cauchy stress field can be expressed as a function of the deformation gradient, i.e.
$$\boldsymbol{S}_m(\boldsymbol{x}, t) = \widetilde{\boldsymbol{S}}(\boldsymbol{F}(\boldsymbol{x}, t), \boldsymbol{x})$$
for any $\boldsymbol{x} \in B$ and $t \in [0, T]$.
2. The function $\widetilde{\boldsymbol{S}}$ is always symmetric for valid deformation gradients, i.e.
$$\widetilde{\boldsymbol{S}}(\boldsymbol{F}, \boldsymbol{x})^T = \widetilde{\boldsymbol{S}}(\boldsymbol{F}, \boldsymbol{x})$$
for any $\boldsymbol{x} \in B$ and any $\boldsymbol{F} \in \mathcal{V}^2$ with $\det \boldsymbol{F} > 0$.

Some notes on the above definition:

- The function $\widetilde{\boldsymbol{S}}$ is called the **elastic stress response function** of the body.
- The dependence on the point $\boldsymbol{x}$ in the stress response function allows us in general to represent the fact that the body may be made up of different materials in different places. The simplest cases have no explicit dependence on $\boldsymbol{x}$, in which case we say the body is **homogeneous**.
- Using the relationship between the Cauchy and Piola–Kirchhoff stresses, we can choose to prescribe either the Cauchy stress by defining a function $\widetilde{\boldsymbol{S}}$ as above, or by prescribing the Piola-Kirchhoff stress by defining a function $\widetilde{\boldsymbol{P}}(\boldsymbol{F})$. We can move between the two approaches using the relation:

$$\boldsymbol{P} = \widetilde{\boldsymbol{S}}(\boldsymbol{F})\mathrm{cof}\boldsymbol{F} = \det \boldsymbol{F}\, \widetilde{\boldsymbol{S}}(\boldsymbol{F})\boldsymbol{F}^{-T} \quad \text{or} \quad \boldsymbol{S} = \frac{\widetilde{\boldsymbol{P}}(\boldsymbol{F})\boldsymbol{F}^T}{\det \boldsymbol{F}}.$$

7.2 Hyperelasticity and conservation of energy

Next, we discuss conservation of energy. If we assume that the body is **isothermal**, so it is held at a constant temperature $\theta = \theta_0$ and there is no heat flow, then we neglect all heat related terms in the energy balance equation.

In particular, in the Lagrangian formulation, we neglect the heat flux $\boldsymbol{\kappa} = \boldsymbol{0}$ and the heat supply field $r = 0$, and the local equation for conservation of energy in Lagrangian form is

$$\rho_0 \frac{\partial \phi_m}{\partial t} = \boldsymbol{P} : \frac{\partial \boldsymbol{F}}{\partial t}.$$

This encodes the fact that the internal energy per unit mass only changes due to internal mechanical work, and no energy is lost as heat.

One way to guarantee that this equation holds is to assume that the Piola-Kirchhoff stress function can be expressed as the derivative of a function $W : \mathcal{V}^2 \to \mathbb{R}$ called the **strain energy density**, i.e.

$$\boldsymbol{P}(\boldsymbol{F}) = DW(\boldsymbol{F}) \quad \text{or in components} \quad P_{ij} = \frac{\partial W}{\partial F_{ij}}(\boldsymbol{F}).$$

Thanks to the chain rule, we find that

$$\rho_0 \frac{\partial \phi_m}{\partial t} = DW(\boldsymbol{F}) : \frac{\partial \boldsymbol{F}}{\partial t} = \frac{\mathrm{D}}{\mathrm{D}t} W(\boldsymbol{F}).$$

and so we find that (up to a constant) $\rho_0 \phi_m = W(\boldsymbol{F})$, where we are implicitly suppressing the dependence on the temperature θ_0. This choice leads to the following definition.

Hyperelastic material

A **hyperelastic material** is an elastic material for which:

1. The Piola–Kirchhoff stress response function satisfies
$$\widetilde{\boldsymbol{P}}(\boldsymbol{F}) = DW(\boldsymbol{F})$$
where $W : \mathcal{V}^2 \to \mathbb{R}$ is the **strain energy density**; and
2. For any $\boldsymbol{F} \in \mathcal{V}^2$ with $\det \boldsymbol{F} > 0$, we have that conservation of angular momentum is guaranteed, i.e.
$$DW(\boldsymbol{F})\boldsymbol{F}^T = \boldsymbol{F}DW(\boldsymbol{F})^T.$$

We will see that the existence of a strain energy density leads to various interesting mathematical consequences.

7.3 Frame indifference

Above, we have seen that conservation of energy suggests a structure for the constitutive laws we can consider: likewise, we now show that the principle of frame indifference also constrains the possible stress reponse functions. In particular, we have the following result.

Proposition 7.1. *If the following symmetry holds for the strain energy density:*
$$W(\boldsymbol{Q}\boldsymbol{F}) = W(\boldsymbol{F})$$
for any $\boldsymbol{F} \in \mathcal{V}^2$ with $\det \boldsymbol{F} > 0$ and any rotation tensor $\boldsymbol{Q} \in \mathcal{V}^2$, then the internal energy and stress tensors are automatically frame indifferent.

Proof. Clearly, W is a frame indifferent field by the definition given at the end of Section 6.5, since
$$W^* = W(\boldsymbol{F}^*) = W(\boldsymbol{Q}\boldsymbol{F}) = W,$$
and so defining the internal energy per unit mass as $\phi_m := W/\rho_0$, it must also be frame indifferent.

Next, we check that the assumption that $\boldsymbol{P} = DW$ means that $\boldsymbol{S}$ also transforms appropriately. Note that
$$\boldsymbol{S}_m = \frac{\boldsymbol{P}\boldsymbol{F}^T}{\det \boldsymbol{F}},$$

and we need to show that $\boldsymbol{S}_m^* = \boldsymbol{Q}\boldsymbol{S}_m\boldsymbol{Q}^T$ in order to verify the requirement of frame indifference. Using the chain rule, we have

$$\frac{\partial W}{\partial F_{ij}} = \frac{\partial W}{\partial F_{kl}^*}\frac{\partial F_{kl}^*}{\partial F_{ij}} = \frac{\partial W}{\partial F_{kl}^*}\frac{\partial}{\partial F_{ij}}\Big(Q_{ka}F_{al}\Big) = \frac{\partial W}{\partial F_{kl}^*}Q_{ka}\delta_{ai}\delta_{jl} = \frac{\partial W}{\partial F_{kj}^*}Q_{ki}.$$

Interpreting the above relationship in tensorial form, we have

$$DW(\boldsymbol{F}) = \boldsymbol{Q}^T DW(\boldsymbol{F}^*) = \boldsymbol{Q}^T DW(\boldsymbol{Q}\boldsymbol{F}).$$

Multiplying on the right by $(\operatorname{cof}\boldsymbol{F})^{-1}$, then $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$ and using the fact that $\det \boldsymbol{Q} = 1$, we have

$$\boldsymbol{S}_m = \frac{DW(\boldsymbol{F})\boldsymbol{F}^T}{\det \boldsymbol{F}} = \frac{\boldsymbol{Q}^T DW(\boldsymbol{Q}\boldsymbol{F})\boldsymbol{F}^T}{\det \boldsymbol{F}} = \frac{\boldsymbol{Q}^T DW(\boldsymbol{Q}\boldsymbol{F})(\boldsymbol{Q}\boldsymbol{F})^T\boldsymbol{Q}}{\det(\boldsymbol{Q}\boldsymbol{F})} = \boldsymbol{Q}^T\boldsymbol{S}_m^*\boldsymbol{Q}.$$

This shows that $\boldsymbol{S}$ transforms correctly, and so the material is automatically frame indifferent. □

7.4 Initial and Boundary conditions

Given a particular constitutive relationship between the Piola-Kirchhoff stress $\boldsymbol{P}$ and the deformation gradient $\boldsymbol{F}$, we have to solve either the evolution problem Equation 7.1 or the force balance problem Equation 7.2. In the former case we need both initial and boundary conditions, while in the second we need just boundary conditions.

7.4.1 Initial conditions

Generally speaking, the initial conditions chosen are based on assumptions about the position and velocity of the material points at the starting point. They usually take the form

$$\boldsymbol{\varphi}(\boldsymbol{x}, 0) = \boldsymbol{\varphi}_0(\boldsymbol{x}) \quad \text{and} \quad \frac{\partial \boldsymbol{\varphi}}{\partial t}(\boldsymbol{x}, 0) = \boldsymbol{v}_0(\boldsymbol{x}),$$

where $\boldsymbol{\varphi}_0$ and $\boldsymbol{v}_0$ are respectively the initial deformation and initial velocity field defined for any $\boldsymbol{x} \in B_0$. In simple cases we might take $\boldsymbol{\varphi}_0(\boldsymbol{x}) = \boldsymbol{x}$, if the body is initially relaxed.

7.4.2 Boundary conditions

The standard boundary conditions for either dynamical or static problems in elasticity are usually of two possible forms; either the position of the boundary points is prescribed, known as **Dirichlet boundary conditions**, or the stress on the boundary is prescribed, known as **Neumann boundary conditions**. Typically we split the boundary of the

body up into two parts on which these conditions are prescribed, Γ_D and Γ_N, and the conditions become:

$$\begin{aligned} \boldsymbol{\varphi} &= \boldsymbol{g} \quad \text{on } \Gamma_D \subset \partial B, \\ \boldsymbol{P}\boldsymbol{\nu} &= \boldsymbol{t} \quad \text{on } \Gamma_N \subset \partial B, \end{aligned}$$

where $\boldsymbol{g}$ and $\boldsymbol{t}$ may depend on time (as may the sets on which they are imposed).

More complex choices of boundary condition do exist, normally arising from frictional contact between the elastic body and some other object.

7.4.3 Full PDE problems

In full, we are now able to state the general PDE problems of elasticity.

Elastodynamics problem

Find $\boldsymbol{\varphi} : B \times [0,T] \to \mathbb{E}^3$ satisfying the following initial-boundary value problem:

$$\begin{aligned} \rho_0 \frac{\partial^2 \boldsymbol{\varphi}}{\partial t^2} &= \nabla_{\boldsymbol{x}} \cdot \widetilde{\boldsymbol{P}}(\boldsymbol{F}) + \rho_0 \boldsymbol{b}_m && \text{in } B \times [0,T], \\ \boldsymbol{\varphi}(\cdot, 0) &= \boldsymbol{\varphi}_0 && \text{in } B, \\ \frac{\partial \boldsymbol{\varphi}}{\partial t}(\cdot, 0) &= \boldsymbol{v}_0 && \text{in } B, \\ \boldsymbol{\varphi} &= \boldsymbol{g} && \text{on } \Gamma_D \times [0,T], \\ \widetilde{\boldsymbol{P}}(\boldsymbol{F})\boldsymbol{\nu} &= \boldsymbol{t} && \text{on } \Gamma_N \times [0,T]. \end{aligned}$$

- This problem is a nonlinear system of 3 coupled PDEs for the 3 components of $\boldsymbol{\varphi}$ which is second-order in both space and time.
- Existence and uniqueness can be proved under assumptions on the stress response function $\widetilde{\boldsymbol{P}}$, the smoothness of the boundaries B and of the data (i.e. the initial and boundary conditions), but this is beyond the scope of this module.
- Typically, solutions of the equations oscillate in time, and we think of this as a **nonlinear hyperbolic** PDE system. Wave-like solutions make physical sense as elastic waves are exactly how sound travels through solid materials.
- Given the complexity of these equations, explicit solutions to these equations are rare, so numerical approximate solutions must be sought. The most common approach is to use Finite Elements.
- A common assumption for practical purposes is that the body force per unit mass $\boldsymbol{b}_m$ arises due to a uniform gravitational field, in which case we may take $\boldsymbol{b} = -g\boldsymbol{e}_3$. Since this function is independent of position, we find that $\boldsymbol{b}_m = -g\boldsymbol{e}_3$.

Elastostatics problem

Find $\boldsymbol{\varphi} : B \to \mathbb{E}^3$ satisfying the boundary value problem:

$$\begin{aligned} \nabla_{\boldsymbol{x}} \cdot \widetilde{\boldsymbol{P}}(\boldsymbol{F}) + \rho_0 \boldsymbol{b}_m &= \boldsymbol{0} \quad \text{in } B \times [0,T], \\ \boldsymbol{\varphi} &= \boldsymbol{g} \quad \text{on } \Gamma_D, \\ \widetilde{\boldsymbol{P}}(\boldsymbol{F})\boldsymbol{\nu} &= \boldsymbol{t} \quad \text{on } \Gamma_N. \end{aligned}$$

- This problem is a nonlinear system of 3 coupled PDEs for the 3 components of $\boldsymbol{\varphi}$ which is second-order in space.
- As for the elastodynamics problem above, existence of solutions can be proved under smoothness assumptions on the body and the data, which is again beyond the scope of this module.
- Once again, numerical solution is usually necessary, and Finite Elements can be used to approximate solutions.
- An important caveat for the elastostatics equations is that non-uniqueness can occur! A prime example is elastic buckling.
- When the material is hyperelastic, the above equation formulation can be show to be the Euler-Lagrange equation of a variational problem. In this case, we seek to minimise the total energy in the system, expressed as

  $$I(\boldsymbol{\varphi}) := \int_B W(\nabla_{\boldsymbol{x}} \boldsymbol{\varphi}) - \rho_0 \boldsymbol{b}_m \cdot \boldsymbol{\varphi} \, \mathrm{d}V_{\boldsymbol{x}},$$

  over all $\boldsymbol{\varphi}$ which satisfy to the boundary conditions. It turns out this is a powerful way to prove existence of solutions to the problem using techniques from the Calculus of Variations.

7.5 Example hyperelastic material laws

In practice, almost all sensible elastic material laws are hyperelastic. The following are all examples of hyperelastic material laws that are used to describe rubber-like materials:

- A **St Venant–Kirchhoff** material has strain energy density

  $$W(\boldsymbol{F}) = \frac{\lambda}{2}(\mathrm{tr}\boldsymbol{E})^2 + \mu \mathrm{tr}(\boldsymbol{E}^2)$$

  where $\boldsymbol{E} = \frac{1}{2}(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I})$ is the Lagrange strain and $\lambda, \mu > 0$ are material parameters.
- A **Neo-Hookean** material has a strain energy density of the form

  $$\begin{aligned} W(\boldsymbol{F}) &:= a\,|\boldsymbol{F}|^2 + \Gamma(\det \boldsymbol{F}) \\ \text{where} \qquad \Gamma(J) &:= c\,J^2 - d \log J, \end{aligned}$$

  where $a, c, d > 0$ are material parameters.

- A **Mooney-Rivlin** material has strain energy density

$$W(\boldsymbol{F}) = a\,|\boldsymbol{F}|^2 + b\,|\mathrm{cof}\,\boldsymbol{F}|^2 + \Gamma(\det \boldsymbol{F}),$$

  where Γ is as for the neo-Hookean material, and $a, b, c, d > 0$ are material parameters.

Note that the neo-Hookean model is just the Mooney-Rivlin model with $b = 0$, so is simply a special case. Note also that all material parameters involved in these equations have units of energy per unit volume, or equivalently of stress (force per unit area).

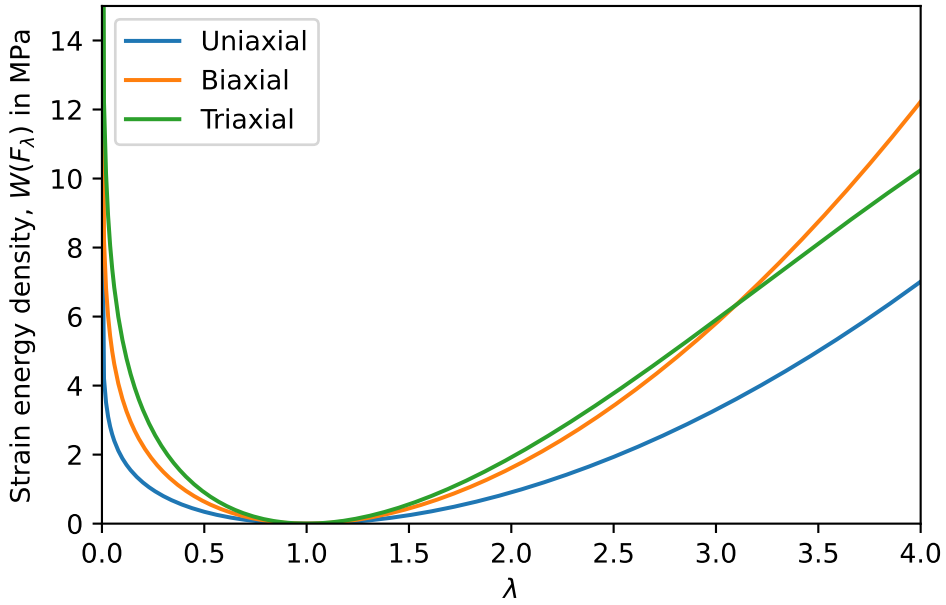


Figure 7.1: The variation of the Neo-Hookean energy density as a function of various forms of stretch. Here, a=0.6MPa and c=0.6MPa, which are values representative of rubber.

Figure 7.1 shows strain energies for a neo-Hookean material where the parameters used are indicative of those which would be used to represent rubber, showing the different results for uniaxial, biaxial and triaxial stretches, i.e. where we use the deformation gradient:

$$\boldsymbol{F}_\lambda = \begin{cases} \lambda \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_3 \otimes \boldsymbol{e}_3 & \text{in the uniaxial case,} \\ \lambda \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \lambda \boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_3 \otimes \boldsymbol{e}_3 & \text{in the biaxial case, and} \\ \lambda \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \lambda \boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \lambda \boldsymbol{e}_3 \otimes \boldsymbol{e}_3 & \text{in the triaxial case.} \end{cases}$$

Note that in each case, the parameters are chosen to guarantee that $\boldsymbol{F}_1 = \boldsymbol{I}$ is a minimum of the energy.

7.6 Material symmetry

Note that elastic materials with a frame indifferent stress response function must satisfy the symmetry

$$\boldsymbol{Q}\widetilde{\boldsymbol{S}}(\boldsymbol{F})\boldsymbol{Q}^T = \widetilde{\boldsymbol{S}}(\boldsymbol{Q}\boldsymbol{F}).$$

Here, we can think of $\boldsymbol{Q}$ as being a change of basis *after* having applied the deformation gradient $\boldsymbol{F}$, which corresponds to a change of frame for observing the material.

If on the other hand we require that

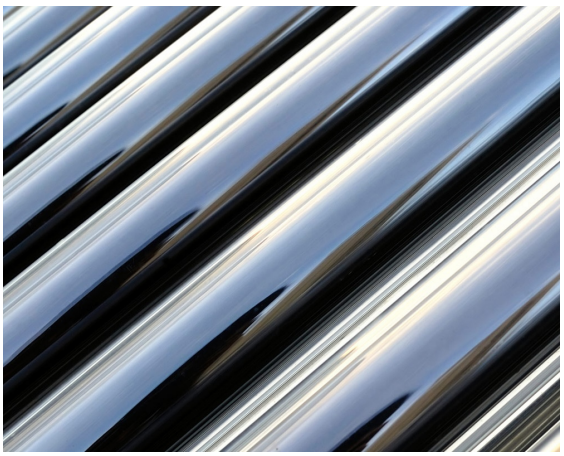
$$\widetilde{\boldsymbol{S}}(\boldsymbol{F}\boldsymbol{R}) = \widetilde{\boldsymbol{S}}(\boldsymbol{F})$$

for $\boldsymbol{R}$ taken from some subgroup of all rotations, we can encode material *symmetry*. We can think of the rotation $\boldsymbol{R}$ as acting *before* the strain $\boldsymbol{F}$ on an infinitesimal piece of the material.

Some examples of when we might wish to assume some material symmetries follow.

- Softwood is often made up of long straight fibres, see Figure 7.2a. Rotating a small piece of such a material about the axis of the fibres will keep the material properties similar. On the other hand, changing the orientation of the fibres (the grain direction) will affect the properties significantly. Grain orientation is an important consideration when using wood in practice.
- Polycrystalline materials and polymeric materials like rubber often behave *isotropically*, i.e. they have material properties which are invariant under any rotation, see Figure 7.2b. In this case, a small sample of the material behaves pretty much identically, no matter which direction it is rotated.

(a) A pine plank.



(b) Steel bars.

Figure 7.2: Examples of a materials with different symmetries.

We will focus on isotropic materials, since this is the strongest symmetry assumption we can make, and so is the most restrictive:

Isotropy

An elastic body is said to be isotropic if the stress response function satisfies the symmetry

$$\widetilde{\boldsymbol{S}}(\boldsymbol{F}\boldsymbol{Q}) = \widetilde{\boldsymbol{S}}(\boldsymbol{F})$$

for all rotations $\boldsymbol{Q}$ and for any admissible deformation gradient $\boldsymbol{F}$.

As we might expect, isotropy imposes significant restrictions on the possible material laws we can expect, and the following result gives a complete characterisation in the hyperelastic case.

Proposition 7.2. *The following 3 conditions are equivalent:*

1. *A homogeneous hyperelastic body is frame indifferent and isotropic.*
2. *The stored energy density can be expressed as a permutation invariant function of the 3 principal stretches λ_i, i.e.*
$$W(\boldsymbol{F}) = \Phi(\lambda_1, \lambda_2, \lambda_3),$$
for some function $\Phi : (0, +\infty)^3 \to \mathbb{R}$.
3. *The stored energy density can be expressed as a function of the three principal invariants of $\boldsymbol{U}$ only, i.e.*
$$W(\boldsymbol{F}) = \widetilde{\Phi}\Big(I_1(\boldsymbol{U}), I_2(\boldsymbol{U}), I_3(\boldsymbol{U})\Big)$$
for some function $\widetilde{\Phi} : (0, +\infty)^3 \to \mathbb{R}$.

Proof. We first show $1 \Rightarrow 2$. In combination, frame indifference and isotropy mean that

$$W(\boldsymbol{F}) = W(\boldsymbol{Q}_1\boldsymbol{F}\boldsymbol{Q}_2)$$

for any two rotations $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$. Using the polar decomposition of $\boldsymbol{F}$, we write

$$W(\boldsymbol{F}) = W(\boldsymbol{R}\boldsymbol{U}) = W(\boldsymbol{U}),$$

since $\boldsymbol{R}$ is a rotation. Then, as $\boldsymbol{U}$ is a symmetric positive definite tensor, we can diagonalise it, implying that

$$W(\boldsymbol{U}) = W(\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T) = W(\boldsymbol{D})$$

where $\boldsymbol{D}$ is a tensor that has components only on the diagonal in a given coordinate frame. These components are the eigenvalues of $\boldsymbol{U}$, which we recall are the principal stretches, and so as a consequence, we can write

$$W(\boldsymbol{F}) = \Phi(\lambda_1, \lambda_2, \lambda_3),$$

where Φ is some function. This function must be permutation invariant since we can introduce rotations which permute the diagonal components of $\boldsymbol{D}$; for example, note that

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda_2 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

and this proves statement 1.

Now we show 2 $\Rightarrow$ 3. We note that the principal stretches are roots of the characteristic polynomial

$$0 = \det(\boldsymbol{U} - \lambda \boldsymbol{I}) = -\lambda^3 + I_1(\boldsymbol{U})\lambda^2 - I_2(\boldsymbol{U})\lambda + I_3(\boldsymbol{U}),$$

and the roots of a polynomial are continuous functions of its coefficients. As such, there is some continuous function that maps the triple of invariants to the triple of stretches (effectively, this is the 'cubic formula'). Composing this function with Φ, it follows that the function $\widetilde{\Phi}$ exists.

Finally, we show 3 $\Rightarrow$ 1. In this case, we begin by noting that $\boldsymbol{U} = \sqrt{\boldsymbol{F}^T\boldsymbol{F}}$, and so each invariant of $\boldsymbol{U}$ is insensitive to left multiplication by a rigid body rotation. Similarly, we note that $(\boldsymbol{FQ})^T(\boldsymbol{FQ}) = \boldsymbol{Q}^T\boldsymbol{F}^T\boldsymbol{FQ}$, and it can be checked directly that

$$\sqrt{(\boldsymbol{FQ})^T(\boldsymbol{FQ})} = \boldsymbol{Q}^T\boldsymbol{UQ},$$

since positive definite square roots are unique. Next, we can use the permutation invariance of the trace to see that

$$\mathrm{tr}(\boldsymbol{U}) = \mathrm{tr}(\boldsymbol{QUQ}^T) \quad \text{and} \quad \mathrm{tr}(\boldsymbol{U}^2) = \mathrm{tr}(\boldsymbol{QU}^2\boldsymbol{Q}^T),$$

and similarly $\det(\boldsymbol{U}) = \det(\boldsymbol{QUQ}^T)$. It follows that all terms in the invariants transform correctly under right multiplication, and therefore the strain energy density is also isotropic. □

7.6.1 Isotropic examples

We have now very significantly narrowed the range of possible stored energy density functions. We will show that the example hyperelastic laws given above are all frame indifferent and isotropic stored energy densities.

- For a **St Venant-Kirchhoff** material, the stored energy density was written using the trace of $\boldsymbol{E} = \frac{1}{2}(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I})$ and $\boldsymbol{E}^2$. We show these are expressible in a permutation invariant way in terms of the principal stretches. First:

  $$2\mathrm{tr}\boldsymbol{E} = \mathrm{tr}(\boldsymbol{F}^T\boldsymbol{F}) - \mathrm{tr}(\boldsymbol{I}) = \mathrm{tr}(\boldsymbol{C}) - 3 = \mathrm{tr}(\boldsymbol{U}^2) - 3.$$

  Diagonalising $\boldsymbol{U}$, we have

  $$\mathrm{tr}(\boldsymbol{U}^2) = \mathrm{tr}(\boldsymbol{QD}^2\boldsymbol{Q}^T) = \mathrm{tr}(\boldsymbol{Q}^T\boldsymbol{QD}^2) = \mathrm{tr}(\boldsymbol{D}^2).$$

The latter expression is the sum of the squares of the principal invariants, and is permutation invariant as required. Inspecting the term $\mathrm{tr}\boldsymbol{E}^2$, we find the only term we need to account for further is $\mathrm{tr}(\boldsymbol{F}^T\boldsymbol{F}\boldsymbol{F}^T\boldsymbol{F})$. Diagonalising once more, we have

$$\mathrm{tr}(\boldsymbol{F}^T\boldsymbol{F}\boldsymbol{F}^T\boldsymbol{F}) = \mathrm{tr}(\boldsymbol{U}^4) = \mathrm{tr}(\boldsymbol{D}^4) = \lambda_1^4 + \lambda_2^4 + \lambda_3^4.$$

Once more, this is a permutation-invariant function of the principal stretches.

- For a **Mooney-Rivlin** material, we note that for the first term,

$$|\boldsymbol{F}|^2 = \mathrm{tr}(\boldsymbol{F}^T\boldsymbol{F}) = \mathrm{tr}(\boldsymbol{U}^2) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2,$$

where λ_i are the principal stretches. Next, we note that

$$|\mathrm{cof}\boldsymbol{F}|^2 = (\det \boldsymbol{F})^2\mathrm{tr}(\boldsymbol{F}^{-1}\boldsymbol{F}^{-T}).$$

Using the polar decomposition, we can write

$$\begin{aligned}\mathrm{tr}(\boldsymbol{F}^{-1}\boldsymbol{F}^{-T}) &= \mathrm{tr}(\boldsymbol{U}^{-1}\boldsymbol{R}^T\boldsymbol{R}\boldsymbol{U}^{-T}) \\ &= \mathrm{tr}(\boldsymbol{U}^{-2}) = \mathrm{tr}(\boldsymbol{D}^{-2}) = \lambda_1^{-2} + \lambda_2^{-2} + \lambda_3^{-2}.\end{aligned}$$

Multiplying by $\det \boldsymbol{U}^2 = \det \boldsymbol{D}^2$, we obtain

$$|\mathrm{cof}\boldsymbol{F}|^2 = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2.$$

Similarly for the final term, we observe that $\det \boldsymbol{F} = \det \boldsymbol{R}\boldsymbol{U} = \det \boldsymbol{U} = \lambda_1\lambda_2\lambda_3$. In each case, we have expressed the terms as a permutation invariant function of the principal stretches, so we are done.

7.7 Linear Elasticity

The equations of elasticity we have derived are challenging to solve exactly due to their inherent nonlinearity which arises from frame indifference. In practice, it often convenient and reasonable to use a linearised form of the equations. Typically, we focus on cases where $\boldsymbol{F}$ stays close to the identity, but any fixed rotation tensor would also be possible.

7.7.1 Linearisation

Suppose that we consider a hyperelastic solid for which $DW(\boldsymbol{I}) = \widetilde{\boldsymbol{P}}(\boldsymbol{I}) = \boldsymbol{O}$, i.e. the reference configuration is stress-free, and moreover assume that the deformation gradient in the material stays close to the identity, i.e.

$$|\boldsymbol{F} - \boldsymbol{I}| \ll 1.$$

Recalling that

$$\boldsymbol{F} = \nabla_{\boldsymbol{x}} \boldsymbol{\varphi} = \boldsymbol{I} + \nabla_{\boldsymbol{x}} \boldsymbol{u},$$

we can Taylor expand the stress response to obtain

$$\begin{aligned}\widetilde{P}_{ij}(\boldsymbol{F}) &= \widetilde{P}_{ij}(\boldsymbol{I}) + \frac{\partial \widetilde{P}_{ij}}{\partial F_{kl}}(\boldsymbol{I})\Big(F_{kl} - \delta_{kl}\Big) + O\big(|\boldsymbol{F} - \boldsymbol{I}|^2\big) \\ &= \frac{\partial \widetilde{P}_{ij}}{\partial F_{kl}}(\boldsymbol{I}) u_{k,l} + O\big(|\boldsymbol{F} - \boldsymbol{I}|^2\big),\end{aligned}$$

where the first term vanishes due to the assumption made above, and in the second term we have used the definition of the displacement gradient. If we drop all quadratic and higher-orer terms, we see that the Piola-Kirchhoff stress can be approximated as simply being a linear function of the displacement gradient. This leads us to define the fourth-order **elasticity tensor** with components

$$\mathbb{C}_{ijkl} := \frac{\partial \widetilde{P}_{ij}}{\partial F_{kl}}(\boldsymbol{I}) = \frac{\partial^2 W}{\partial F_{ij} \partial F_{kl}}(\boldsymbol{I}).$$

A fourth-order tensor is a linear map from second-order tensors to second-order tensors, and we can define components in various ways, for example via the expressions

$$\mathbb{C}_{ijkl} = \boldsymbol{e}_i \cdot \left(\left(\mathbb{C}(\boldsymbol{e}_k \otimes \boldsymbol{e}_l)\right) \boldsymbol{e}_j\right) = (\boldsymbol{e}_i \otimes \boldsymbol{e}_j) : \left(\mathbb{C}(\boldsymbol{e}_k \otimes \boldsymbol{e}_l)\right),$$

where $\{\boldsymbol{e}_i\}$ are the basis vectors in a fixed Cartesian coordinate frame.

A nice feature of this linearisation is that we can recast the elastodynamics and elastostatics problems in terms of the displacement field $\boldsymbol{u}$ alone, since the acceleration and velocity fields are

$$\frac{\partial^2 \boldsymbol{\varphi}}{\partial t^2} = \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \quad \text{and} \quad \frac{\partial \boldsymbol{\varphi}}{\partial t} = \frac{\partial \boldsymbol{u}}{\partial t}.$$

In particular, dropping higher-order terms, we obtain the following linearised version of our elastodynamics and elastostatics problems.

Linear elastodynamics problem

Find $\boldsymbol{u} : B \times [0,T] \to \mathbb{E}^3$ satisfying the following initial-boundary value problem:

$$\begin{aligned}\rho_0 \frac{\partial^2 \boldsymbol{u}}{\partial t^2} &= \nabla_{\boldsymbol{x}} \cdot (\mathbb{C} \nabla_{\boldsymbol{x}} \boldsymbol{u}) + \rho_0 \boldsymbol{b}_m && \text{in } B \times [0,T], \\ \boldsymbol{u}(\cdot, 0) &= \boldsymbol{u}_0 && \text{in } B, \\ \frac{\partial \boldsymbol{u}}{\partial t}(\cdot, 0) &= \boldsymbol{v}_0 && \text{in } B, \\ \boldsymbol{u} &= \boldsymbol{g} - \boldsymbol{x} && \text{on } \Gamma_D \times [0,T], \\ (\mathbb{C} \nabla_{\boldsymbol{x}} \boldsymbol{u})\, \boldsymbol{\nu} &= \boldsymbol{t} && \text{on } \Gamma_N \times [0,T].\end{aligned}$$

- This problem is a linear system of 3 coupled PDEs for the 3 components of $\boldsymbol{u}$ which is second-order in both space and time.
- Note we must shift the exact positions for Dirichlet boundary conditions to displacements, and we replace the stress-strain relation everywhere by the linearised form.
- Typically, solutions of the equations oscillate in time, and we think of this as a **linear hyperbolic** PDE system. Wave-like solutions make physical sense as elastic waves are exactly how sound travels through solid materials. Unlike for the nonlinear theory we have developed above, the principle of superposition does apply to solutions of the linear elastodynamics equations.

Linear elastostatics problem

Find $\boldsymbol{u} : B \to \mathbb{E}^3$ satisfying the boundary value problem:

$$\begin{aligned} \nabla_{\boldsymbol{x}} \cdot (\mathbb{C} \nabla_{\boldsymbol{x}} \boldsymbol{u}) + \rho_0 \boldsymbol{b}_m &= \boldsymbol{0} && \text{in } B \times [0, T], \\ \boldsymbol{u} &= \boldsymbol{g} - \boldsymbol{x} && \text{on } \Gamma_D, \\ (\mathbb{C} \nabla_{\boldsymbol{x}} \boldsymbol{u})\, \boldsymbol{\nu} &= \boldsymbol{t} && \text{on } \Gamma_N. \end{aligned}$$

7.7.2 Symmetries of the elasticity tensor

The elasticity tensor $\mathbb{C}$ has $3^4 = 81$ components; however, we will see that as a consequence of our assumptions, there are actually only 21 independent entries.

First, let's observe that

$$\widetilde{\boldsymbol{S}}(\boldsymbol{F}) = \frac{\widetilde{\boldsymbol{P}}(\boldsymbol{F}) \boldsymbol{F}^T}{\det \boldsymbol{F}}.$$

If we Taylor expand about $\boldsymbol{I}$, then as long as $\widetilde{\boldsymbol{P}}(\boldsymbol{I}) = \boldsymbol{O}$, various terms vanish and we have

$$\widetilde{S}_{ij}(\boldsymbol{F}) = \frac{\partial P_{ij}}{\partial F_{kl}}(\boldsymbol{I}) u_{k,l} + O(|\boldsymbol{F} - \boldsymbol{I}|^2) = \mathbb{C}_{ijkl} u_{k,l} + O(|\boldsymbol{F} - \boldsymbol{I}|^2).$$

The fact that we require symmetry of $\boldsymbol{S}$ to ensure that angular momentum balance holds entails that $\mathbb{C}_{ijkl} = \mathbb{C}_{jikl}$ for any indices i, j, k, l. This is called the *minor symmetry* of the elasticity tensor.

If $\mathbb{C}$ were derived as the second derivative of a strain energy density W, then it must also satisfy the *major symmetry*, i.e.

$$\mathbb{C}_{ijkl} = \frac{\partial^2 W}{\partial F_{ij} \partial F_{kl}} = \frac{\partial^2 W}{\partial F_{kl} \partial F_{ij}} = \mathbb{C}_{klij}.$$

Putting both of these symmetries together, we see that we also have

$$\mathbb{C}_{ijkl} = \mathbb{C}_{ijlk},$$

and as a result, we have

$$\mathbb{C}_{ijkl}u_{k,l} = \tfrac{1}{2}\mathbb{C}_{ijkl}(u_{k,l} + u_{l,k}) = \mathbb{C}_{ijkl}\varepsilon_{kl},$$

where $\boldsymbol{\varepsilon} = \mathrm{sym}(\nabla_{\boldsymbol{x}}\boldsymbol{u})$ is the infinitesimal strain.

If we wish to show there are 21 independent components of the elasticity tensor, we need only note that there are 21 possible choices of independent pairs taken from the collection

$$\{(1,1),(2,2),(3,3),(1,2),(2,3),(3,1)\}.$$

7.7.3 Isotropy

In practice, we do not often need 21 independent coefficients to describe a material due to symmetry. In the isotropic case, we need only 2, which we now discuss.

Recall that an isotropic hyperelastic solid satisfies

$$W(\boldsymbol{Q}_1\boldsymbol{F}\boldsymbol{Q}_2) = W(\boldsymbol{F})$$

for any rotations $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$. Using this relationship with $\boldsymbol{Q}_1 = \boldsymbol{Q}$ and $\boldsymbol{Q}_2 = \boldsymbol{Q}^T$ and applying the chain rule, we can check that

$$\boldsymbol{Q}(\mathbb{C}\boldsymbol{\varepsilon})\boldsymbol{Q}^T = \mathbb{C}(\boldsymbol{Q}\boldsymbol{\varepsilon}\boldsymbol{Q}^T)$$

for any rotation $\boldsymbol{Q}$ and for any tensor $\boldsymbol{\varepsilon}$. $\mathbb{C}$ is a linear map on tensors, and it can be shown that the only way to construct a linear map with the necessary symmetries is to demand that

$$\mathbb{C}\boldsymbol{\varepsilon} = \lambda\,\mathrm{tr}(\boldsymbol{\varepsilon})\boldsymbol{I} + 2\mu\,\mathrm{sym}(\boldsymbol{\varepsilon}),$$

where λ and μ are the only 2 independent coefficients required. These coefficient are called the Lamé parameters, and are not the only possible ways to parametrise isotropic elasticity, as we discuss next.

7.7.4 Other elastic moduli

While the Lamé parameters are one way to parametrise linear elasticity, there are other possible choices. Important examples are:

- The **bulk modulus** κ,
- The **shear modulus** G (this is identical to μ used above),
- The **Young's modulus** E, and
- The **Poisson ratio** ν.

We will discuss each of these parameters in turn.

7.7.4.1 Bulk modulus

Suppose that we take a piece of material and uniformly expand or compress it, and ask what stress is induced in the material. In particular, if we consider a deformation gradient of the form

$$\boldsymbol{F} = (1 + \tfrac{1}{3}\alpha)\boldsymbol{I},$$

then the volume change induced is

$$\det \boldsymbol{F} = (1 + \tfrac{1}{3}\alpha)^3 = 1 + \alpha + O(\alpha^2),$$

so α is approximately the relative change in the volume (at least for deformations where we stay close to the identity, which is the standing assumption we make in linear elasticity). Looking at the corresponding displacement gradient, we have

$$\nabla_{\boldsymbol{x}} \boldsymbol{u} = \boldsymbol{F} - \boldsymbol{I} = \tfrac{1}{3}\alpha \boldsymbol{I},$$

and hence

$$\mathbb{C}\nabla_{\boldsymbol{x}} \boldsymbol{u} = \lambda\alpha\boldsymbol{I} + \tfrac{2}{3}\alpha\mu\boldsymbol{I} = \alpha(\lambda + \tfrac{2}{3}\mu)\boldsymbol{I}.$$

In this case, a spherical state of stress is induced in the material, with a magnitude which is proportional to the fractional change in volume. The bulk modulus is the appropriate coefficient, i.e.

$$\kappa = \lambda + \tfrac{2}{3}\mu.$$

7.7.4.2 Shear modulus

Next we consider small shear deformations, where

$$\boldsymbol{F} = \boldsymbol{I} + \alpha \boldsymbol{e}_1 \otimes \boldsymbol{e}_2.$$

In this case, the displacement gradient is

$$\nabla_{\boldsymbol{x}} \boldsymbol{u} = \alpha \boldsymbol{e}_1 \otimes \boldsymbol{e}_2,$$

and so plugging in, we obtain

$$\mathbb{C}\nabla_{\boldsymbol{x}} \boldsymbol{u} = \mu\alpha(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_1).$$

Here, we see that μ is the coefficient of proportionality between a shear deformation and the shear stress induced, so we define $G = \mu$ to be the shear modulus of the material.

7.7.4.3 Young's modulus and Poisson ratio

A final option is to consider a uniaxial tension experiment in which we induce a deformation of the form

$$\boldsymbol{F} = (1+\alpha)\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + (1-\beta)\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + (1-\beta)\boldsymbol{e}_3 \otimes \boldsymbol{e}_3.$$

Here we think of α as being the fractional increase in length in the $\boldsymbol{e}_1$ direction, and by introducing β, we allow ourselves the flexibility to have a reduction in the width of the sample in the perpendicular directions. In this case,

$$\nabla_{\boldsymbol{x}} \boldsymbol{u} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & -\beta & 0 \\ 0 & 0 & -\beta \end{pmatrix}$$

and hence plugging in, we find

$$\mathbb{C}\nabla_{\boldsymbol{x}} \boldsymbol{u} = \begin{pmatrix} (\lambda+2\mu)\alpha - 2\lambda\beta & 0 & 0 \\ 0 & \lambda\alpha - 2(\lambda+\mu)\beta & 0 \\ 0 & 0 & \lambda\alpha - 2(\lambda+\mu)\beta \end{pmatrix}$$

Note that if the material is stress free on the surfaces of the sample that are not being strained, then the 22 and 33 components of the stress should vanish. To achieve this, we should choose

$$\beta = \frac{\lambda}{2(\lambda+\mu)}\alpha.$$

This special ratio of coefficients is called the Poisson ratio, denoted

$$\nu = \frac{\lambda}{2(\lambda+\mu)},$$

and unlike the other parameters in linear elasticity we have been considering it is dimensionless.

If we pick β in this way, then we find that the stress induced in the 11 component is

$$(\lambda+2\mu)\alpha - 2\lambda\beta = \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}\alpha,$$

and the coefficient of proportionality is the Young's modulus, denoted

$$E = \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}.$$

7.7.4.4 Discussion

In the cases above, we showed that all of the elastic moduli we have described can be obtained from the Lamé parameters, but given any pair of these parameters, we can transform to any other pair. For example, in terms of Young's modulus E and the Poisson ratio ν, the isotropic elasticity tensor becomes

$$\mathbb{C}\boldsymbol{\varepsilon} = \frac{E\nu}{(1+\nu)(1-2\nu)}\mathrm{tr}(\boldsymbol{\varepsilon})\boldsymbol{I} + \frac{E}{1+\nu}\mathrm{sym}(\boldsymbol{\varepsilon}).$$

You can see in this case that something bad happens if ν tends to $\frac{1}{2}$, since the coefficient in the first term blows up. This corresponds to the case where the material is *incompressible*, i.e. where volume change is not permitted. We can see this in an alternative expression for the bulk modulus:

$$\kappa = \frac{E}{3(1-2\nu)}.$$

Unlike for fluids, in the vast majority of real elastic materials, incompressibility is not normally a feature of elastic solids, and most materials have a Poisson ration of around 0.2–0.4. The other elasticity coefficients all tend to be positive; for some example values for real materials at room temperature, see Table 7.1. In terms of units, with exception of the Poisson ratio, all of the moduli discussed are expressed in units of energy per unit volume; in SI units, the natural units are $\mathrm{kg\,m^{-1}\,s^{-2}}$. Energy per unit volume is equivalent to force per unit area, and the unit most often used is therefore the Pascal, or equivalently Newtons per metre squared. Elastic moduli for real materials tend to be in the range of 10^6–10^{12} Pascals, so it is common to use Megapascals (MPa) and Gigapascals (GPa) to express the moduli in practice.

Table 7.1: Approximate elastic moduli for some materials at room temperature.

Material	κ (GPa)	G (GPa)	ν	E (GPa)
Aluminium	75	25	0.36	70
Copper	140	47	0.35	125
Diamond	540	460	0.2	1100
Lead	41	5	0.44	14
Silicone Rubber	2	0.01	0.48	0.05
Stainless Steel	140	78	0.27	195